

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: 5 July 2020

GENERAL

The author is grateful to Ulrich Schumann for acknowledging that the paper is interesting. The author deemed the comments generated by Schumann constructive for improving the coefficient of model accuracy (CMA) being introduced. Four comments that required response are below.

COMMENT 1

While working on contrail cirrus modeling to extend his previous research studies (Schumann 2012; Schumann and Graf 2013; Schumann and Heymsfield 2017), Schumann noted that high values of coefficient of model accuracy (CMA) are only obtainable

C1

when the model is nearly perfect in representing the observations. He thinks achieving high values of CMA becomes a difficult task when observations contain random errors or in the case where model-observation agreement “goodness” is sensitive to small shifts in small-scale structures. Schumann remarked that it was demanding for him to find an optimum criterion for “goodness” while applying CMA.

REPLY

The author agrees with the remarks from Ulrich Schumann that high values of the version of CMA presented in the discussion paper can be obtained when the model is nearly effective. This may make it demanding to fulfill the criterion set to determine acceptability of model results.

This problem was solved as follows. CMA is now expressed as a function of (i) non-parametric normalized linear regression coefficient (trend slope), (ii) difference between the variances of observed and modeled data, and (iii) difference between the means of observed and modeled data.

In determining the measures of co-variation, the comparison baseline was changed from $2 \cdot A_x$ to $3 \cdot A_x$ where A_x is the mean of the observed data. Stepwise derivation of CMA was revised and can be found in Eqs. (A1)-(A10) (see Figures 1-2) of this document. In the revised manuscript, Eqs. (5)-(14) of the discussion paper will be replaced with Eqs. (A1)-(A10) (see Figures 1 and 2 of this reply). The MATLAB and R codes for computing MCA were revised and can be found provided as supplementary materials to this reply. Furthermore, the MATLAB and R codes in the discussion paper will be revised accordingly during the revision of the manuscript.

Conclusively, it was found that high values of CMA can easily be attained with the revised formula.

COMMENT 2

Schumann suggested that comparisons made in this paper should comprise the Taylor

C2

skill score, as given in Eq. (4) of Taylor (2001).

REPLY

Taylor skill score (TSS) as given in Eq. (4) of Taylor (2001) was included in the comparison of “goodness-of-fits”. A part from TSS and CMA, other “goodness-of-fits” included the Nash Sutcliffe Efficiency NSE (Nash and Sutcliffe, 1970), Index of Agreement IOA (Willmott, 1981), Root Mean Squared Error (RMSE), and Model Average Bias (MAB). NSE, RMSE, MAB and IOA can be found in Eqs. (15) to (18) of the discussion paper.

Two rainfall-runoff models including NAM and the Hydrological Model focusing on Sub-flows’ Variation (HMSV) of Onyutha (2019) were calibrated using a number of objective functions based on the Generalized Likelihood Uncertainty Estimation GLUE (Beven and Binley, 2001). Apart from CMA (Eq. A9), other objective functions included the Nash Sutcliffe Efficiency NSE (Nash and Sutcliffe, 1970), Index of Agreement IOA (Willmott, 1981), Root Mean Squared Error (RMSE), Model Average Bias (MAB), and TSS (Eq. A11). NSE, RMSE, MAB and IOA can be found in Eqs. (15) to (18) of the discussion paper.

For a selected objective function, each hydrological model was calibrated 5000 times using GLUE strategy. The values of the objective functions were graphically compared. The best modeled series were those obtained using the set of parameters which yielded the best value of the objective function. The best set of parameters was obtained with the maximum values of CMA, TSS, R-squared, and IOA while RMSE and MAB were required to be at their minimum.

Revised results of the “goodness-of-fits” can be found in Figures 3-6 in this document. The detail description of Figures 3-6 will also be included in the revised manuscript. Changes were also made to Tables A1-A4 in the discussion paper but will be included in the revised manuscript.

Specifically, it was found that for a given CMA, values of TSS were generally larger. In

C3

other words, TSS gets closer the maximum value of 1 faster than the CMA.

COMMENT 3

Schumann noted that CMA makes use of the number of times a data point $x(i)$ appears in observed series x , as seen in line 497 of the discussion paper. He remarked that occurrences of data in reality may be characterized by small round-off errors so that x is nearly equal to a set of values in the observations. The question to answer was on how the situation of such a near-equality can be accounted for.

REPLY

Generally, to minimize the effects of round-off errors, the floating point precision can be increased from float to double though this could require large computational resource.

