**Non-Gaussian Ensemble Filtering and Adaptive Inflation for Soil Moisture Data Assimilation**

Emmanuel C. Dibia,a Rolf H. Reichle,b Jeffrey L. Anderson,c and Xin-Zhong Liang a,d

a Department of Atmospheric and Oceanic Science, University of Maryland, College Park, College Park, Maryland

b Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland

c National Center for Atmospheric Research, Boulder, Colorado

d Earth System Science Interdisciplinary Center, College Park, Maryland

*Corresponding author*: Xin-Zhong Liang, xliang@umd.edu

ABSTRACT

The rank histogram filter (RHF) and the ensemble Kalman filter (EnKF) are assessed for soil moisture estimation using perfect model (identical twin) synthetic data assimilation experiments. The primary motivation is to gauge the impact on analysis quality attributable to the consideration of non-Gaussian forecast error distributions. Using the NASA Catchment land surface model, the two filters are compared at 18 globally-distributed single-catchment locations for a 10-year experiment period. It is shown that both filters yield adequate estimates of soil moisture, with the RHF having a small but significant performance advantage. Most notably, the RHF systematically increases the normalized information contribution (NIC) score of the mean absolute bias by 0.05 over that of the EnKF for surface, root-zone and profile soil moisture. The RHF also increases the NIC score for the anomaly correlation of surface soil moisture by 0.02 over that of the EnKF (at a 5% significance level). Results also demonstrate that the performance of both filters is somewhat improved when the ensemble priors are adaptively inflated to offset the negative effects of systematic errors.

1. Introduction

Subseasonal-to-seasonal forecasting at 2-week to 12-month lead times has become an increasingly relevant research area for the environmental modeling community due in large part to a stronger societal demand for reliable information about high-impact climate events (e.g., floods, droughts) further in advance of their onset. Several features in the Earth system have been investigated to determine to what extent they may be exploited as sources of predictability. One such source is the land surface, which directly modulates atmospheric flows, as indicated by the significant moisture and energy fluxes at the land-air boundary. Soil moisture variability can be regarded as a proxy for these interactions (Koster and Suarez 2003, Seneviratne et al. 2010). Earlier studies have shown that satellite and in-situ soil moisture observations, when assimilated into a land surface model, improved the representation of soil moisture and other land surface states for short timescales (a few days) (Draper and Reichle 2015); within a coupled land-atmosphere modeling framework, improved land surface estimates resulted in more realistic evapotranspiration and precipitation (Koster et al. 2002). Earlier work has also demonstrated that enhanced soil moisture initialization has similar implications for the reproduction of large-scale atmospheric modes with anomalies persisting on intradecadal and multidecadal timescales (e.g., Kuenzer et al. 2009, Ford et al. 2016). Given a fixed set of soil moisture observations, we wish to examine how much useful information may be extracted by data assimilation systems having different levels of complexity.

Modern land surface assimilation has roots stemming from the Kalman filter (Kalman 1960), a state estimation method that takes into account the input uncertainties. The Kalman filter analysis is optimal if (i) the observation error and forecast error are mutually uncorrelated, unbiased, and white (temporally uncorrelated), (ii) the model state is linearly related to the observations, and (iii) the model dynamics are linear. Land surface models are driven by nonlinear processes, which suggests that an assimilation system would be more effective at producing improved state estimates if it did not rely so heavily on linearity assumptions.

The extended Kalman filter (EKF; Evensen 1992) explicitly allows for the first-order treatment of weakly nonlinear model processes. The EKF requires the state-dependent linearization (or tangent linear model; TLM) of the model operator to propagate error estimates forward in time. The code in operational land models is not differentiable analytically (due to, for example, the presence of conditional statements), so the linearization must be determined numerically. This additional computational expense increases with the number of model variables that are included in the analysis state vector (Reichle et al. 2002a). Deciding which variables to include and which to disregard when performing this computation requires a substantial amount of application-specific expertise to ensure that the resultant TLM approximation is useful. The EKF’s use of closure approximations and its assumption of linear error growth may still lead to filter failure if nonlinearities are strong enough. One could simply increase the frequency at which the state update (or analysis) is performed or revise the TLM to include higher-order terms (e.g., Miller et al. 1994). Such engineering has been shown to mitigate the threat of filter instabilities; however, the enhanced filter performance may not be enough to justify the extra computational burden.

The ensemble Kalman filter (EnKF; Burgers et al. 1998) entirely avoids the TLM and allows for the fully nonlinear propagation of the forecast errors. This is achievable because the complete probability distribution of the errors is approximated with a finite ensemble of model states at each update step. An added benefit of the EnKF is that it can more readily account for dynamic horizontal forecast error correlations (difficult for the EKF due to computational reasons) and permits the more careful treatment of model error. The EnKF has been shown to be capable in a broad range of applications (Houtekamer and Zhang 2016). In the context of soil moisture state estimation, it has been shown that the EnKF performs at least as well as the EKF when controlling for computational effort (Reichle et al. 2002a). But even though the ensemble of the EnKF contains information about higher-order statistical moments that is advanced between analyses, the actual update equation explicitly considers only its mean and covariance, making the EnKF suboptimal for problems dominated by nonlinearity and/or non-Gaussianity.

