Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2019-289-AC1, 2020

Geoscientific Model Development Discussions

Interactive comment on "The Sailor diagram. An extension of Taylor's diagram to two-dimensional vector data" *by* Jon Sáenz et al.

Jon Sáenz et al.

jon.saenz@ehu.es

Received and published: 13 January 2020

Our reply to the original comments by the reviewer is written in italics.

The paper addresses a relevant and often appearing issue: comparing vector quantities. It reviews the different approaches developed so far, giving appropriate credit to those, and adds the idea of a novel graphics presentation as "sailor diagram". This is potentially a useful tool for a vast range of applications, several examples are chosen from different fields for illustration. The deviation of method is clearly outlined and valid, reproducibility is excellent. The title is excellently chosen, abstract is concise and the term "sailor diagram" justified in the paper. Language and maths are clear.

Thank you for your kind comments.

Figures are less clear. Grey squares in all figures are hard to spot (and important).

We appreciate this comment by Reviewer 1 and we agree with his/her appreciation. We are reworking the package so that in a new version these grey squares will be larger and they will be on top of all the lines.

Although it is nice that the figures relate to real world examples, for introducing the concept it would be helpful to have figures showing clearly the benefits and limitations of the sailor diagram. Figure 2a is a very good one. The others are not easy to interpret, i.e., helping less to understand the concept. Applicability and interpretation, and general presentation would benefit from clearer examples. Given that the graphics are a central idea of the paper, following revisions are suggested, with the intention to improve understanding and uptake of the Sailor diagram for other researchers: 1) List and number the features of the Sailor diagram clearly, eg., like i) size of ellipse depicting covariance ii) direction of ellipse indicating error main axis iii) squares indicating bias for both components iv) options for scaling as indicated in Fig. 12) give one (possibly synthetic) example figure illustrating clearly each feature (e.g.,datasets disagreeing on i) and agreeing on ii) and iii). For iv) note what scaling comes with which advantage / disadvantage.

We again appreciate this constructive comment by Reviewer 1. We find it a very useful suggestion. We have prepared a new figure (it will be Figure 1 in the revised version of the manuscript). In it, we have taken a real-world example (one year -2018- of hourly wind data in a grid point in front of Los Angeles from ERA5, 38 degrees North and 124 degrees West). These data are considered as the reference dataset. From this reference dataset, new synthetic datasets with artificial sources of error are produced. In the first case (MOD1), a constant bias is added to the reference dataset. Thus, the orientation and lengths of MOD1 are the same as the ones from the reference dataset, but the bias is different from zero. This addresses source of error iii, as indicated by the reviewer.

A second model (MOD2) is generated by rotating 30° counter-clockwise the wind fields. This introduces an artificial rotation of the EOFs without affecting the fractions of variance of the principal axes. The rotation changes the position of the average wind and it also introduces some bias. This second synthetic dataset addresses sources of error iii (bias) and ii (rotation), as suggested by reviewer.

As a completely artificial example, we have also generated model MOD3, which only involves a random resampling in time of wind measurements. Thus, the bias source of error must be zero and the EOFs must also be the same, but due to the fact that the correlation coefficient between datasets now is not the same, in this case the sailor diagram shows a perfect graphical agreement, but the legend shows that the RMSE is not zero, so that this part is due to the lower correlation of the wind. This case addresses the lack of temporal correlation of the data despite perfect agreement in length of major axes of the ellipse (case i suggested by reviewer), angle (case ii) and no bias (case iii).

The last synthetic case (MOD4) is produced by multiplying the original data by a scale factor (2), so that the axes of the ellipses are changed (case i suggested by reviewer), but without any rotation of the axes and a different bias due to the scaling factor. We have produced versions centered and uncentered of the Sailor plot for these synthetic datasets and we intend to use this as the first Figure of the paper so as to address this interesting point raised by the reviewer. We agree with him/her that this addition will improve the interpretability of the paper. We enclose the new figure below.

Figure 1: Sailor diagrams representing each of the synthetic datasets including the artificial sources of errors described above in the text. Two versions are included, one uncentered (left) and another one where all the ellipses are centered over the reference value (right).

These considerations are clearly reflected in the values of the table exported by the package (call to function SailoR.Table), as shown below. This table will also be added

to the last version of the paper as Table 1.

