Response to Dr. John Ding

Here we provide our response to comment from Dr. Ding including the plan to revise the manuscript in response to his comments. The original review comments are in black, and our responses in red.

SC1 The NSE criterion

Having had taken a postgraduate-level course from Nash (1968-69) when he was a visiting professor at my alma mater, the University of Guelph in Ontario, Canada, I believe I was privy to his thinking behind the development of the NSE criterion, a variant of the sum of squared error measure (Nash and Sutcliffe, 1970).

I could be wrong because of the passage of time, but his thinking was that in the absence of a model, the best estimate of the future flows was the average or mean observed flow value $\bar{O}$, thus its appearance in the denominator of the criterion below (e.g., Ding, 1974, Equation 47):

$$R^2 = 1 - F/F_0$$

$$F_0 = \sum_{t=0}^{L-1} (Q_{t+1} - \bar{O})^2$$

$$F = \sum_{t=0}^{L-1} (Q_{t+1} - O_{t+1})^2$$

in which R^2 is the model efficiency; F_0 and F are the initial variance from the mean value and the residual variance from the observed values, respectively; *O* and *Q* are the observed and simulated flow series, respectively; *t* is the index of timestep Δt; and *L* is the length, or number of ordinates, of the observed hydrograph.

To my way of thinking then, a better estimator would instead be based be a one-step ahead forecast $\hat{O}_{t+1} = O_t + (O_t - O_{t-1})$. Replacing the *O* term in Equation (2) by $\hat{O}_{t+1}$ produces Equations (4) and (5):

$$F_1 = \sum_{t=0}^{L-1} \left(Q_{t+1} - \hat{O}_{t+1}\right)^2$$

$$R_1^2 = 1 - F/F_1$$

This would make F_1 smaller in value than F_0, and the modified NSE criterion, R_1^2, a lower score one. The drawback was to always overshoot, by one timestep, the turning points, the peaks and troughs of an observed hydrograph.

References

Ding, J. Y.: Variable unit hydrograph, J. Hydrol., 22, 53-69, 1974.

Nash, J. E.: A course of lectures on parametric or analytical hydrology. Great Lakes Institute, University of Toronto, Ont., PR 38, 1968-69.
Nash, J.E. and Sutcliffe, J.V.: River flow forecasting through conceptual models, part I

- A discussion of principles, J. Hydrol., 10(3), 282-290, 1970.

First of all, we appreciate your efforts to share the history of earlier work on optimization metric development. Based on the Nash and Sutcliffe's paper, development of NSE was motivated for the model optimization and the NSE efficiency given for a particular model is a ratio of the model skill, here sum of squared of error, to the simplest flow estimation (i.e., observed mean flow).

More generally the NSE is formulated as a skill score $S=1-e/e_{ref}$, where e is some error metric and e_{ref} is the reference benchmark. While the observed variance is used as a benchmark in the NSE, other benchmarks can be used as well (e.g., see Schaefli and Gupta, 2007). To our understanding, your proposal is to use persistence as a benchmark, an approach used in the climate community.

NSE has been used predominantly for model optimization (esp., hydrologic models) for their practical applications. In the meantime, there has been substantial studies on the metrics along with evolution of optimization techniques (i.e., multi-criteria-based optimization). This has been discussed in Gupta et al., 2009.

If your intent is to advocate for a wider set of benchmarks, then we wholeheartedly agree on this point (e.g., Pappenberger et al., 2015); however, this is not the main objective of this contribution. We consider adding some brief descriptions on the meaning of NSE along with the historical developments of the optimization metrics.

References

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F., 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, *Journal of Hydrology*, 377, pp.80–91.

Pappenberger, F., Ramos, M.H., Cloke, H.L., Wetterhall, F., Alfieri, L., Bogner, K., Mueller, A. and Salamon, P., 2015. How do I know if my forecasts are better? Using benchmarks in hydrological ensemble prediction. *Journal of Hydrology*, *522*, pp.697-713.

Schaefli, B. and Gupta, H.V., 2007. Do Nash values have value?. *Hydrological Processes: An International Journal*, *21*(15), pp.2075-2080.