Many other advanced data assimilation algorithms have been introduced that may be beneficial for approaching the nonlinear problems highlighted thus far. Such methods include reduced-rank filters (Verlaan and Heemink 1997), particle filters (Gordon et al. 1993, Xiong et al. 2006, Nakano et al. 2007), and the increasingly popular and broad suite of hybrid methods that consider input from both sequential and variational components when performing the analysis. Of particular interest here is the rank histogram filter (RHF) of Anderson (2010), which is similar to the EnKF but does not assume that the ensemble prior and observation distributions are normal.

It has been demonstrated that the RHF can be a competitive alternative to the EnKF even for global numerical weather prediction problems (Anderson 2010). The models used in these earlier examples have typically had dynamics meant to mimic the chaotic nature of a geophysical fluid such as the atmosphere or the ocean (Zhang et al. 2009). Land surface models, in contrast, are dissipative – any ensemble of initial perturbations will collapse toward a single common trajectory with time if something is not done to forcibly maintain its spread (Reichle et al. 2002b). The dynamical regime of the model undoubtedly affects the performance characteristics of any ensemble data assimilation method. Therefore, a more comprehensive assessment of the RHF is needed that explores how the RHF behaves in a land surface application. The main objective of the present paper is to examine the performance differences between the EnKF and the RHF for soil moisture estimation within the context of a synthetic experiment using the NASA Catchment land surface model. A secondary objective is to assess the impact of adaptive error covariance inflation (section 2c) on the performance of the EnKF and RHF. We begin in Section 2 with a review of the two ensemble filter methods. Section 3 discusses the details of the land surface model and outlines the experimental design. Section 4 presents results, and Section 5 provides a conclusion.

**2. Review of ensemble filtering methods**

a. Ensemble Kalman Filter (EnKF)

Given an ensemble of model forecasts, , , each of dimension , the EnKF updates the state, *xi*, of each ensemble member in the following manner

(1)

where the  and superscripts are the forecast (prior) and analysis (posterior) quantities, respectively, corresponding to values before and after the update of the state estimate. Each member is given a unique observation vector, , obtained from a random draw from a normal distribution having a mean equal to the actual unperturbed observation vector, , and a covariance equal to the user-defined observation error covariance matrix,. The forward operator, , computes the observation-space equivalent of the model state. The innovation increment, , is weighted by , the Kalman gain (or influence matrix), which is given explicitly in Eq. (2)

|  |  |  |
| --- | --- | --- |
|  |  | (2) |

where is the forecast error covariance matrix. The ensemble formulation of the analysis problem circumvents the need to ever represent  explicitly in computer memory as a matrix (Keppenne 2000).  can now be diagnosed as the sample covariance of the ensemble forecast as follows

(3)

where  is the normalized forecast perturbation matrix; the columns of which are formed by the normalized deviations of each ensemble member forecast from the ensemble mean

(4)

(5)

The mean of the posterior ensemble, , is the optimal state estimate, which, if we assume that all operators are linear, can also conveniently be written as follows in terms of the mean ensemble forecast and mean observation values.

(6)

Note that in Eq. (6), is the mean of the ensemble of , not the value of the actual observation, .

*b. Rank histogram filter (RHF)*

The RHF is an ensemble filtering method that can accommodate any distribution for the ensemble prior or observation. Figures 6 and 7 in Anderson (2010) provide useful schematics that show how the RHF works. Here, we will briefly describe its update procedure.

To construct the ensemble prior, the RHF applies the forward operator, , to each member’s state vector, , to compute the ensemble estimates of the observation. The values are then arranged in increasing order to partition the real number space into  intervals. The RHF assumes that the prior’s cumulative density in each of these intervals is equal to . Between members, the probability of the distribution is made uniform (constant); in the unbounded regions (the tails), the probability is set to match that of a Gaussian with variance equal to the ensemble variance and mean chosen such that the  cumulative density condition is met. Here, the observation likelihood is specified as Gaussian with mean  and variance . To make the update computation less expensive, the observation likelihood is approximated as a piecewise linear function in between the values of each member in the prior distribution (the tails are left unchanged as partial Gaussians). The prior and likelihood are then multiplied pointwise. After multiplication, the posterior ensemble is chosen from the resulting distribution by determining the points on the real number line that have cumulative densities satisfying for . Finally, each member’s analysis increment (posterior minus prior) is regressed linearly onto all components of its corresponding analysis state vector.

Available observations are assimilated serially, one at a time, which assumes that the observation errors are independent. Given the use of Gaussian tails, the range of possible state values is not constrained in the regions of the distributions outside of the ensemble. This is obviously not appropriate for bounded quantities. It is possible to force boundedness on the posterior ensemble as was mentioned by Anderson (2010), but the subsequent regression may still result in physically impossible values for the unobserved quantities. (This problem is solved by the marginal adjustment RHF of Anderson (2020), a recent modification of the RHF.) As is discussed in Section 3.c, we impose extra-analysis physical consistency checks to reduce the harmful effects of unbounded posterior distributions.

In summary, perhaps the most important feature of the RHF is that, unlike the EnKF, the RHF does not restrict the ensemble distribution of model states (prior or posterior) to be Gaussian. This novel aspect of the RHF is the focus of the present study.

