Supplement of Hydrol. Earth Syst. Sci., 24, 615–631, 2020
https://doi.org/10.5194/hess-24-615-2020-supplement

Hydrology and Earth System Sciences
Open Access
EGU

Supplement of

Dual state/rainfall correction via soil moisture assimilation for improved streamflow simulation: evaluation of a large-scale implementation with Soil Moisture Active Passive (SMAP) satellite data

Yixin Mao et al.

Correspondence to: Bart Nijssen (nijssen@uw.edu)

Supplemental Material

S1. The ensemble Kalman smoother (EnKS) version of the Soil Moisture Analysis Rainfall Tool (SMART)

The Soil Moisture Analysis Rainfall Tool (SMART) is a rainfall correction scheme developed and updated by Crow et al. (2009; 2011) and Chen et al. (2012). It is based on sequential assimilation of soil moisture (SM) measurements into a simple Antecedent Precipitation Index (API) model to obtain SM increments. It then linearly relates these increments to rainfall accumulation errors. In the study we extended the ensemble Kalman filter (EnKF) version of SMART developed by Crow et al. (2011) to an ensemble Kalman smoother (EnKS) version with probabilistic rainfall estimates.

Following Crow et al. (2009; 2011), the API model is used to capture the response of moisture storage (represented by the *API* state) to rainfall input:

$$API_t = \gamma API_{t-1} + P_t \tag{S1}$$

where t is a time step index; P is the original uncorrected precipitation observation [mm] and γ is a loss coefficient (dimensionless) that accounts for storage loss through evaporation, drainage, etc. In the ensemble version of SMART (Crow et al., 2011), Eq. (S1) is converted to:

$$API_t^{(j)} = \gamma API_{t-1}^{(j)} + \eta_t^{(j)} P_t + \omega_t^{(j)} \tag{S2}$$

where the superscript *(j)* denotes the *j*th ensemble member; η is multiplicative noise with mean 1 added to the observed precipitation to represent random precipitation forcing error; and ω is zero-mean Gaussian noise to represent random API model structure and parameterization error. The API state can be related directly to SM content via rescaling (Crow et al., 2009). The rescaled SM measurement, θ, can therefore be assimilated to update the *API* states via the standard EnKS technique both at the measurement time step and during the data gap before the measurement time step. Mathematically, if two adjacent measurements come in at time k and time m with $m - k \geq 1$, then the measurement at time m is used to calculate the gain K and API increment δ for each time step i at time step m as well as during the gap (i.e., $k < i \leq m$):

$$K_i = \frac{T_{im}}{T_m + R_m} \tag{S3}$$

and

$$\delta_i^{(j)} = API_i^{+(j)} - API_i^{-(j)} = K_i \cdot (\theta_m + \kappa_m^{(j)} - API_m^{-(j)}) \tag{S4}$$

where K is the Kalman gain; T_{im} is the covariance matrix between *API* states at time i and m; R is the measurement error variance for the rescaled SM measurements; the superscript *(j)* denotes the *j*th ensemble member; the superscripts "-" and "+" denote *API* states before and after an update, respectively; and κ is zero-mean Gaussian noise added to represent the random SM measurement error. T_{im} is calculated as:

$$T_{im} = \frac{1}{M-1} \sum_{j=1}^{M} (API_i^{-(j)} - \overline{API}_i^{-}) \cdot (API_m^{-(j)} - \overline{API}_m^{-}) \tag{S5}$$

where M is the ensemble size; $\overline{API}_t^{-}$ is the ensemble-mean *API* states before update.

The SMART algorithm then uses ensemble-mean *API* increment δ to estimate the rainfall correction amount via a simple linear relation. We extended this relation to produce an ensemble of corrected rainfall time series (instead of the single rainfall estimates in past studies) where each ensemble member of the perturbed rainfall time series is corrected by the corresponding member of δ.:

$$[P_{corr}^{(j)}]_l = [\eta^{(j)} P^{(j)}]_l + \lambda [\delta^{(j)}]_l \tag{S6}$$

where "[ ]" denotes temporally aggregated P or δ (in the SMART study in this paper, this window was set to the 3-hour native SMART time step without aggregation); l is the new time index for the aggregated windows; and λ is a scaling factor that can either be calibrated or set to a prescribed constant. Finally, negative P_{corr} resulted from Eq. (S6) are reset to zero, and the final corrected precipitation time series is (multiplicatively) rescaled to be unbiased over the entire simulation period toward the original precipitation observation time series.

S2. Mathematical details of ensemble Kalman filter (EnKF) in the state update scheme

The ensemble Kalman filter (EnKF) method is a commonly used data assimilation (DA) techniques in hydrology. The EnKF technique applied in this study directly follows Mao et al. (2019). Below will briefly review its mathematical details.

