Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2020-51-AC3, 2020

Geoscientific Model Development Discussions
Open Access
GMDD
Interactive comment

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

Charles Onyutha

conyutha@kyu.ac.ug

Received and published: 24 October 2020

GENERAL

The author is thankful to Barry Croke for his meticulous comments which were all constructive in enhancing both quality and content of the paper. Furthermore, the author is grateful for the recognition that there is merit in the coefficient of model accuracy (CMA) being proposed.

MAJOR COMMENTS

The author agrees with Barry that there are various versions of R-squared. Kvalseth (1985) in his paper (The American Statistician, 39:4, 279-285, DOI: 10.1080/00031305.1985.10479448) elaborated on nine versions of R-squared (see

Eqs (1)-(9) and (11) of cited paper). It is important that clarifications are required on the existing variants of R-squared and these will be made in the revised manuscript.

In Geoscience (take for instance, the field of hydrology), the versions of R-squared which are increasingly being used are (1) the square of the Pearson product moment correlation coefficient, and (2) coefficient of determination or the Nash-Sutcliffe Efficiency NSE (Nash and Sutcliffe 1970, J. Hydrol., 10, 282–290, doi:10.1016/0022-1694(70)90255-6). These two R-squared versions (1) and (2) as correctly noted by Barry are not the same based on the information from Bardsley (2013) published in Hydrol. Proc. DOI: 10.1002/hyp.9914. The author also realizes that clarification, for instance, on the two common versions of R-squared was lacking in the discussion paper and this will be worked on during revision of the manuscript. Furthermore, the author agrees that information in the manuscript is somewhat mixed up regarding which version is being treated. Although the version of R-squared based on linear regression is rampantly used, the revised manuscript will focus on the NSE as suggested. This follows the conclusion of Kvalseth (1985) that generally the best form of R-squared appears to be what hydrologists refer to as the NSE. To make the need of focusing on NSE emphatic, the title will slightly be changed from " From R-squared to coefficient of model accuracy for assessing "goodness-of-fits"" to " From R-squared or Nash Sutcliffe Efficiency to coefficient of model accuracy for assessing "goodness-of-fits"". Furthermore, discussion regarding the version of R-squared based on linear regression will be removed from the revised manuscript. The criticism that R-squared still remains the same when X and Y are swapped will also be removed from the revised manuscript. Generally, due to the need to put NSE central to the manuscript, several sentences in the abstract, conclusion, discussion, and other sections will be removed or revisited.

Results of comparison of NSE, CMA, Taylor Skill Score TSS (Taylor (2001), J. Geophys. Res., 106, 7183−7192, DOI: 10.1029/2000JD900719), and the version of R-squared recommended by Kvalseth (1985) (or Eq. 11 of the paper by Kvalseth (1985)) can be seen in Figures 1 and 2 of this reply to Barry Croke's comments. The need to consider

TSS in comparison CMA with existing metrics was recommended by Ulrich Schumann in his short comment on this discussion paper (see Geosci. Model Dev. Discuss., DOI: 10.5194/gmd-2020-51-SC1). Root Mean Squared Error (RMSE) was not considered for comparison because it is deemed comparable to NSE since both of them are based on the sum of squared residuals. Figure 1 shows comparison of values of the various "goodness-of-fit" metrics NSE, CMA, TSS, and IoA. Figure 2 comprises comparison of observed and modeled flow based on calibration of two models including the Hydrological Model focusing on Sub-flows' Variation (HMSV) and Nedbør-Afstrømnings-Model (NAM). Generally, the results show the acceptability of the CMA being introduced. The reason behind the performance of CMA can be explain in terms of a careful derivation of the proposed CMA's formula. Based on comments from other reviewer's, CMA's formula from the discussion paper requires some modification and this will be done in the revised manuscript as described next.

For CMA to offer an all-encompassing solution to the shortcomings of R-squared, it has three components including (i) absolute value of rank-based coefficient of correlation (f) between observed (X) and modeled (Y) series, (ii) difference in the variances of X and Y, and (iii) bias of the mean of Y from the mean of X.

The first CMA part comprising the term f quantifies co-variability of X with Y. Values of f ranges from zero to 1. When X and Y are perfectly correlated, f takes the value of one. When there is no correlation between X and Y, f yields a value of zero.