*c.* *Addressing systematic error with inflation*

By design, the EnKF updates the magnitudes of the ensemble covariances during each analysis such that the norm of the error covariance matrix after the analysis is typically smaller and necessarily no larger than it was before the analysis; the user-defined observation error covariance matrix is usually kept static, meaning that, with successive analysis cycles, more and more of the information provided by the observations will tend to be ignored and the data assimilation system will eventually stop altering the state estimate, a phenomenon known as filter divergence.

Filter divergence can often be avoided via appropriate use of a noise application scheme like the one employed during all of the land model integrations in this study (see Section 3.b). Such a strategy is an inflation method because it helps ensure that the ensemble spread does not get too small. Several different inflation formulations have been proposed in the literature, each being shown to generally improve ensemble representativeness and performance (e.g., Anderson and Anderson 1999, Whitaker and Hamill 2012).

A common limitation of these inflation approaches is their use of static inflation parameters. A given parameter set may adequately inflate the ensemble for one time period but may over- or under-inflate the ensemble for another. Obtaining a set of manually tuned time-varying inflation parameters for large models is not feasible, prompting the introduction of adaptive techniques that leverage the real-time ensemble statistics to determine the appropriate inflation values. Such methods commonly rely on the ensemble’s innovation (e.g., Anderson 2007, Reichle et al. 2008). In particular, the inflation method of Anderson (2009a) computes the Bayesian update of a multivariate Gaussian inflation distribution after considering the impact of an observation on the expected innovation variance, which can be described as follows if considering the case of when a single observation is to be assimilated

|  |  |  |
| --- | --- | --- |
|  |  | (7) |

Here, , , and , the inflation factor, are scalar quantities. It is further assumed that the observation errors and forecast errors are uncorrelated. Note that since is assumed to be a normally-distributed variable, it can take on any real value so long as the expression above is satisfied; it could even be negative if the observation error is especially large (the restriction of to nonnegative values was externally imposed). Once an observation becomes available, the inflation factor in Eq. (7) is used to update the prior estimate of the inflation factor, the result of which is multiplicatively applied to the ensemble covariance. The adjusted ensemble then assimilates the observation in the usual manner. To prevent the auxiliary update of the inflation factor from diverging, it is damped toward unity after the analysis is performed

|  |  |  |
| --- | --- | --- |
|  |  | (8) |

where is a damping factor.

As stated previously, Anderson (2009a) models the inflation factor using a normal distribution, where the forecasted inflation distribution is the most recent (damped) posterior inflation distribution. This study uses the adaptive inflation algorithm of Gharamti (2018), which improves on that of Anderson (2009a) by instead modeling the inflation factor using an inverse-gamma distribution, yielding a more stable implementation that completely avoids negative values, and better combats overdispersion.

**3. Land model and experiment setup**

*a. Land surface model*

The land surface model used here is the Catchment model of the Goddard Earth Observing System (GEOS) modeling and assimilation framework (Koster et al. 2000). The catchment-based approach considers the effects of subgrid heterogeneity on the horizontal structure of hydrological land surface processes. The model has been shown to be viable when used standalone (Reichle et al. 2019) as well as when coupled to a general circulation model (Gelaro et al. 2017).

The model prognostic variables include soil temperature and heat content, snow water equivalent, and soil water. The latter are water excess and deficit variables that measure the total amount of water stored in the catchment and the departure of a layer’s water content from the equilibrium vertical profile. The excess and deficit variables include the surface excess, root zone excess, and catchment deficit, which correspond to three nested layers of 0-5 cm, 0-100 cm, and 0 cm to the (spatially varying) bedrock depth, respectively. These soil water prognostic variables are used to diagnose the corresponding volumetric soil moisture values – surface moisture content (sfmc), root zone moisture content (rzmc), and profile moisture content (prmc). The distribution of rzmc is used to spatially partition a catchment fractionally into three different hydrological regimes: (i) saturated - ground surface completely saturated, (ii) transpiration-sufficient – surface not saturated but transpiration proceeds without stress or severe water limitation (moisture-stress evapotranspiration is not supported), and (iii) wilting - soil moisture is below the wilting point and therefore too dry for any transpiration to occur. Different parameterization schemes for the surface energy balance and soil water transfer are used in the different hydrological regimes, a distinction that is ignored in more conventional, layer-based land surface models.

*b. Ensemble spread and stochastic perturbations*

To maintain the ensemble spread, each member is treated during the model integration with regular noise applications (Reichle et al. 2002a, Reichle and Koster 2003). The noise varies according to an autoregressive model of order 1 (AR(1)), which can be described as follows

(9)

(10)

(11)

 is a random noise sample from a standard-normal distribution,  is the interval between noise applications, and  is the correlation timescale for the component of the state that will ultimately be perturbed. The formulation in Eq. (9) – Eq. (11) explicitly enforces a temporally correlated error. Once  is obtained, the state, , is modified according to the following logic:

*If*  *must be nonnegative, then*

(12)

*otherwise*

(13)

The expression in Eq. (12) uses the inverse transform method, which makes it possible to multiplicatively apply  after converting it to a sample of a lognormal distribution with mean 1 and standard deviation . Multiplicative noise application is necessary for variables that cannot be negative (e.g., downwelling shortwave radiation, precipitation).  is a fixed parameter used to scale the variance of the unit-variance noise of Eq. (9) to a level consistent with the expected errors of the perturbed state variable. After the perturbations are applied, the updated quantities are passed through a physical consistency check procedure, and, if necessary, are adjusted to remain within the physical limitations encoded in the model; this same check is applied at every call to the model core, regardless of whether a perturbation scheme is being used.