The EnKF algorithm was applied to each SMAP pixel individually. The EnKF method is based on a propagation model and a measurement model:

$$x_{k+1} = f(x_k, u_k) + \omega_k \tag{S7}$$

$$\tilde{y}_k = Hx_k + v \tag{S8}$$

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), which, in our application, is top-layer VIC-simulated SM estimates in every finer-resolution VIC grid cell that is associated with a SMAP pixel; u is model meteorological forcing, in our context rainfall; $f()$ is a land surface model that propagates states to the next time step, in our context the VIC model; 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, i.e., $\tilde{y} = SM_1^{obs}$ where SM_1^{obs} is the SMAP observation at its native coarser resolution; H is an observation operator that relates model states x to measurements $\tilde{y}$, in our contex the areal-averaged first-layer SM state from the multiple VIC grid cells; and v is random 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 (all three VIC SM layers in our context) 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 (again, all three VIC SM layers) to represent random errors from model structure and parameterization;

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

$$K_k = P_k H^T \cdot (H P_k H^T + R)^{-1} \tag{S9}$$

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:

$$P_k = \frac{1}{N-1} \sum_{j=1}^{N} (\hat{x}_k^{-(j)} - \bar{\hat{x}}_k^{-})(\hat{x}_k^{-(j)} - \bar{\hat{x}}_k^{-})^T \tag{S10}$$

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 (only the first and second VIC SM layers from the top) is individually updated as:

$$\hat{x}_k^{+(j)} = \hat{x}_k^{-(j)} + K_k \cdot (\tilde{y}_k + v_k^{(j)} - \hat{y}_k^{-(j)}) \tag{S11}$$

where $\hat{y}_k^{-(j)}$ is the simulated measurement at time *k* for the *j*th ensemble member, i.e., $\hat{y}_k^{-(j)} = H\,\hat{x}_k^{-(j)}$; $v_k^{(j)}$ is random noise added to represent measurement error whose error statistic is consistent with *R* in Eq. (S9).

S3. The sensitivity of the SMART rainfall correction performance to the γ parameter

The unit-less γ parameter in Eq. (1) in the main manuscript was tuned such that the API model (approximately) optimally captured the SM dynamic observed by SMAP. We further carried out a sensitivity analysis of the rainfall correction performance to γ. Specifically, we varied γ to see its impact on the correlation coefficient improvement and percent RMSE reduction (PER). Figures S1 and S2 show the domain-median of both evaluation metrics, respectively, after correction at different γ values (in the manuscript, γ = 0.98 was used). We see that around our chosen value γ = 0.98, the sensitivity of rainfall correction performance to γ is relatively small, and γ = 0.98 results in optimal PER when evaluating SMART results at 1-day and 3-day accumulation periods (although performance is even slightly better at γ = 0.99 for the other measures shown). However, we also see that the correction performance is significantly degraded if γ is far from our chosen value (i.e., if γ < 0.95). These results should generally confirm that our selected γ value in the manuscript is reasonable and roughly optimal.

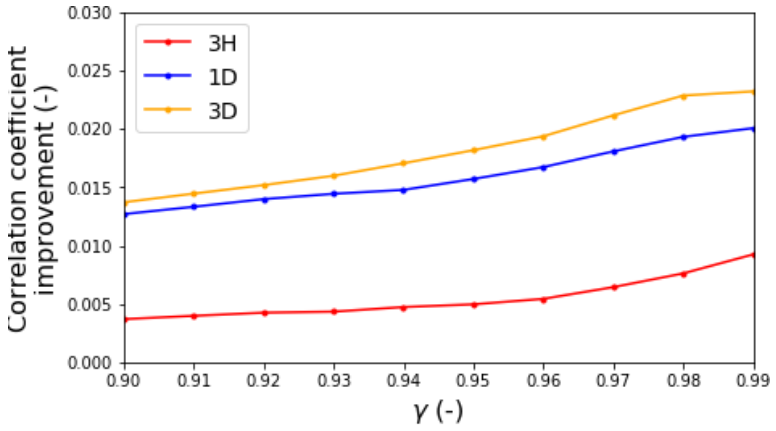

Figure S1. Domain-median correlation coefficient improvement of IMERG rainfall after SMART correction (with respect to the NLDAS-2 reference) using different γ values. Improvement is evaluated for 3-hour (3H), 1-day (1D) and 3-day (3D) accumulation intervals.