Direct application of arithmetic operators (such as subtraction and addition) to values which are nearly equal can lead to huge round-off errors. This can be a problem for any “goodness-of-fits” not only the CMA. The computer codes provided are purposefully basic to depict the stepwise procedure adopted in the derivation of CMA. However, these codes can provide starting points regarding improvement of the CMA computation procedure to take into account any possible round-off errors (if any). Such an improvement could be in a way to enhance precision without the requirement of large computational resources. Furthermore, dealing with technicality required to minimize round-off errors in computation is deemed to be an open problem in scientific computing. This requires answering questions regarding accuracy, efficiency, precision, robustness, suitable algorithms or computer programs, data structures, and computing architectures.

Instead of using $e(i)$ which represents how many times a data point $x(i)$ appears in observed series x , we can use $v(i)-u(i)$ as expressed in Eqs. (A6)-(A7) (see Figure 2 of this document) MATLAB or R codes included as supplementary materials to this reply.

COMMENT 4

C4

The last comment was that all abbreviations used in the text should be defined.

REPLY

The author recognizes that it was by mistake that abbreviations such as IOA were not defined in the discussion paper by mistake. During revision of the paper, all abbreviations will be defined.

REFERENCES

- Beven, K. J., and Binley, A. M.: The future role of distributed models: Model calibration and predictive uncertainty, *Hydrol. Process.*, 6, 279–298, doi:10.1002/hyp.3360060305, 1992.
- DHI.: Reference manual, DHI water & environment, in: MIKE11 – A modeling system for rivers and channels, Hørsholm, Denmark: DHI water & environment, pp. 278–325, 2007.
- Madsen, H.: Automatic calibration of a conceptual rainfall–runoff model using multiple objectives, *J. Hydrol.*, 235(3–4), 276–288, doi:10.1016/S0022-1694(00)00279-1, 2000.
- Nash, J. E., and Sutcliffe, J. V. River flow forecasting through conceptual models part I – a discussion of principles, *J. Hydrol.*, 10, 282–290, doi:10.1016/0022-1694(70)90255-6, 1970.
- Nielsen, S. A. and Hansen, E.: Numerical simulation of the rainfall-runoff process on a daily basis, *Nordic Hydrol.*, 4(3), 171–190, 1973.
- Onyutha, C.: Hydrological model supported by a step-wise calibration against sub-flows and validation of extreme flow events. *Water*, 11, 244, doi:10.3390/w11020244, 2019.
- Schumann, U.: A contrail cirrus prediction model. *Geosci. Model Dev.*, 5, 543–580, doi: 10.5194/gmd-5-543-2012, 2012.

C5

- Schumann, U., and Graf, K.: Aviation-induced cirrus and radiation changes at diurnal timescales. *J. Geophys. Res.*, 118, 2404–2421, doi: 10.1002/jgrd.50184, 2013.
- Schumann, U., and Heymsfield, A.: On the lifecycle of individual contrails and contrail cirrus. *Meteor. Monogr.*, 58, 3.1–3.24, doi: 10.1175/AMSMONOGRAPHIS-D-16-0005.1, 2017.
- Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, 106, 7183–7192, doi: 10.1029/2000JD900719, 2001.
- Willmott, C. J., Robeson, S. M. and Matsuura, K.: A refined index of model performance, *Int. J. Climatol.*, 32, 2088–2094, doi: 10.1002/joc.2419, 2012.

Please also note the supplement to this comment:

<https://gmd.copernicus.org/preprints/gmd-2020-51/gmd-2020-51-AC1-supplement.zip>

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

C6

Consider X and Y as the observed and modeled data sets, respectively. Let H be the series resulting from application of penalties to Y . Let us take \bar{x} and \bar{h} as the means of the x_i 's and h_i 's, respectively. The values of h can be derived based on Eq. (8) of the discussion paper.

Adjustment was made to the comparison baseline ξ used in Eqs. (8) and (9) of the discussion paper. Although CMA comprised $\xi = 2\bar{x}$ in the earlier version of CMA, ξ was revised to become $3\bar{x}$. Setting $\xi = 3\bar{x}$ makes CMA to reach its maximum value of 1 faster than when $\xi = 2\bar{x}$ is used. Thus, for $1 \leq i \leq n$, $\omega_{x,i} = \omega_{h,i} = 0$ if $h_i = x_i = \xi$, otherwise

$$\omega_{x,i} = \min\left\{(h_i - \xi)^2, (x_i - \xi)^2\right\} \quad (A1)$$

$$\omega_{h,i} = \max\left\{(h_i - \xi)^2, (x_i - \xi)^2\right\} \quad (A2)$$

where \min and \max respectively denote the minimum and maximum of any two values. Normally errors of opposite signs can cancel each other during their summation. This was avoided in Eqs. (A1)-(A2) by squaring of the terms in brackets on the right hand side. The measure β of deviations of the data points from ξ can be given by

$$\beta = \begin{cases} 0 & \text{if } \sum_{i=1}^n h_i = 0 \text{ or } \sum_{i=1}^n \omega_{h,i} = 0 \\ \sum_{i=1}^n \omega_{x,i} \times \left(\sum_{i=1}^n \omega_{h,i}\right)^{-1}, & \text{otherwise} \end{cases} \quad (A3)$$