*c. Experiment design*

The domain for the synthetic experiment consists of 18 globally-distributed locations mainly clustered in North America and Europe (Reichle et al. 2019). The locations were selected because each is also a validation site for the NASA SMAP soil moisture products. As a group, the sites have a wide array of soil and vegetation properties that provide challenges to real-world land-assimilation efforts. Site selection was made in anticipation of future work that will rely on the in-situ measurements for validation of real-data assimilation experiments (please see Appendix A for validation site details). The synthetic experiment is carried out at each site independently from the others. MERRA-2 data (Gelaro et al. 2017) is used for the surface meteorological forcing inputs. The Catchment model parameters and boundary conditions are those used in Version 4 of the SMAP L4\_SM system (Reichle et al. 2019).

Conducting the experiment in a synthetic environment allows for the fine-grained control of every aspect of the experiment. The potential usefulness of the results obtained in a real-world context is made more apparent when the parameters of the experiment are set in a sensible way. An important consideration is how to simulate model error, which can be difficult to characterize given that it is unknown in the real world. Studies focusing on land data assimilation have tried to account for model error in a synthetic framework by using distinct meteorological forcing datasets and/or model parameters for the truth and data assimilation runs (e.g., Reichle and Koster 2003). Here, we follow a similar rationale and prescribe model error by requiring that the truth be generated using meteorological forcing for the years 2000-2009 while the data assimilation experiments are run with forcing corresponding to the years 2010-2019; additionally, the data assimilation ensembles were initialized with the 1 January 2010 model states taken from the end of the experiment period; stated differently, we are trying to estimate the truth given intentionally incorrect (or error-prone) meteorological forcing and initial conditions. A schematic of the approach is provided in Figure 1 and is explained in more detail below.

Diagram

Description automatically generated

Fig. 1. Schematic of the synthetic experiment showing the 10-year offset of the meteorological forcing used to drive the truth run (single-member) and that used to drive the assimilation run (100-member ensemble).

The Catchment model initial conditions were obtained as follows. Firstly, the model state was spun-up from a cold start by running the model four times through the 20-year period from 1 January 2000 through 31 December 2019. After this spinup, the model was run from 1 January 2000 through 31 December 2009. This 10-year period serves as the truth for the assimilation experiments (Figure 1). Daily sfmc observations of the truth were generated by perturbing the true value with a random sample of normally distributed noise with mean zero and standard deviation equal to 0.03 m3 m-3. The standard deviation was chosen to be roughly equal to typical errors in SMAP soil moisture retrievals (Colliander et al. 2022).

Once the observations and truth states had been recorded, the model was reset to 1 July 2009 at which point a 100-member ensemble was initialized, each with an identical copy of the truth state at that time (Figure 1). To spin up the ensemble, each member was run for another six months through 31 December 2009 with a unique set of AR(1) perturbations applied to a subset of the meteorological forcing and prognostic variables (see Section 3.b). The perturbations were samples from either lognormal or normal distributions with mean 1 and 0, respectively. The timescale of the AR(1) correlation for each perturbation time series was chosen to provide a physically meaningful error timescale; there was no prescription of cross-correlated errors (see Table B1 in the appendix). The ensemble was then run from 1 January 2000 through 31 December 2009, using the (intentionally wrong) 2010-2019 meteorological forcing data and 1 January 2010 ensemble of initial conditions. The ensemble trajectory during this ten-year period was recorded as the open loop simulation. The open-loop establishes the base-level performance of the modeling system without the benefits of soil moisture data assimilation.

Finally, the data assimilation experiments consist of ensemble simulations that are identical to the OL but with the assimilation of the synthetic daily observations of surface soil moisture. To mimic the quality controls in place for real-world satellite soil moisture retrieval algorithms, soil moisture observations were not assimilated if any one of the ensemble members indicated that snow was on the ground or that the soil was frozen. After an analysis was performed, the updated quantities were passed through a physical consistency check procedure, and, if necessary, were adjusted to remain within the physical limitations encoded in the model.

There were four assimilation experiments performed, each one highlighting a different filter/inflation combination – (i) EnKF without adaptive inflation, (ii) RHF without adaptive inflation, (iii) EnKF with adaptive inflation, (iv) RHF with adaptive inflation. (Note that from here onward, ‘inflation’ is used to refer to the adaptive inflation algorithm of Gharamti (2018) described in Section 2.c, and not to the general application of model prognostics and forcing perturbations outlined in Section 3.b). Each case used the exact same Catchment model, meteorological forcing, initial ensemble, and member-specific temporal sequence of AR1 perturbations and had access to the same set of sfmc observations.

The update of the state and of the inflation parameters is handled wholly by NCARs Data Assimilation Research Testbed (DART; Anderson et al. 2009b), which allows for any discretized computer model to be interfaced with a suite of sequential ensemble data assimilation methods and common inflation techniques. In these experiments, the prior state vector provided to DART is not the entire model state vector but rather a subset consisting of only the surface excess and root zone excess model prognostic variables, which matches the EnKF state vector of the SMAP L4\_SM system (Reichle et al. 2019; their section 2.2.2). The adaptive inflation algorithm was set to inflate only the ensemble prior; the mean and standard deviation of the inflation distribution were initialized to 1.0 and 0.6, respectively, and were not allowed to be negative; before the inflation factor was updated and applied to the ensemble, the difference between the inflation factor and 1 was multiplied by a damping constant (equal to 0.9), the result of which was then added to 1 (see Eq. (8)).