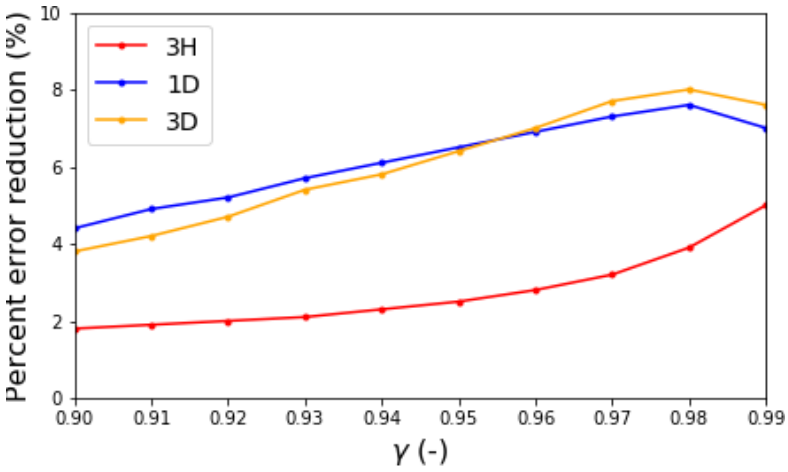

Figure S2. Same as Fig. 1, but for RMSE reduction (PER) evaluations.

S4. The impact of the λ parameter in the SMART rainfall correction scheme

In the SMART rainfall correction scheme, λ is a unit-less scaling factor that linearly relates the API state increment to rainfall correction amount. It can either be calibrated or set to a prescribed constant. We experimented with two strategies of determining λ in this study: 1) calibrating a temporally constant λ at each SMAP pixel separately to optimize the rainfall correlation with respect to the NLDAS-2 benchmark rainfall, and 2) setting λ to a spatial constant of 0.1, which is applicable for any region that may not have a good rain gauge coverage.

The rainfall correction results from the two strategies are shown in Fig. S3, in which column 1 shows the improvement of correlation coefficient *r* after SMART correction with λ tuned at each pixel to maximize *r* (with respect to the NLDAS-2 benchmark), and column 2 shows results obtained using a domain-constant value of $\lambda = 0.1$. Simply setting $\lambda = 0.1$ results in slightly smaller correlation improvement compared to the optimal λ case for all examined temporal accumulation periods (3-hour, 1-day and 3-day), especially for locations in the eastern and western ends of the domain. In general, these reductions are small, and since constant-λ is a more generally applicable case, we selected the $\lambda = 0.1$ strategy for all the SMART results presented in the main manuscript.

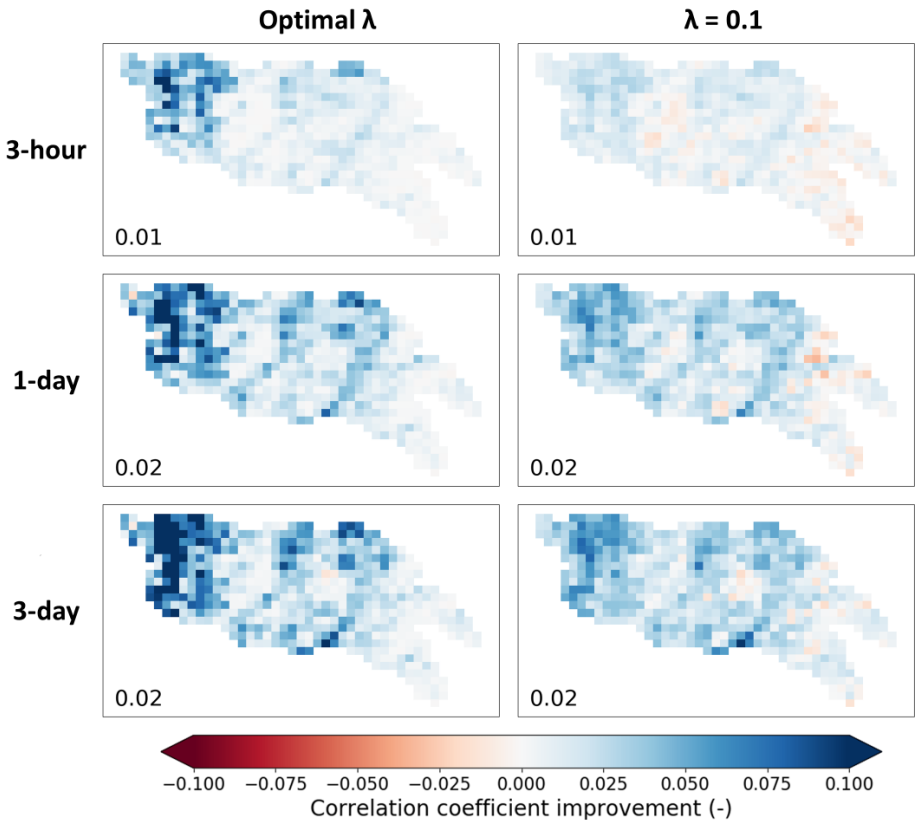

Figure S3. Maps of correlation coefficient improvement after SMART EnKS rainfall correction. The left column shows the results with λ tuned at each pixel to optimize the correlation coefficient of corrected rainfall relative to the NLDAS-2 benchmark, and the right column shows the results with domain-constant $\lambda = 0.1$ (this column is identical to the left column in Fig. 3 in the main manuscript). Each row shows results based on different temporal accumulation periods: 3-hourly, 1-day and 3-day aggregation, respectively. The number on the lower left corner of each subplot shows the domain-median correlation improvement.