The values of β range from zero to one. An ideal model yields $\beta = 1$. When $\beta = 0$, it means the outputs of the model can simply be represented by the mean of observed data. The term β gives insight about the co-variation of X and Y . Apart from quantifying deviations of the data points from the comparison baseline in terms of β , we could obtain another expression (λ) as a function of the standard deviations of X and Y denoted by σ_x and σ_y , respectively such that

$$\lambda = \begin{cases} 0 & \text{if } \sigma_x = \sigma_y = 0 \\ \min(\sigma_x, \sigma_y) \times \left(\max(\sigma_x, \sigma_y)\right)^{-1}, & \text{otherwise} \end{cases} \quad (A4)$$

The values of λ vary from zero to one. For the case when the variances of both X and Y are the same, we get $\lambda = 1$. However, when $\lambda = 0$ it means either $\sigma_x = 0$ and $\sigma_y \neq 0$ or $\sigma_x \neq 0$ and $\sigma_y = 0$.

We need another term θ as a function of \bar{x} and \bar{h} such that

$$\theta = \begin{cases} 1 & \text{if } \bar{h} = \bar{x} = 0 \\ \min\{\bar{h}^2, \bar{x}^2\} \times \left(\max\{\bar{h}^2, \bar{x}^2\}\right)^{-1}, & \text{otherwise} \end{cases} \quad (A5)$$

The term θ varies from zero to one and takes care of the differences in the means of X and Y . When, the means of the observed and modeled series are the same, $\theta = 1$. However $\theta = 0$ means either $\bar{x} = 0$ and $\bar{h} \neq 0$ or $\bar{x} \neq 0$ and $\bar{h} = 0$. The larger the difference between \bar{x} and \bar{h} , the smaller the θ and vice versa.

Fig. 1. Part 1 of CMA revised derivation

C7

The next step is to compute the linear regression coefficient (or trend slope) such that it varies from zero to one. As already presented in the discussion paper, consider u_i as the number of times the i^{th} data point is exceeded by others within the given sample. Again, for the given sample, take v_i as the number of times the i^{th} data point exceeds others. We can non-parametrically rescale Y and X in terms of the difference between exceedance and non-exceedance counts of data points to obtain series d_y and d_x , respectively using

$$d_{y,i} = v_{y,i} - u_{y,i} \text{ for } 1 \leq i \leq n \quad (A6)$$

$$d_{x,i} = v_{x,i} - u_{x,i} \text{ for } 1 \leq i \leq n \quad (A7)$$

where u_y and v_y respectively denote u and v applied to y 's. Similarly, u_x and v_x respectively represent u and v applied to x 's. For an illustration, consider the dataset Y with $n = 10$ such that $y = \{5, 8, 3, 6, 2, 5, 8, 1, 4, 6\}$. It means $u_y = \{4, 0, 7, 2, 8, 4, 0, 9, 6, 2\}$, $v_y = \{4, 8, 2, 6, 1, 4, 8, 0, 3, 6\}$, and $d_y = v_y - u_y = \{0, 8, -5, 4, -7, 0, 8, -9, -3, 4\}$.

Non-parametric normalized linear trend slope f for regression of Y on X can be computed using

$$f = \left(\sum_{i=1}^n (d_{y,i} \times d_{x,i})\right) \times \left(\sqrt{\sum_{i=1}^n d_{y,i}^2 \times \sum_{i=1}^n d_{x,i}^2}\right)^{-1} \quad (A8)$$

Finally, CMA can be computed using

$$CMA = |f| \times \beta \times \theta \quad (A9)$$

where $|f|$ denotes the absolute value of f . The values of CMA range from 0 to 1. CMA equal to one indicates a perfect model (no errors). However, CMA equal to zero indicates that the model is not better than the comparison baseline (such as the mean of observed data).