*d. Validation*

1) Performance metrics and normalized information contribution

The impact on model skill of assimilating soil moisture observations was evaluated using several performance metrics, namely the anomaly correlation coefficient (), the mean absolute bias (), and the unbiased root mean square error (). Each metric was applied to the three diagnostic soil moisture content variables output by the model (sfmc, rzmc, prmc).

The computation for the is identical to that of the regular Pearson correlation coefficient but with the extra step of first removing the mean seasonal cycle from each of the two time series. The mean seasonal cycle was determined using the mean trajectory of the open loop (no meaningful difference in performance was seen when the climatologies of the respective experiments were used instead of the open loop mean). On a given day, the average value across all available years was taken from the mean open-loop run; the resulting time series (of 365 daily values) was then smoothed with an equal-weighted, 15-day average centered on the day of interest; leap-day quantities (February 29th) were discarded in all input time series prior to the calculation of the anomaly correlation coefficient. The was determined by taking the absolute value of the difference between two time series, then taking the average value of the result. The computes the root mean square error of two time series after having first subtracted the respective mean values from each.

Confidence intervals for the were computed using a Fisher-Z transform of the anomaly correlation values. As in Draper et al. (2012), the effective sample size of the input time series was reduced to account for inherent autocorrelations in the data. This consideration leads to progressively larger uncertainties for moisture variables with more temporal memory. The confidence intervals for the and were computed using a two-tailed t-test with 10,000 replicates.

The open loop estimates provide a baseline skill level without assimilation. For each performance metric, , the skill improvement gained from the assimilation relative to the open loop baseline can be measured as a normalized information contribution (; Kumar et al 2009)

(14)

Here, is the value corresponding to perfect performance for the given . For example, if looking at , then  would be equal to 1. The measures how well the analysis does when compared to the open loop run, where positive/negative values imply that data assimilation helped/hurt model performance relative to the open loop. A equal to 1 (the highest possible value, regardless of the particular ) means that the analysis exactly matched the truth over the entire experiment, while a  equal to 0 means that the analysis performed just as well as the open loop (i.e., assimilating observations had no effect). A negative-valued indicates that assimilating observations led to a decline in skill. Here, the and quantities are regarded as being independent, meaning that an expression for the error in estimates can be obtained by summing the uncertainty components of Eq. (14) in quadrature.

(15)

The confidence intervals for the NIC scores are computed as the uncertainty in Eq. (15), , multiplied by 2. The difference in any two values is considered statistically insignificant if the respective confidence intervals overlap.

2) Normalized Innovations

Innovation statistics will also be used to help characterize filter performance. The EnKF assumes that the observations and model estimates are unbiased and uncorrelated, meaning that a histogram of the innovations, after having been scaled with the inverse of the expected error standard deviation , ought to closely approximate a standard normal distribution (zero mean and standard deviation of one). Given the synthetic nature of the experiment, the observation error, , is known exactly (set as a scalar quantity equal to the square of 0.03 m3 m-3). So a normalized innovation distribution for the EnKF with standard deviation not equal to 1 suggests that the ensemble’s simulated errors of the observation are incorrect, which can only be due to sampling error, nonlinearity and/or a violation of the EnKF’s basic assumptions. The RHF does not make bias or correlation assumptions, so there is not a known target distribution for its corresponding normalized innovation histogram, but it will regardless be presented alongside that of the EnKF to further highlight any departure of the soil moisture assimilation problem from the EnKF’s underlying assumptions.

**4. Results and discussion**

*a. Soil moisture*

Improved accuracy of the soil moisture state has the most potential to improve near-surface atmospheric flows in regions with strong land-atmosphere coupling (Koster and Suarez 2003). One such location is the Great Plains of the United States, where the Little Washita test site is positioned. For portions of the analysis that follow, we focus our attention on the Little Washita site, given that it is representative of where we would expect to see maximal impact of assimilating soil moisture observations. In some cases, we also illustrate results for the sites at Carman, St. Josephs, and Tonzi Ranch to highlight interesting aspects. We did not detect any meaningful general dependence of the assimilation performance on site-specific environmental features (soil properties, land cover type, climate) or the ensemble itself (e.g., innovations statistics). If such a signal is to be detected, a much larger number of sites is likely needed, which is beyond the scope of the present study.