S5. Investigation of cross-correlation of errors in the dual system

S5.1. Background and methods

It is well known that correlated errors in different parts of a Kalman filter result in sub-optimal filter outputs. Therefore, in the original paper detailing the dual state/rainfall correction system, Crow and Ryu (2009) advised that the corrected rainfall (informed by the SM measurements) should not be fed back into the state EnKF correction scheme into which the

same SM measurements are assimilated. Instead, corrected rainfall and states should be combined via an offline model simulation (see Fig. 1 and Sect. 2.4.3 in the main manuscript). Later studies that applied the dual correction system all followed this general guideline (e.g., Chen et al., 2014; Alvarez-Garreton et al., 2016). However, although this guideline avoids first-order error correlation in the system, it does not eliminate the possibility of error cross-correlation. Specifically, the corrected rainfall and the updated states are informed by the same SM measurement, thus they potentially inherit the same error from the SM measurements. See Sect. 2.4.3 in the main manuscript for more details.

To further investigate this issue, we designed a set of synthetic experiments and applied in an arbitrary small domain within the Arkansas-Red (a box around the Little Arkansas subbasin, see Table 1 and Fig. 2 in the main manuscript for its location). Synthetic measurements, instead of the real SMAP measurements, were generated and assimilated into the dual correction system so that we have complete control over all the error statistics and correlation, which is impossible in a real-data case. Specifically, a single perturbed VIC realization (with perturbed forcing and states) was treated as the synthetic "truth". Synthetic measurement can then be generated at daily interval by degrading the true surface-layer SM by adding random measurement errors. Precipitation perturbation was assumed to be temporally auto-correlated (first-order autoregressive noise with parameter $\phi = 0.9$), and all the other error assumptions and dual correction setup were consistent with those described in Sect. 2.4 in the main manuscript.

We generated two sets of synthetic measurements based on the same truth with the same measurement error statistics but mutually independent realizations of errors. Then, two scenarios of dual correction were designed and carried out (see Fig. S4 for illustration):

Scenario 1: a single time series of synthetic SM measurements were assimilated into both the state update and the rainfall correction schemes. This scenario mimics the issue in the real-data dual system with error cross-correlation in the two schemes and potentially degraded streamflow;

Scenario 2: two different time series of SM measurements (with mutually independent errors) were assimilated into the two schemes separately. This scenario completely avoids the issue of error cross-correlation.

The final runoff performance from the dual correction system were evaluated toward the truth, and the runoff performance from the two scenarios was compared. Differences in the performance of the two scenarios would indicate degradation caused by error cross-correlation present in Scenario 1. For these synthetic experiments, runoff was evaluated locally at each grid cell without routing, since we know the true condition locally.

S5.2. Results

Deterministic and probabilistic results from the two scenarios were compared in Fig. S5 and Fig. S6. Clearly, runoff results show only very little difference between the two scenarios in terms of both PER and NENSK (see Sect. 2.5 in the main manuscript for details of the two metrics). This is true for both the total runoff and the fast- and slow-response runoff components separately. This suggests that the streamflow performance is not noticeably degraded by assimilating the same SM retrievals to both the state update and rainfall correction schemes. Although cross-correlated error theoretically exists in the system, it is not persistent enough to cause problematic streamflow results. In other words, we are not significantly over-using the information contained in SM retrievals in the system. This is true both from a deterministic sense and in terms of probabilistic representation. We also experimented the case where the synthetic measurements were assumed to have temporally auto-correlated errors instead of white errors, which in theory creates an enhanced risk of degradation in the subsequent streamflow, but drew similar conclusions as above (results not shown).

The synthetic results in this section verifies that we can safely assimilate a single time series of SMAP retrievals into both parts of the dual correction system without significantly degrading the final streamflow estimates.

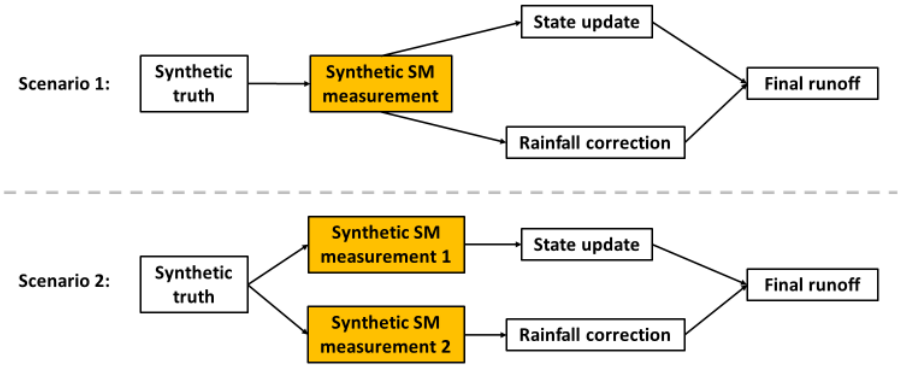

Figure S4. Illustration of the synthetic experiments for investigating error cross-correlation.

