

Geosci. Model Dev. Discuss., referee comment RC1
<https://doi.org/10.5194/gmd-2022-64-RC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on gmd-2022-64

Anonymous Referee #1

Referee comment on "Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not" by Timothy O. Hodson, Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2022-64-RC1>, 2022

Review comments for "Root mean square error (RMSE) or mean absolute error (MAE): when to use them or not"

General:

The manuscript provides an interesting discussion on the choice of RMSE and MAE from the likelihood perspective. When neither RMSE nor MAE is optimal, the author lists several options: refining the model, transforming the data, using robust statistics, and constructing a better likelihood. However, the suggested options are not quite helpful for users of the standard statistical metrics. For instance, the likelihood-based inference is deemed the most versatile by the author, but most readers are not expected to explore this option in their future applications. If possible, providing some examples with illustrations will be very helpful.

Overall the paper is well written and quite informative, although there are some questionable remarks and inaccurate statements as detailed below.

Major points:

The abstract mostly states the motivation of the paper and a not well-substantiated opinion. The reader might need to rewrite the abstract to better reflect the actual content of this paper.

In Section 6 of the paper, "Why not use both RMSE and MAE?", the author argues against using both RMSE and MAE. It is stated that presenting both metrics is "unnecessary and potentially confusing" "If the evidence strongly supports one over the other". In most applications, such evidences to strongly support one metric over the other are not easily attainable. While decomposing one metric into several independent components suggested by the author is viable, using multiple metrics is still a practical way to avoid mistaken conclusions caused by merely relying on one metric.

Specific points:

Abstract Lines 3-4, "Some of this confusion arises from a recent debate between Willmott and Matsuura (2005) and Chai and Draxler (2014), in which either side presents their arguments for one metric over the other. Neither:

While Chai and Draxler (2014) argued against favoring MAE over RMSE by Willmott and Matsuura (2005), they did not favor RMSE over MAE. That is clearly stated in the abstract, as quoted below.

"The RMSE is more appropriate to represent model performance than the MAE when the error distribution is expected to be Gaussian. In addition, we show that the RMSE satisfies the triangle inequality requirement for a distance metric, whereas Willmott et al. (2009) indicated that the sums-of-squares-based statistics do not satisfy this rule. In the end, we discussed some circumstances where using the RMSE will be more beneficial. However, we do not contend that the RMSE is superior over the MAE. Instead, a combination of metrics, including but certainly not limited to RMSEs and MAEs, are often required to assess model performance."

Lines 33-34: "That recent shift may explain why Willmott and Matsuura (2005) and Chai and Draxler (2014) were unaware of the historical justification for MAE and RMSE; neither were they the first to overlook it":

The author might be entitled to have such a judgement of others' unawareness or overlook, but it is better to avoid such opinions in a scientific paper.

Equation 10: A close parenthesis is missing

Line 180: Please add a comma after "Laplace".

Line 209. "... though wrongly suggest that MAE only applies to uniformly distributed errors":

It is not accurate. Although Chai and Draxler (2014) gave one example with "uniformly distributed errors" to show that the MAE would be a good metric for such cases, it was not suggested that they are the ONLY cases where MAE would be appropriate. This statement is a misinterpretation of the paper.