Figure 2 shows the sfmc and rzmc time series at the Little Washita validation site during June, July, and August (JJA) 2009, the final year of the experiment period. The same plot was created for different years (not pictured) to see if there was a trend in the difference between the experimental configurations that were tested (e.g., EnKF versus the RHF), but no such trend was found; hence the year 2009 was chosen arbitrarily for presentation purposes. Pictured are the truth trajectory and the ensemble mean values of the open loop and data assimilation runs; also shown are the 100 individual members of the open loop ensemble (grey lines). The peaks in the soil moisture trajectories coincide with precipitation events and are typically followed by relatively slower dry-down periods, which are marked by little to no precipitation. Note that the open loop ensemble spread is small during precipitation events and increases during the dry-down periods. For both sfmc and rzmc, the mean of the open loop at this site is generally drier than the truth at the Little Washita site, meaning that the assimilation of sfmc observations should add moisture to correct the systematic dry bias coming from the model error specification (the incorrect meteorological forcing). Qualitatively, this was achieved for each of the four data assimilation runs as indicated by the smaller differences between their ensemble mean values and the truth (relative to the difference between the open loop and the truth). Also noticeable are the modest improvements in correlation with the truth. None of the filters responded particularly well to surface soil moisture observations following relatively strong precipitation events during this time period. There seems to be rapid peaking at the onset of the wetting episodes but not of a high enough magnitude. This last point can be explained by recalling that the analysis is effectively a weighted average of the model simulation (ensemble mean) and the assimilated observation, the weights being inversely proportional to the variance of the respective errors. The observation error was intentionally held fixed for the entirety of the experiment period whereas the ensemble estimate of the model forecast error (the ensemble spread) was changing over time. During heavy precipitation events, the model ensemble was increasingly confident that the soil was nearing saturation as evidenced by relatively small values of ensemble spread. As a result, the observation information is not given as much relative weight in the averaging process, sometimes leading to insufficient soil wetting. The severity of this effect is mitigated as the precipitation ceases and the soil dries, allowing the ensemble spread to increase to more climatologically normal values (boosting responsiveness of the ensemble to information introduced by the assimilated observations).

Chart, histogram

Description automatically generated

Fig. 2. Time series of volumetric soil moisture content (m3 m-3) for (a) sfmc and (b) rzmc at Little Washita during JJA 2009. (OL: Open Loop, EnKF: Ensemble Kalman Filter, RHF: Rank Histogram Filter, EI: EnKF with adaptive inflation enabled, RI: RHF with adaptive inflation enabled.)

Figure 3 shows the average (Figure 3a), (Figure 3b), and (Figure 3c) metrics for sfmc, rzmc, and prmc, along with their corresponding 95% confidence intervals. The EnKF had a positive for each of the 3 soil moisture variables, indicating a systematic improvement over the open loop. The  scores are ~0.51 for sfmc, ~0.41 for rzmc, and ~0.42 for prmc. The for the EnKF was positive for sfmc and rzmc with respective values of ~0.16 and 0.03 but was near-zero for prmc (~ 0). The assimilation of soil moisture observations with the EnKF improved model skill in terms of the bias for sfmc but seems to have led to no real performance benefit for the deeper soil moisture reservoirs of the catchment (rzmc and prmc) relative to the open loop. The EnKF scores are all positive with a value of ~0.32 for sfmc and a more modest ~0.25 for both rzmc and prmc. Generally, the EnKF provides a clear net benefit over the open loop for model estimates of soil moisture.

Graphical user interface, chart

Description automatically generated

Fig. 3. Normalized information contribution metrics for (a) anomaly correlation coefficients, (b) mean absolute bias, and (c) unbiased root mean square error (ubRMSE). Metrics are averaged across all 18 validation sites. The entire 10-year experiment period was considered. Confidence intervals pictured are 95%.

The RHF had a of ~0.52 for sfmc, ~0.41 for both rzmc and prmc, performing significantly better than the EnKF for sfmc and just as well for rzmc and prmc. For , the RHF showed scores of ~0.23, ~0.12, and ~0.08 for sfmc, rzmc, and prmc, respectively, performing significantly better than the EnKF in each case. Each of the values for the RHF were positive, meaning that the RHF had a positive effect on model performance in terms of model bias relative to the open loop. The RHF had values of ~0.33 for sfmc and ~0.25 for both rzmc and prmc, effectively identical to those for the EnKF. In all cases, values for the RHF were positive and never significantly worse than those for the EnKF. The most notable difference between the two filters is seen for the bias, where the RHF shows itself to be the significantly better filter, increasing the for each of the three soil moisture variables by ~0.05 over the EnKF.

With the addition of adaptive inflation, the RHF achieved values at least as good as its non-inflated variant, but in only one case was that improvement statistically significant ( for sfmc). Similarly, the use of adaptive inflation with the EnKF always improved performance but significantly so in only two cases ( for sfmc and rzmc). In terms of , the inflated RHF performed significantly better than all of the other filter experiments for sfmc, but such a difference did not translate to the unobserved soil moisture variables (rzmc and prmc), for which no one filter configuration showed itself to be best. The and subpanels illustrate that inflation generally does improve the *MAB* and *ubRMSE* values but never in a statistically significant sense.

The value magnitudes across all experiments and metrics indicate that the impact of soil moisture assimilation on skill was always positive and never negative. The largest improvements over the open loop were observed for the anomaly correlations ( ranging from ~0.40 to ~0.53) and the smallest skill improvements were for the bias ( ranging from ~0 to ~0.24). The metrics show that the RHF performed somewhat better than the EnKF and that adaptive inflation mostly results in minor skill improvements for both filters compared to the respective noninflated filter variants.