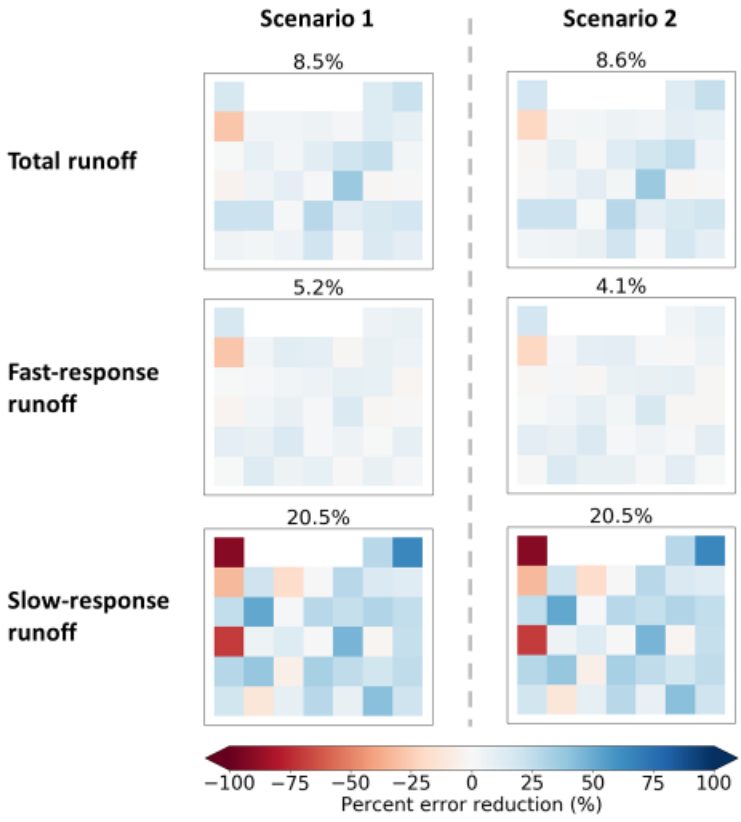

Figure S5. PER of synthetic daily runoff results from the error cross-correlation experiment. Blue color indicates runoff improvement after dual correction while red color indicates degraded runoff. The two columns show the results from the two assimilation scenarios described in Sect.

S5. The three rows show results of total runoff, fast-response runoff and slow-response runoff, respectively. The number on top of each subplot indicates the domain-median PER.

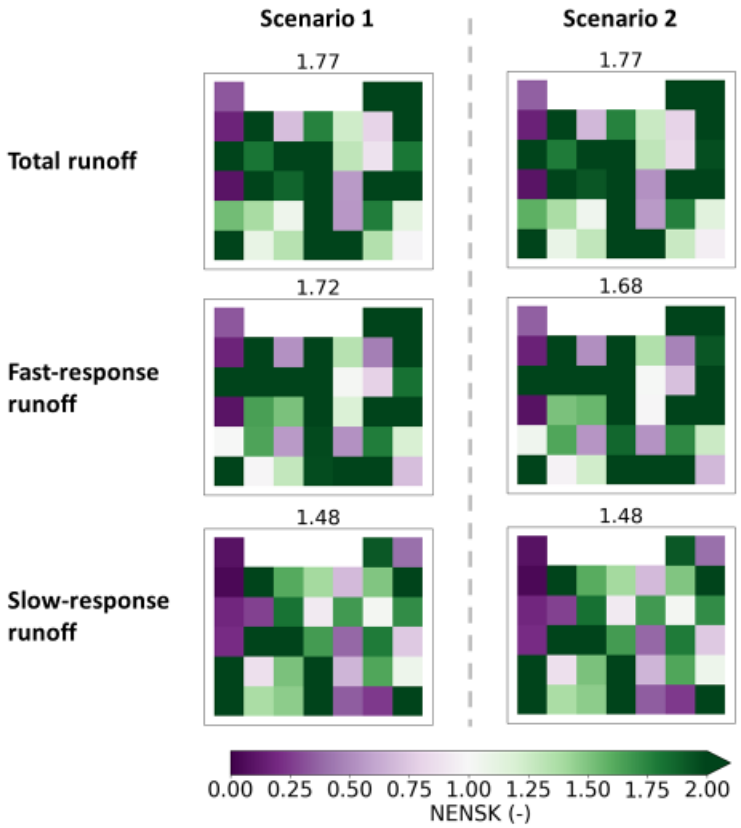

Figure S6. Same as Fig. S5 but for NENSK. Lighter color (either green or purple) indicates closer-to-one (thus better) NENSK.