Instead of using β , we can use λ to computed the alternative coefficient of model accuracy (ACMA) such that

$$ACMA = |f| \times \lambda \times \theta \quad (A10)$$

Like CMA, the values of ACMA also range from 0 to 1. For an ideal model (where there are no errors), ACMA is equal to one. When ACMA is zero, it means the modeled series is the same as the mean of the observed data. For data with large variance or coefficient of variation, a good skill score SC can be given by $SC = \beta \times \theta$.

Fig. 2. Part 2 of CMA revised derivation

C8

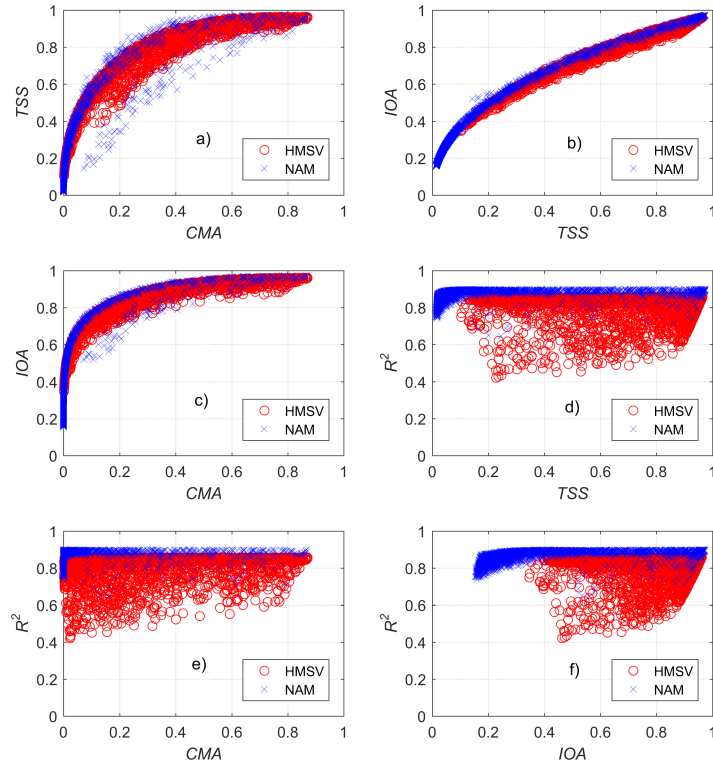


Fig. 3. Figure 3: Observed versus modeled flow from (a-g) HMSV and (h-n) NAM rainfall-runoff models calibrated using (a, h) NSE, (b, i) R2, (c, j) CMA, (d, k) IOA, (e, l) RMSE, (f, m) MAB, and (g, n) TSS

C9

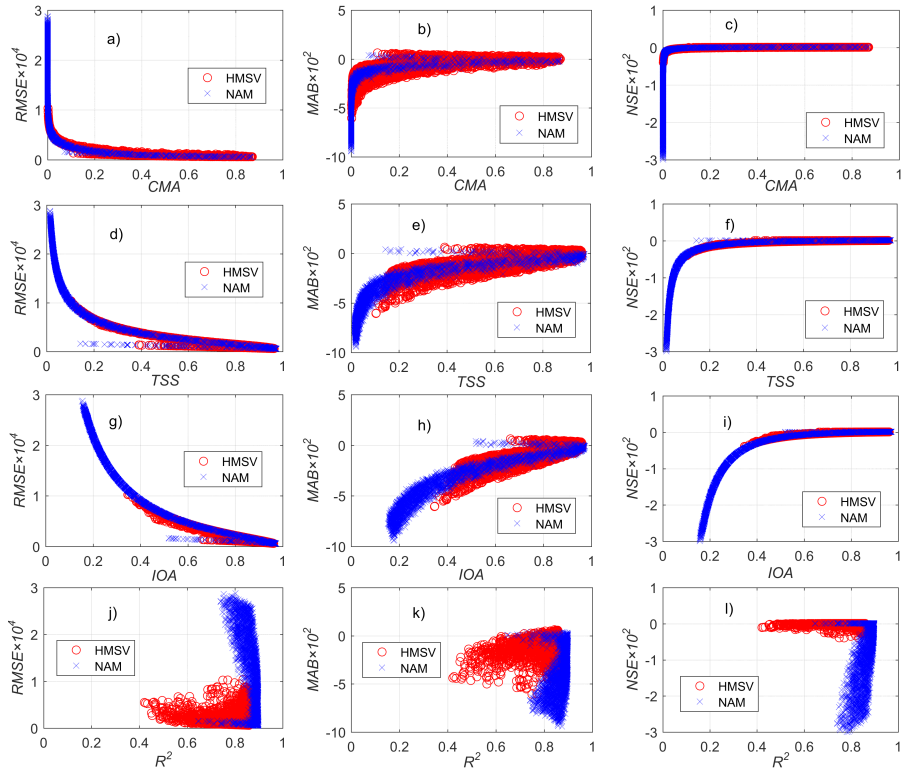


Fig. 4. Figure 4: Comparison of modeled and observed flow in terms of a) variance, b) skewness, c) kurtosis, and d) coefficient of variation (CV), as well as the mean of e) long-term daily flow, f) minimum fl