In the second part, we obtain another term Wg in terms of the amounts by which X and Y vary from a common baseline (Cb). This Cb is taken to be three times the mean of observed series (Xm); in other words, Cb=3Xm. In Cb, Xm is multiplied by 3 not 2 or 1 due to the need to maintain a reasonableness of CMA. Values of Cb greater (less) than 3 makes CMA conservative (exaggerated). Finally, Wg is given by the ratio of w1 to w2 where where w1= sum of the minimum values of (X-Cb)ˆ2 and (Y-Cb)ˆ2, and w2= sum of the maximum values of (X-Cb)ˆ2 and (Y-Cb)ˆ2. When deviations of the values of X from Cb are the same as those for Y, Wg yields a value of one. It is important to note that, here, Wg can have a value of one regardless of how far apart Xm is from mean of Y (or Ym). Furthermore, Wg can have a value regardless of whether the X and Y are correlated or not. This means, CMA should take into account how biased Ym is from Xm. This leads us to the third term of CMA denoted by Vm.

The third part of CMA is such that Vm = 1 if Xm = Ym otherwise Vm is the ratio of of u1 to u2 where u1 is the squared minimum value among Xm and Ym, and u2 is the squared maximum value among Xm and Ym.

In summary,

CMA= f $\times$ Vm $\times$Wg ....(1)

CMA ranges from 0 to 1. When CMA is zero, it indicates that the model is not better than the comparison baseline (or mean of observed series). CMA value of one indicates that there are no model errors. In the case when CMA is one, it means (i) X and Y are perfectly correlated, (ii) the deviations of X and Y from Cb are the same, and (iii) Xm is equal to Ym.

Mathematical expressions involved in the derivation of CMA will be clearly presented in the revised manuscript. The reviewer noted that there are a few typos in the paper. The author agrees with the reviewer. Corrections will be made in the revised manuscript.

MINOR COMMENTS

1. Line 64: The reviewer noted that the use of both upper and lower letters in representing series and its individual values is confusing. The author agrees with the reviewer on this. One style of representation or notation for series and its individual values will consistently be adopted as advised by the reviewer.

2. Line 105-109: Regarding penalty, the reviewer wondered if we are making use of x=observed(x)-mean of observed and y=modeled-mean of observed (x). However, we are making use of observed (x) and modeled (y).

3. Line 113: Regarding sub-scripts in equations 9 and 10, we are determining the deviation of x and h from the comparison baseline (Cb). For instance, min(xi,hi) means whichever of the ith values of x and h is less than the other. On the other hand, max(xi,hi) means whichever of the ith values of x and h is greater than the other. So, max(xi,hi) is not the same as max(xi, hi). These components in the revised equation for CMA are in terms of w1 and w2 as explained shortly before. The number of terms in the revised CMA formula were reduced for clarity.

4. Line 116: Since we are characterizing variance, we make use of squared deviations. Furthermore, effect of considering absolute and squared values can influence CMA. For instance if we have a=0.2 and b=0.4, it means aˆ2=0.04, bˆ2=0.16. Dividing these two terms a and b we obtain a/b=0.5 and aˆ2/bˆ2=0.25. In other words, the use of square values instead of absolute values is to balance the contribution of the term Wg (see Eq. 1) in CMA.

5. Line 152: The author agrees that including the observed value as a denominator in the formula of MAB (taken to mean model average bias) can be problematic. In the revised manuscript, MAB will be left out so as to allow comparison of "goodness-of-fits" which do not have units.

6. In the revised manuscript, "poorly fit model" will be changed to "poorly fitted model".

7. Line 176: In the revised manuscript "tend o zero" will be changed to "tend to zero".

8. Line 190: The author agrees with the reviewer on the fact that hydrological papers that refer to R-squared or the coefficient of determination are using NSE rather than the R-squared for a linear regression. For clarity in the revised manuscript, focus will be given to NSE instead of R-squared based on linear regression.