S6. Sensitivity analysis of SMART rainfall correction performance to rain/no rain threshold

We have added a sensitivity analysis of SMART rainfall correction performance to rain/no rain threshold. Specifically, we altered the threshold of classifying IMERGE rain/no rain (this threshold was essentially set to zero in the manuscript and SMART only corrected time steps during which non-zero rainfall occurs), and observed its impact on the rainfall correction

results (i.e., categorical metrics at different rainfall scales as well as correlation improvement and PER).

Figures S7 to S9 show the SMART correction results with different rain/no rain thresholds. For categorical metrics (Fig. S7), having a rain/no rain accumulation threshold of 1 mm/3 hours or 2 mm/3 hours mitigates the issue of worsened POD at small rainfall events comparing to zero threshold, but also removes the (although small) FAR improvement. For mid-ranged rainfall events, a positive threshold mitigates the issue of worsened FAR as in the zero-threshold case, but POD improvement becomes smaller. For larger rainfall events, POD improvement and TS improvement become slightly smaller (i.e., closer to zero) when using a positive rain/no rain threshold (note that the small positive rain/no rain threshold value can be considered as a "larger" rainfall event percentile wise at some pixels with overall low precipitation, therefore affecting the categorical metrics toward the right side on the categorical metrics plots).

In addition to the categorical metrics, setting the rain/no rain threshold to either 1 mm/3 hours or 2 mm/3 hours slightly lowers values of correlation coefficient improvement and PER versus the baseline case of applying a rain/no rain threshold of zero accumulation (Figures S8 and S9).

In summary, there is no obvious optimized (non-zero) value for the rain/no rain threshold since there is a trade-off between POD and FAR performance. Although the overall TS at smaller rainfall events improves with a non-zero threshold, the correction for larger events, which SMART is most suitable for, slightly worsens. Therefore, a positive rain/no rain threshold does not benefit correlation coefficient and PER (which are sensitive to both POD and FAR performance). Based on this analysis, we selected a zero rain/no rain threshold for all SMART correction results presented in the main manuscript.