C10

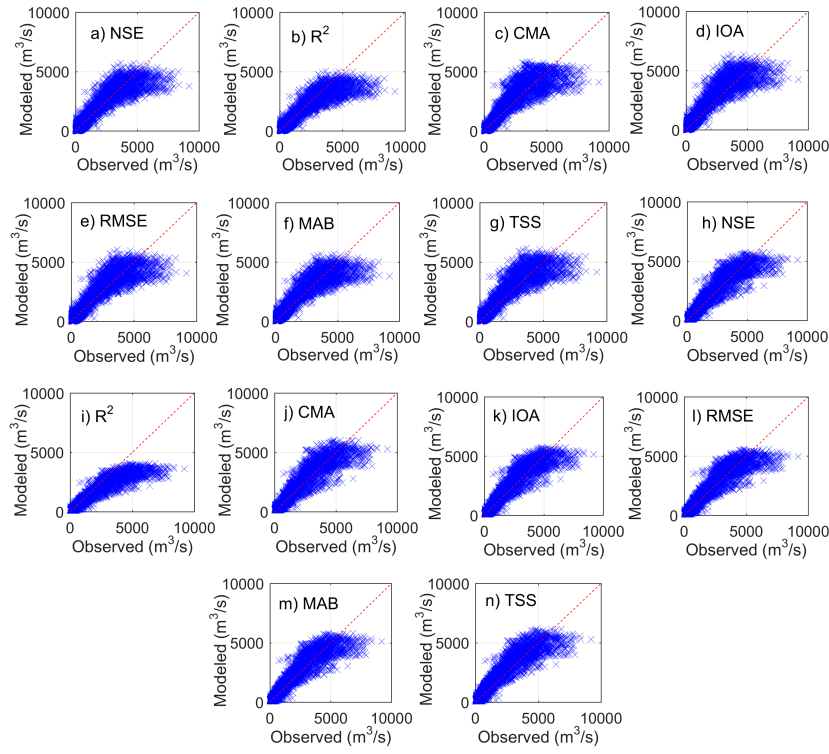


Fig. 5. Observed versus modeled flow from (a-g) HMSV and (h-n) NAM rainfall-runoff models calibrated using (a, h) NSE, (b, i) R², (c, j) CMA, (d, k) IOA, (e, l) RMSE, (f, m) MAB, and (g, n) TSS

C11

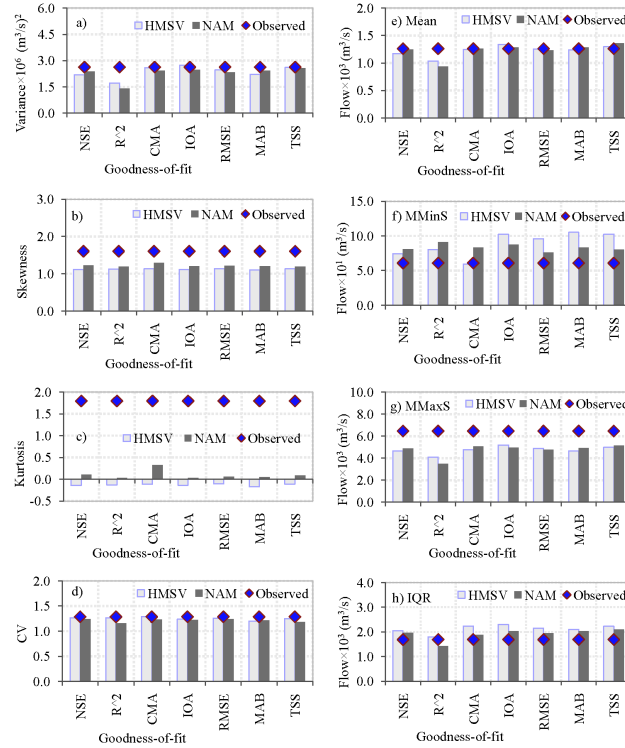


Fig. 6. Comparison of modeled and observed flow in terms of a) variance, b) skewness, c) kurtosis, and d) coefficient of variation (CV), as well as the mean of e) long-term daily flow, f) minimum fl

C12