*b. Ensemble representativeness*

The rank histograms for the open loop and for each of the four assimilation experiments are shown in Figure 4 for a representative set of locations, including St Josephs, Carman, Tonzi Ranch, and Little Washita. The rank histograms were created by considering every prior ensemble used over the experiment period, then ordering (or ranking) all the members of this augmented set from low to high according to the surface soil moisture state value; these bins are then populated using all the available observations (over the 10 years). The bins corresponding to each experiment are not identical, but the observations used to populate the bins are the same. To reduce noise in the illustrations, the counts of adjacent bins were averaged (the smoothing has no impact on the major features of the histograms). It is important to note that a rank histogram says nothing about whether the ensemble or the observations are representative of the actual system variable being estimated (soil moisture in this case). It simply provides a way to gauge if an ensemble is representative of the space spanned by the observations. One-to-one correspondence between the two would be evidenced by a flat line for a given experiment. None of the ensembles in Figure 4 has a histogram with such a horizontal profile. At St Josephs (Figure 4a), the rank histogram for the open loop shows some underdispersion (as evidenced by the gentle ‘U’ shape) and a forecast ensemble that is biased wet compared to the observations. Data assimilation does not seem to correct for either issue in any meaningful way as evidenced by the close tracking of the filter and open loop histogram profiles; in fact, data assimilation seems to exaggerate the bias (note the higher left edge of the histograms of the assimilation experiments compared to that of the open loop).

All the rank histograms at Carman (Figure 4b) also show biased and underdispersive ensembles but instead with forecasts that tend to be too dry with respect to the observations; there is no clear distinction between the open loop and the assimilation experiments at this location, except, again, at the left-most edge where the assimilation experiments show slightly wetter ensembles than for the open loop.

The open loop rank histogram at Tonzi Ranch (Figure 4c) does not show an obvious global dispersion problem, but it does exhibit a forecast ensemble that is too dry compared to the observations. The assimilation experiments reduce the severity of the bias (given by the relatively flatter outer left edge).

At Little Washita (Figure 4d), the rank histogram profile for the open loop indicates a forecast ensemble biased wet compared to the observations; but the filter experiments show more general underdispersion with little to no bias (as evidenced by their roughly symmetrical ‘U’ shapes). The results imply that data assimilation leads to small changes in the ensemble’s representativeness of the observation space; most of these changes are concentrated on the left side of the histogram figures, where assimilating the soil moisture observations clearly leads to generally wetter ensembles. The main aspects of the rank histograms do not vary coherently with filter type or whether inflation was used, nor do they seem to depend on climatology (dry versus wet).

Chart, line chart, histogram

Description automatically generated

Fig. 4. Rank histograms of SFMC observations for each assimilation experiment at (a) St. Josephs, (b) Carman, (c) Tonzi Ranch, and (d) Little Washita. All 10 years of available data were considered.

Figure 5 shows time series of the open loop and ensemble prior standard deviations of surface and root zone soil moisture at the same four locations during JJA 2009. At St Josephs (Figure 5a), the standard deviation for the open loop ensemble had a JJA 2009 mean value of 0.026 m3 m-3, ranging from 0.01 to 0.042 m3 m-3 over the 3-month time period. The mean standard deviations were somewhat lower for the RHF (0.024 m3 m-3) and higher for the EnKF (0.29 m3 m-3), both having roughly similar ranges of ~0.01 to ~0.045 m3 m-3. The time series for the inflated cases were identical to those of their respective noninflated counterparts. The variability of the time series does not show a clear dependence on experimental configuration. At Carman (Figure 5b), the time series for the open loop standard deviations shows a mean value of approximately 0.031 m3 m-3 and a 3-month range of 0.012 to 0.045 m3 m-3. The RHF standard deviation mean is 0.024 m3 m-3 while that for the EnKF is 0.026 m3 m-3. Inflation does not impact the mean for the EnKF but does lead to a small reduction for the RHF to a value of 0.023 m3 m-3. Here, the time series corresponding to the data assimilation experiments do not track the open loop case as closely (e.g., in early August) as was seen for St Josephs. The open loop at Tonzi Ranch (Figure 5c) had a mean standard deviation of 0.062 m3 m-3, with a range of 0.037 to 0.078 m3 m-3. The spreads in the ensemble priors are significantly lower than that for the open loop, with mean values of 0.046 m3 m-3 for the EnKF and 0.042 m3 m-3 for the RHF (the values were the same for the respective inflated variants of each filter). At Little Washita (Figure 5d), the ensemble standard deviations for the open loop showed a mean value of 0.026 m3 m-3 (ranging from 0.010 to 0.038 m3 m-3). The EnKF variants are relatively less dispersed, having mean values of 0.032 m3 m-3 and 0.033 m3 m-3 for the non-inflated and inflated forms, respectively. The RHF variants track each other closely, with means of 0.027 m3 m-3 and a common range of 0.011 to 0.046 m3 m-3.

Generally, the RHF tends to have smaller spreads than the EnKF (the same is true for the relative differences between the inflated variants); the EnKF variants, as we have previously shown, exhibit positive performance gains over the open loop, but the large standard deviations relative to the RHF can be interpreted as evidence that the EnKF successfully compensates for the slight inappropriateness of its assumptions by boosting the ensemble spread.

The ensemble spread in the inflated EnKF closely tracks that of the non-inflated EnKF, and the same is true for the RHF variants. In fact, at St. Josephs and Tonzi Ranch there is no discernible difference in ensemble spread between the inflated and non-inflated filter versions across the entire JJA 2009 period (Figure 5a,c), and only very small differences are seen at Carman and Little Washita (Figure 5b,d). It remains unclear why the inflation does not lead to larger differences in ensemble spread during JJA 2009.