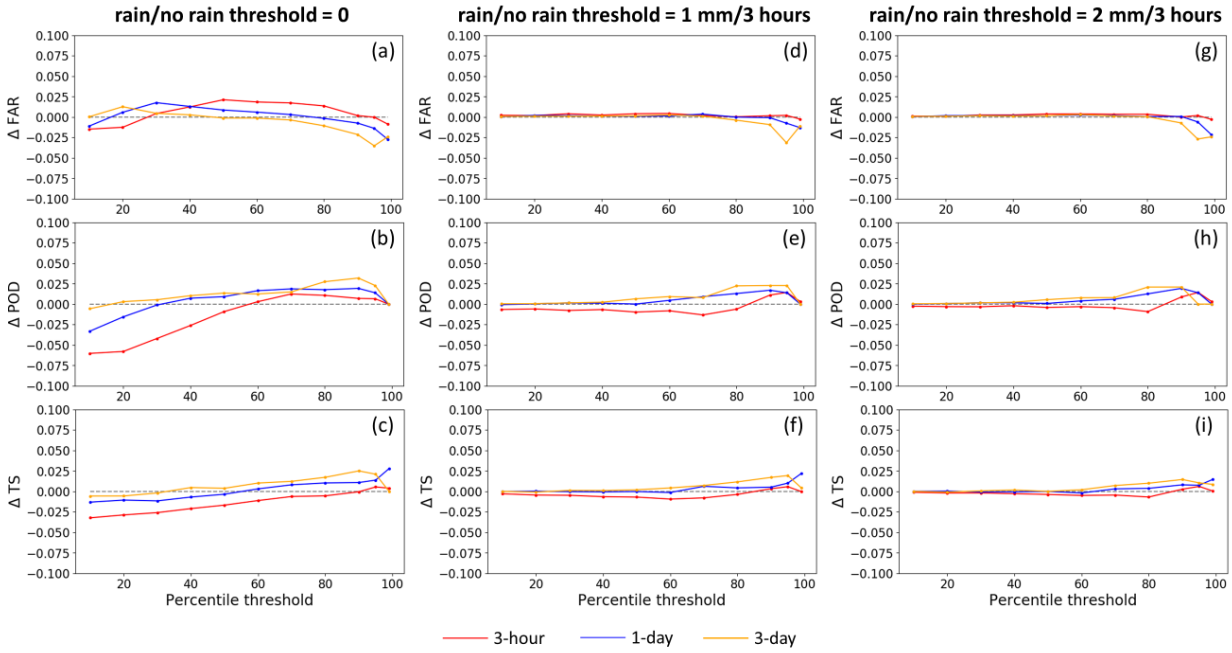

Figure S7: Change in categorical metrics (FAR, POD and TS) before and after SMART correction for 3-hourly, 1-day and 3-day accumulation periods. The left column (panels a, b and c) is the same as in Fig. 4 (right column) in the main text with SMART only correcting IMERG rainfall events with non-zero accumulation. The middle and right columns show the same metrics with SMART only correcting IMERG rainfall for events where accumulation rates exceed thresholds of 1 mm/3 hours and 2 mm/3 hours, respectively.

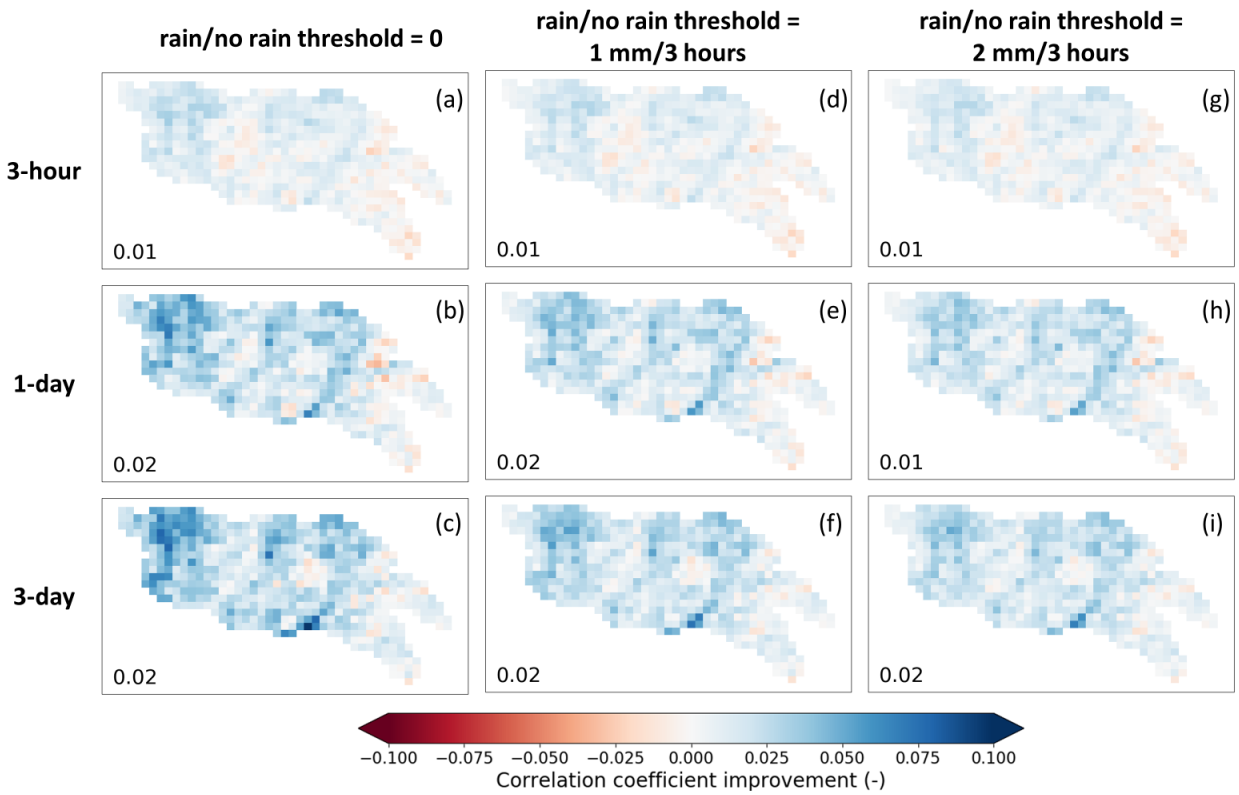

Figure S8: Correlation coefficient (with respect to the NLDAS-2 reference precipitation) improvement before and after SMART correlation for 3-hourly, 1-day and 3-day accumulation periods. As in Fig. 7, the left column (panels a, b and c) is the same as in Fig. 4 (right column) in the main text with SMART only correcting IMERG rainfall events with non-zero accumulation. The middle and right columns show the same metrics with SMART only correcting IMERG rainfall for events where accumulation rates exceed thresholds of 1 mm/3 hours and 2 mm/3 hours, respectively.

