Supplemental Material

Journal of Hydrometeorology

A Framework for Diagnosing Factors Degrading the Streamflow Performance of a Soil Moisture Data Assimilation System

https://doi.org/10.1175/JHM-D-18-0115.1

Supplemental Material

A framework for diagnosing factors degrading the streamflow performance of a soil moisture data assimilation system

Yixin Mao[1], Wade T. Crow[2] and Bart Nijssen[1]

1: Department of Civil and Environmental Engineering, University of Washington, Seattle, WA

2: Hydrology and Remote Sensing Laboratory, Agricultural Research Service, USDA, Beltsville, MD

Corresponding author: Bart Nijssen (nijssen@uw.edu)

S1. Mathematical details of ensemble Kalman filter (EnKF)

The ensemble Kalman filter (EnKF) method is one of the most commonly used data assimilation (DA) techniques in hydrology. EnKF was first introduced by Evensen [1994] and has subsequently been applied to a large number of land DA applications [e.g., Crow and Ryu, 2009; Chen et al., 2014; Massari et al., 2015]. It represents model error by an ensemble of model run replicates, which are combined with measurements sequentially to update model states. Specifically, the EnKF method is based on a propagation model and a measurement model:

$$x_{k+1} = f(x_k, u_k) + \omega_k \quad \text{(S1)}$$

$$\tilde{y}_k = Hx_k + v \quad \text{(S2)}$$

where subscript k is a discrete time index; x is a column vector of model states to update (the column vector length is the total number of state variables to update); u is model meteorological forcing; *f()* is a land surface model that propagates states to the next timestep; ω lumps together modeling errors during propagation from various sources including forcing data error, model structure error and parameterization error; $\tilde{y}$ is measurement data, in our context surface SM measurements; H is an observation operator that relates model states x to measurements $\tilde{y}$; and ν is measurement error.

In a standard EnKF, an ensemble size of N model replicates is propagated and updated sequentially over time in the following way:

1) An ensemble of initial model states is first generated by perturbing the initial deterministic model states to represent initial state error;

2) For each ensemble member, the land surface model is run until the next measurement time with perturbed meteorological forcing to represent forcing error. Model states are directly perturbed as well to represent random errors from model structure and parameterization;

3) Once an observation time is reached, the Kalman gain *K* is calculated as:

$$K_k = P_k H^T \cdot (H P_k H^T + R)^{-1} \tag{S3}$$

where *R* is the measurement error variance, and the forecast state error covariance matrix P_k is estimated by sampling across the propagated ensemble states:

$$P_k = \frac{1}{N-1}\sum_{j=1}^{N}(\hat{x}_k^{-(j)} - \bar{\hat{x}}_k^{-})(\hat{x}_k^{-(j)} - \bar{\hat{x}}_k^{-})^T \tag{S4}$$

where $\hat{x}_k^{-(j)}$ is the propagated state vector at time *k* for the *j*th ensemble member, and $\bar{\hat{x}}_k^{-}$ is the mean of $\hat{x}_k^{-(j)}$ across all ensemble members;

4) Following the calculation of *K*, each ensemble member of states is individually updated as:

$$\hat{x}_k^{+(j)} = \hat{x}_k^{-(j)} + K_k \cdot (\tilde{y}_k + v_k^{(j)} - \hat{y}_k^{-(j)}) \tag{S5}$$

where $\hat{y}_k^{-(j)}$ is the simulated measurement at time *k* for the *j*th ensemble member, i.e., $\hat{y}_k^{-(j)} = H\,\hat{\mathrm{x}}_k^{-(j)}$; $v_k^{(j)}$ is random noise added to represent measurement error whose error statistic is consistent with *R* in Equation (S3).

S2. Mathematical details of the evaluation metrics used in the study

1) Percent error reduction (PER)

PER is defined as the percent reduction in the root-mean-squared error (RMSE) compared to the open-loop baseline:

$$PER = \left[1 - \frac{RMSE_a}{RMSE_{open}} \right] \times 100 \qquad \text{(S6)}$$

where the subscripts *open* and *a* denote the open-loop and DA analysis runs, respectively.

2) Kling-Gupta efficiency (KGE)

The Kling-Gupta efficiency (KGE) [Gupta et al. 2009] combines the performance in terms of correlation, variance and bias:

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \qquad \text{(S7)}$$

where r is the correlation coefficient between simulated and observed streamflow; α is the ratio of their standard deviations; and β is the ratio of their means. KGE ranges from negative infinity to 1 with values closer to 1 indicating better performance.

3) Percent continuous rank probability score reduction (PSR)

Continuous rank probability score (CRPS) measures the deviation of the cumulative distribution function (CDF) of an ensemble from that of a reference (observation in the real-data case or truth in the synthetic case) [Hersbach, 2000]. If we assume the reference has zero uncertainty, then its CDF (denoted by $F_t^r(s)$ where t denotes timestep and r denotes observation) is a unit step function:

$$F_t^o(\mathrm{s}) = \begin{cases} 0, & \mathrm{s} < \mathrm{y}_t^r \\ 1, & s \geq \mathrm{y}_t^r \end{cases} \qquad \text{(S8)}$$

where y_t^r is the reference variable value and *s* is a random variable. CRPS is then calculated as the temporal mean of the CDF deviation from the reference:

$$CRPS = \frac{1}{n}\sum_{t=1}^{n}\int_{-\infty}^{\infty}[F_t^a(\mathrm{s}) - F_t^r(\mathrm{s})]^2 ds \tag{S9}$$

where $F_t^a(s)$ denotes the CDF of an analysis ensemble at time *t*, and *n* is the total length of the time series. In practice, the continuous CDF of an analysis, $F_t^a(s)$, is empirically estimated by the finite ensemble. Note that CRPS penalizes both a deviation of the ensemble mean from the observation and a large ensemble spread. A smaller (i.e., closer-to-zero) CRPS value indicates better ensemble performance.

