The objective function, Nash–Sutcliffe efficiency scale, and autoregressive projections

Parameter optimization of a hydrologic model needs to specify an objective or penalty function for the model to meet.

The classical Nash and Sutcliffe 1970 efficiency scale (NSE) expressed by Eq. (10) can be recast using the original notation as: $R^2 = 1 - F/F_0$. It has both an objective function in residual variance F, which is sum of squares of the simulation error (SSE) and an observed-mean-flow (μ_o) benchmark embedded in initial variance F_0, a fixed value. There is a one-to-one correspondence between NSE and F, and optimizing NSE is same as optimizing F. But this is not necessarily true in its variants, including an earliest known one, Ding 1974, Eqs. (40) and (47) therein.

NSE is a measure of correlation as well as others between simulation and observation as shown in a componentized form in Eq. (11). What it needs physically as well as statistically is at least one auxiliary benchmark to help interpret its intermediate scores between a perfect score of 1 for an observed or reference hydrograph, i.e. a perfect model, and of 0 for the (primary) benchmark model, μ_o. Establishing auxiliary benchmarks or baselines will help address one question about the popular performance metric: how close to 1 are NSE values reachable by models, e.g., Nearing et al. 2022, Table 1 therein.

The concept of two-parameter ($\omega 1$:$\omega 2$) homothetic transformation hydrographs represents a first step toward searching for such auxiliary benchmarks, as described for a twin-peak synthetic hydrograph in Sections 3.1, 3.2, Equation (21), and presented in Figures 1, 2 and 3.

I've put forward a simplest second-order autogressive process of the streamflow, $AR(2, c = 0, c_1 = 2, c_2 = -1)$, as a replacement of the primary benchmark, μ_0, e.g., Ding 2018. This, a slope-based projection hydrograph, instead could be considered a secondary benchmark, e.g., Azmi et al. 2021, SC1 and AC1 therein. In the same vein, a simplest third-order $AR(3, 0, 2, -2, 1)$, a curvature-based projection hydrograph, could be a tertiary one.

AR(2) and AR(3) projection hydrographs can be generated for the twin-peak example hydrograph. Scoring them would yield NSE values, calibration free.

I encourage the authors to pursue this AR projection approach in a future study. For the example hydrograph, I for one would be interested in what are NSE scores for AR(2) and AR(3) benchmarks, and whether the higher score of the two is lower than but close to the values shown in Fig. 3(a) for both BB (Bad-Bad) and BG(Bad-Good) transformations.

References

Azmi, E. et al. (2021). "Technical note: "Bit by bit": a practical and general approach for evaluating model computational complexity vs. model performance". In: *Hydrology and Earth System Sciences* 25.2, pp. 1103–1115. DOI: `10.5194/hess-25-1103-2021`. URL: `https://hess.copernicus.org/articles/25/1103/2021/`.

Ding, J. (2018). "Interactive comment on "On the choice of calibration metrics for "high flow" estimation using hydrologic models" by Naoki Mizukami et al." In: DOI: `https://doi.org/10.5194/hess-2018-391-SC1`.

Ding, JY (1974). "Variable unit hydrograph". In: *Journal of Hydrology* 22.1-2, pp. 53–69.

Nash, J Eamonn and Jonh V Sutcliffe (1970). "River flow forecasting through conceptual models part I—A discussion of principles". In: *Journal of hydrology* 10.3, pp. 282–290.

Nearing, G. S. et al. (2022). "Technical note: Data assimilation and autoregression for using near-real-time streamflow observations in long short-term memory networks". In: *Hydrology and Earth System Sciences* 26.21, pp. 5493–5513. DOI: `10.5194/hess-26-5493-2022`. URL: `https://hess.copernicus.org/articles/26/5493/2022/`.