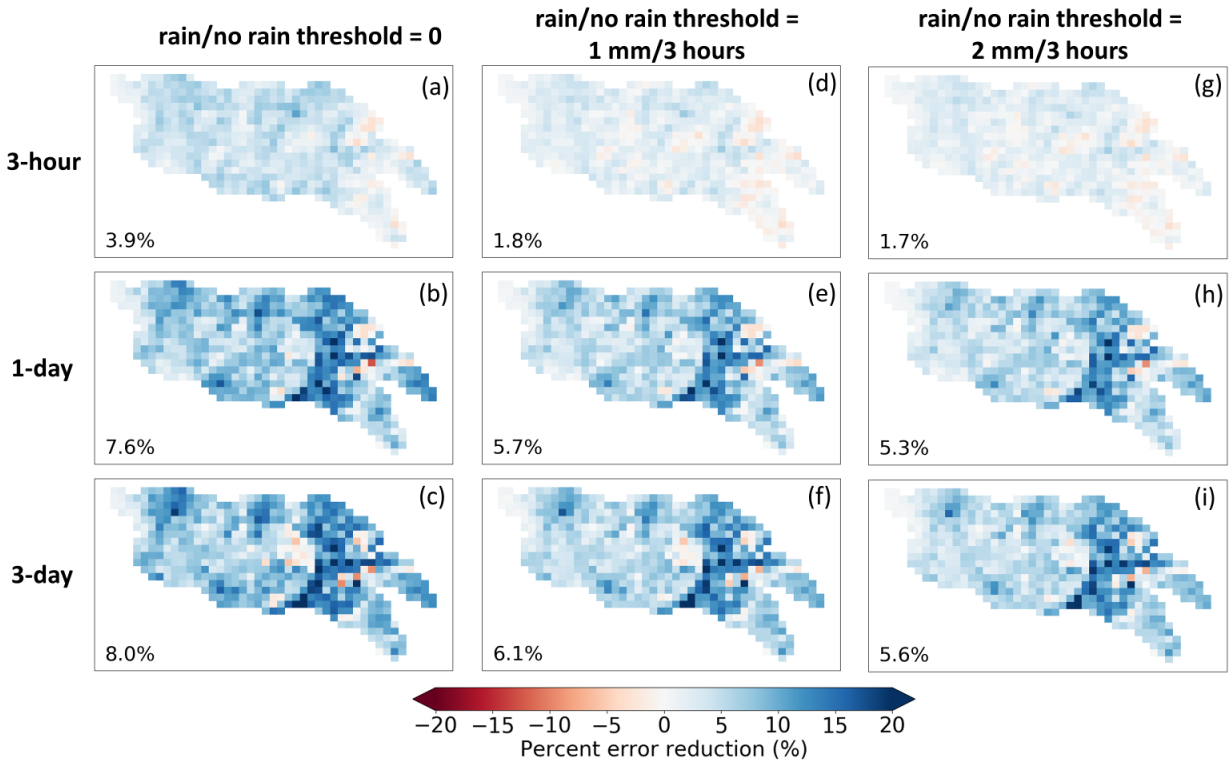

Figure S9: Same as Fig 8, but for percent RMSE reduction (PER; with respect to the NLDAS-2 reference precipitation). The left column (panels a, b and c) is the same as in Fig. 5 (left column) in the main text.

References:

Alvarez-Garreton, C., Ryu, D., Western, A. W., Crow, W. T., Su, C.-H., and Robertson, D. R.: Dual assimilation of satellite soil moisture to improve streamflow prediction in data-scarce catchments, Water Resour. Res., 52(7), 5357-5375, doi:10.1002/2015WR018429, 2016.

Chen F., Crow, W. T., and Holmes, T. R. H.: Improving long-term, retrospective precipitation datasets using satellite-based surface soil moisture retrievals and the Soil Moisture Analysis Rainfall Tool, J. Appl. Remote Sens., 6(1), 063604, doi:10.1117/1.JRS.6.063604, 2012.

Chen, F., Crow, W. T., and Ryu, D.: 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, 2014.

Crow, W. T., and Ryu, D.: 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, 2009.

Crow W. T., Huffman, G. J., Bindlish, R., and Jackson, T. J., Improving satellite-based rainfall accumulation estimates using spaceborne surface soil moisture retrievals, J. Hydrometeorol., 10, 199-212, doi:10.1175/2008JHM986.1, 2009.

Crow, W. T., van den Berg, M. J., Huffman, G. J., and Pellarin, T.: Correcting rainfall using satellite-based surface soil moisture retrievals: The Soil Moisture Analysis Rainfall Tool (SMART), Water Resour. Res., 47, W08521, doi:10.1029/ 2011WR010576, 2011.

Mao Y., Crow, W. T., and Nijssen, B.: A framework for diagnosing factors degrading the streamflow performance of a soil moisture data assimilation system, J. Hydrometeorol., 20(1), 79-97, doi:10.1175/JHM-D-18-0115.1, 2019.