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

GMDD

Interactive comment

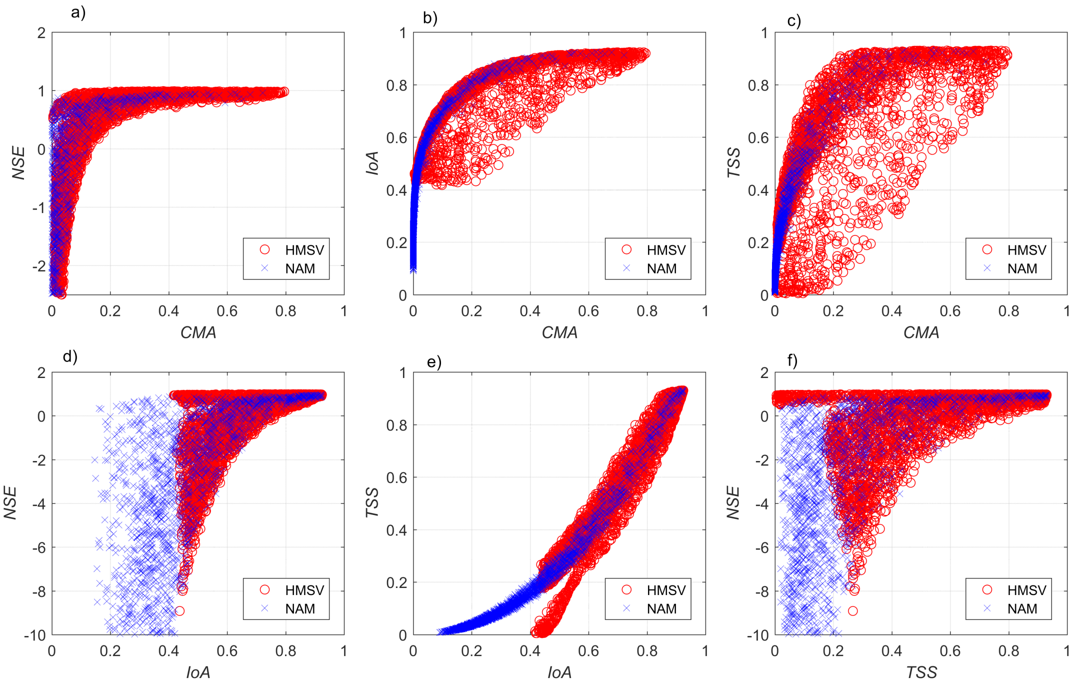

Fig. 1. Comparison of various "goodness-of-fits" including a)-b) CMA, IoA, TSS, NSE

Printer-friendly version

Discussion paper

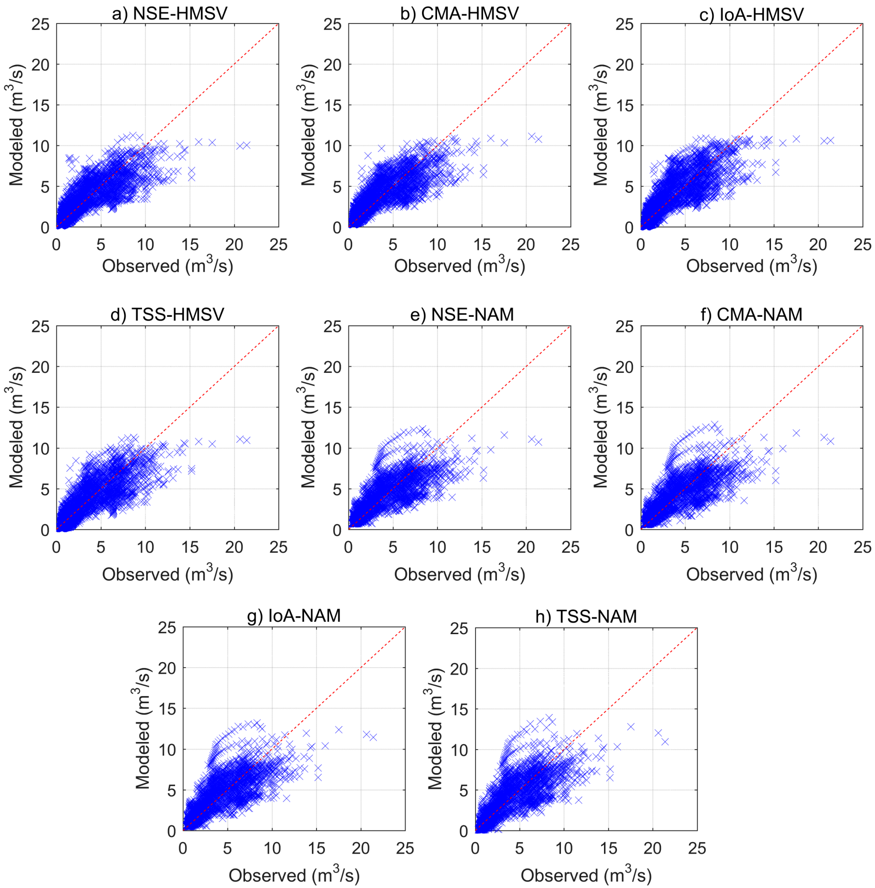

Fig. 2. Comparison of observed and modeled flow from a)-d) HMSV and e)-h) NAM