Analogous to PER, the percent CRPS reduction (PSR) quantifies the percent reduction in CRPS of a DA analysis compared to the open-loop baseline (note that here the baseline is the open-loop ensemble instead of the deterministic open-loop run):

$$PSR = \left[1 - \frac{CRPS_a}{CRPS_{open}}\right] \times 100 \tag{S10}$$

4) Normalized ensemble skill (NENSK)

NENSK measures the ensemble-mean error normalized by ensemble spread:

$$NENSK = \frac{ENSK}{ENSP} \tag{S11}$$

where the ensemble skill (ENSK) and ensemble spread (ENSP) are calculated as:

$$ENSK = \frac{1}{n}\sum_{t=1}^{n}(\overline{\mathrm{y}}_t^a - \mathrm{y}_t^o)^2 \tag{S12}$$

$$ENSP = \frac{1}{n}\sum_{t=1}^{n}(y_t^{a,(j)} - \bar{y}_t^{a})^2 \qquad \text{(S13)}$$

where $\bar{y}_t^a$ denotes the ensemble-mean value at timestep t, and $y_t^{a,(j)}$ denotes the value of the jth ensemble member. In an ideal situation where an ensemble is a correct representation of analysis uncertainty, the observed or true condition looks like one of the realizations of the ensemble [Anderson, 1996; Wilks, 2011] and therefore NENSK should be approximately one [Talagrand et al., 1997; Wilks, 2011] (NENSK > 1 indicates an under-dispersed ensemble while NENSK < 1 indicates an over-dispersed ensemble). This metric has been used to verify ensemble hydrologic variables [e.g., De Lannoy et al., 2006; Brocca et al., 2012; Alvarez-Garreton et al., 2014].

References:

Alvarez-Garreton, C., D. Ryu, A. W. Western, W. T. Crow, and D. E. Robertson (2014), The impacts of assimilating satellite soil moisture into a rainfall-runoff model in a semi-arid catchment, *J. Hydrol.*, 519, 2763-2774, doi:10.1016/j.jhydrol.2014.07.041.

Anderson, J. L. (1996), A method for producing and evaluating probabilistic forecasts from ensemble model integrations, *J. Clim*, 9, 1518-1530, doi: 10.1175/1520-0442(1996)009<1518:AMFPAE>2.0.CO;2.

Brocca, L., T. Moramarco, F. Melone, W. Wagner, S. Hasenauer, and S. Hahn (2012), Assimilation of surface-and root-zone ASCAT soil moisture products into rainfall–runoff modeling, *IEEE Trans. Geosci. Remote Sens.*, 50(7), 2542-2555, doi: 10.1109/TGRS.2011.2177468.

Chen, F., W. T. Crow, and D. Ryu (2014), Dual forcing and state correction via soil moisture assimilation for improved rainfall–runoff modeling, *J. Hydrometeorol.*, *15*(5), 1832–1848, doi:10.1175/JHM-D-14-0002.1.

Crow, W. T., and D. Ryu (2009), A new data assimilation approach for improving hydrologic prediction using remotely-sensed soil moisture retrievals, *Hydrol. Earth Syst. Sci.*, 12(1-16), doi:10.5194/hess-13-1-2009.

De Lannoy, G. J. M., P. R. Houser, V. R. N. Pauwels, and N. E. C. Verhoest (2006), Assessment of model uncertainty for soil moisture through ensemble verification, *J. Geophys. Res.*, 111, D10101, doi: 10.1029/2005JD006367.

Evensen, G. (1994), Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, *99*(C5), 10143-10162, doi:10.1029/94JC00572.

Gupta, H. V., H. Kling, K. K. Yilmaz, and G. F. Martinez (2009), Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, *J. Hydrol.*, 377, 80-91, doi: 10.1016/j.jhydrol.2009.08.003.

Hersbach, H. (2000), Decomposition of the continuous ranked probability score for ensemble prediction systems, *Weather Forecast.*, 15(5), 559-570, doi: 10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2.

Massari, C., L. Brocca, A. Tarpanelli, and T. Moramarco (2015), Data Assimilation of Satellite Soil Moisture into Rainfall-Runoff Modelling: A Complex Recipe?, *Remote Sens.*, 7, 11403-11433, doi:10.3390/rs70911403.

Talagrand, O., R. Vautard, and B. Strauss (1997), Evaluation of probabilistic prediction systems, technical report, Eur. Cent. for Medium-Range Weather Forecast., Reading, UK.

Wilks, D. S. (2011), Statistical methods in the atmospheric sciences (3rd edition), Elsevier/Academic Press, Amsterdam; Boston.