Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2020-51-RC1, 2020

Geoscientific Model Development Discussions
Open Access

Interactive comment on "From R-squared to coefficient of model accuracy for assessing "goodness-of-fits"" *by* Charles Onyutha

Anonymous Referee #1

Received and published: 7 September 2020

Review of: From R-squared to coefficient of model accuracy for assessing "goodness-of-fits" by Charles Onyutha

I have examined part of this paper and found some of it to be well written and interesting. I learned a few things from it, which in my view puts the contribution in the above average category. Nevertheless, I find that it is too long for what it is, and does not make fair comparisons with other metrics. I offer several suggestions as to how it could be improved.

Abstract: the abstract is ineffective as it does not compactly present the major findings. Much of this material reads like an "introduction", starting with the subjective first sentence. My suggested rewrite of this sentence is:

A new measure of "goodness of fit" eliminates several of the well known shortcomings of the widely used correlation coefficient Rˆ2, including its insensitivity to bias when models are compared to measurements.

The abstract should then go on to describe the new metric, list its specific properties, then list its similarities, differences, advantages and disadvantages as compared to other metrics. Each sentence should be compact and deliver new and interesting information. No fluff or opinion is appropriate.

Line 45 ff This PP is very interesting, though I am concerned that misinterpretations may be included. Fundamentally, the reason that the values of R or Rˆ2 do not depend on whether y is plotted against x, or vice versa, is evident in the definition provided by eq. 1; specifically, there is no difference in how x and y are treated in that mathematical definition, so that x and y are interchangeable and mathematically symmetric. That is, if all of the x's were replaced by y's, and all the y's were replaced by x's, the equation would look the same. Thus, given any two column table of data, either column could be defined as x and the other as y, and the result returned by eq 1 would be the same.

Equation 4 and line 75ff. I have confirmed that Eq 4 is correct, but the claim that the "deviations of X and Y from their means to obtain γ are assessed independently" is not. Specifically, the author has mistakenly concealed that dependence in his substitution for "m". The reader can refer back to eq 2 to see that m depends on the PRODUCT of the x and y deviations, and eq 4 depends on m; this conclusively refutes the author's statement.

I don't have time to plough through the author's derivation, but it is clear that computing this would be very difficult compared to many of the simple, single-formula metrics that are currently available to compare models and measurements.

I am unimpressed with the choice of the dataset used to illustrate application of the model. Data that are widely and readily accessible would be better, for example, data from a government website would be better. That original data set could be manipulated in various simple ways to compare the manipulation, now representing the “model”, with the real data, and different metrics applied and compared.

The extended narrative of the performance of different metrics is ineffective. A table whose first column lists the various properties, with additional columns to the right providing values for the particular metric that heads that column, would be far better and would allow direct and simple comparisons of the properties of each metric with all others. Is the metric constrained to range from 0 to 1? Are effects of bias easily excluded? How many different formulae are needed to compute the metric? Ease of computation: if not a trivial calculation, are automated programs readily and widely available, as they are for Rˆ2? Does the metric have real physical significance? Is it widely used? Etc.

My conclusion is that the author has more work to do. A shorter, compact paper would be more effective.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2020-51, 2020.