3) Explain the underlying assumptions and the limitations (i.e., what could go wrong with the interpretation). For instance, in Fig 1, the almost orthogonal major axes – are they caused by the two EOF being approx. same size and some noise deciding on whether the correct EOFS are aligned in the graphics?

The reviewer is right that in this case, the standard deviation of the reference dataset seems the same along both EOFs, despite it is not exactly the same (4.3 m/s versus 4.2 m/s). For the rest of models, it is not so close (4.21/3.62, 4.75/4.29, 3.68/3.47, 4.23/3.99). Therefore, we are sure the results are not purely due to degeneracy of the eigenvalues of the covariance matrix in this case. The reviewer raises a valid point in this comment, and we will modify the package to check this case, raising a warning for the user in the next release of the package. However, in our experience, this will be more the exception than the norm for realistic vector time series.

Are Fig. 1 (major and minor axis) thus showing a possible pitfall of interpretation of the Sailor diagram? What other limitations and possible pitfalls do exist?

Besides the fact that the Sailor diagram presents in graphical terms the bias and the EOFs, the legend at the lower-right corner of the diagram shows also the RMSE of the time series. Even though the EOFs were exactly the same (see MOD3 in Figure 1 in this response), the RMSE legend, which presents in an aggregated index the amount of the error due both to the bias and the EOFs would still point to the dataset showing the smallest error. We will stress this point in the final version of the manuscript.

4) Remove figures not adding information. The whole section 2 (data description) is not necessary for the understanding of the principle of the Sailor diagram and can be shortened significantly, just serving the understanding of real world examples.

We tried to thoroughly describe all the datasets used to ease replicability, but we will shorten this section in the last version of the manuscript following the suggestion by

the reviewer.

It is not clear for what Fig. 3a is needed – and its explanation is full of abbreviations (check "per" and "pers"). Somebody not familiar with these particular data sets cannot extract sensible information from section 4.4.

We intended to show that the SailoR diagram, besides playing a role in the evaluation of typical vector quantities such as wind or current can also be used for additional vector quantities, such as wave energy flux. In this particular case, we find interesting that the EOFs from techniques which involve the resampling of the original dataset such as persistence and analogs ("pers", "analo" and "analorf") show the smallest rotation of the EOFs. As we mentioned above in this reply, the information about the RMSE in the bottom-right legend shows that the combination of bias and EOFs leads to smaller RMSE in the case of the "analo" dataset. We would prefer to keep this figure, because we think it is interesting, but we agree with the reviewer that it needs a better explanation along the lines written in this reply. However, we could also remove it, in case the reviewer still thinks it must be removed.

5) Figure 4 (right) needs clarification. It is impossible to relate the color codes to the 2 clusters of ellipses. Why are there exactly 2 clusters of ellipses? Furthermore, it is unclear what the centres should denote. Why are there 2 grey ellipses in the upper cluster? It is unclear what is intended to show. I cannot draw conclusions from this figure. Either clarify or remove this figure.

We understand the comments by the reviewer. In this case, again, we find the Figure should be kept in the final version of the manuscript, since it illustrates one of the problems of one of the ways that the Sailor diagram can be used to assess ensembles.

Regarding color codes, we can change the color palette, but it will be of no help, and the reason is that the ellipses are centered almost at the same point (but not exactly at the same point) for each of the members of the ensemble (clusters defined by realizations of the same model). In this case, the figure shows that the bias of each of the

realizations by each of the models is quite the same (but not exactly the same). Since this figure is an uncentered version of the Sailor diagram, every realization (that are being taken as independent simulations in this case) implies that there is a grey ellipse centered at every mean value. The intra-model bias is much smaller than the inter-model one, and that is the reason that there are two clusters of ellipses (one cluster involving all the realizations from MIROC and another cluster involving all the realizations by IPSL). The centre of each ellipse denotes the mean of every realization, and since they are very close, it seems that they are the same, but they are not. We add here a centered version of the plot which could also be used. However, we, as authors, would prefer the original one since, in the case of the Figure 2 below, the individual realizations can be barely identified. Figure 2: Centered version of Figure 4 (right) of the manuscript.

6) It is commendable you provided an R package SailoR. It would be good to state clearly in section 3.7 which figure is included in the manual (instead "some of these plots")

We will clarify this in the final version of the manuscript and the new version of the manual, when we upload the modified version of the package considering all the suggestions made by reviewers.

7) For better visibility, consider plotting the squares on top of the lines, to change grey to black, and to enlarge the size of the squares.