Of note is that the standard deviations of the open loop are not necessarily always greater than or equal to those of the noninflated filter configurations, which seems nonsensical if one neglects the boundedness of the sfmc problem being considered here; when an ensemble is especially wet or dry, its spread tends to collapse as the mean approaches the extremes. One example of this behavior can be seen clearly at Little Washita by comparing Figures 1a and Figure 5d, where the open loop shows significantly smaller standard deviations than the filters during time periods for which the filter ensemble sfmc values are moderate and the open loop ensemble has a relatively much wetter soil column. Across sites, it was also observed that inflation values tended to be closer to one during periods when the mean of the open loop was closer to the truth trajectory.

Chart

Description automatically generated

Fig. 5. Time series examples of volumetric soil moisture content (m3 m-3) standard deviation of data assimilation experiment ensembles (priors) and of the open loop ensemble for JJA 2009 at (a) St. Josephs, (b) Carman, (c) Tonzi Ranch, and (d) Little Washita.

*c. Measuring non-Gaussianity*

Figure 6 illustrates the normalized innovation distributions of the assimilation experiments at the same representative locations as in Figure 4 and over the entire 10-year experiment period. Distributions for the EnKF and RHF are shown in the left column, and distributions for the filter variants with adaptive inflation are shown in the right column. For reference, the dotted line pictures the unit normal distribution (mean = skewness = 0, variance = 1), which is the distribution of the normalized innovations of the EnKF when all its assumptions are satisfied (see Section 1). Each of the distributions in Figure 6 is biased and skewed, clearly showing the systematic departure of the soil moisture assimilation problem considered here from those assumptions.

At St Josephs (Figures 6a and 6b), the dry biases for the non-inflated filters are the same at -0.17; the standard deviations are greater than one (1.40 and 1.46 for the EnKF and RHF, respectively), suggesting that the ensembles are underestimating the expected background error covariance; the skewness values for the EnKF and the RHF are nearly identical at -0.41 and -0.40, respectively. The negative skewness values imply that the probability density tends to be somewhat more concentrated at the higher end of the distribution when compared to a Gaussian (the relationship is reversed for positive skewness). The use of inflation (Figure 6b) didn’t lead to any meaningful change in the mean, standard deviation, or skewness for either filter; whatever difference in statistics was observed seems to depend more on filter type (RHF vs EnKF) than it does on use of inflation.

Carman (Figures 6c and 6d) instead shows wet biases (0.15 and 0.18 for the EnKF and RHF, respectively); the standard deviations are 1.35 for the EnKF and 1.39 for the RHF; also shown are negative skewness values with -0.34 and -0.27 for the EnKF and RHF. As at St Josephs, inflation leads to negligible change in the distribution statistics considered here.

At Tonzi Ranch (Figures 6e and 6f), the innovations indicate a systematic wet bias relative to expectation with values of 0.13 and 0.15 for the EnKF and RHF; the respective standard deviations are 1.12 and 1.16, with skewness values of -0.38 for the EnKF and -0.34 for the RHF. The normalized innovation statistics for the corresponding inflated cases are nearly identical. As at Carman and St Josephs, the differences in skewness are more dependent on filter type than on whether inflation was used.

The bias of the normalized innovations at Little Washita was 0.04 for the EnKF and 0.08 for the RHF (observations tended to be wetter than the forecasts); the EnKF had a standard deviation and skewness of 1.67 and -0.74, respectively, with the RHF having similar values (standard deviation of 1.76 and skewness of -0.67). Consideration of inflation (Fig. 6h) shows little to no change for the RHF, which had a reduction of 0.01 in the mean and one of 0.02 in both the standard deviation and skewness; somewhat larger differences are present between the EnKF variants where now the inflated case shows an increase of 0.03 in the bias and a 0.02 reduction in the standard deviation of the normalized innovations distribution.

When looking at similar plots for all 18 locations (not pictured), there is no apparent relationship between the mean, bias, and skewness. Also of note is that, for each filter, inflation always led to a skewness magnitude that was less than or equal to that for the corresponding noninflated case. The RHF, inflated or not, tended to lead to skewness values closer to zero than the EnKF. These results imply that adaptive inflation was more likely to generate ensembles with relatively dry outlier members; these drier outliers were also more common for the EnKF than for the RHF. The results also show that the standard deviations of the normalized innovations distributions were generally larger for the RHF than for the EnKF, irrespective of inflation usage, a finding that is in line with the tendency of the RHF to have smaller ensemble spreads than the EnKF (as previously discussed).

Shape, polygon

Description automatically generated

Fig. 6. Representative set of normalized innovations distributions at (a,b) St Josephs, (c,d) Carman, (e,f) Tonzi Ranch and (g,h) Little Washita for the (a,c,e,g) non-inflated and (b,d,f,h) inflated filter variants. Histograms were computed over the entire assimilation period (1 January 2000 – 31 December 2009). Quantities listed in key show the mean, standard deviation (sdev) and skewness of each distribution.

To better quantify and understand the effects of non-Gaussianity on filter performance, we use the Kolmogorov-Smirnoff (KS) statistic to test if a sample comes from a normal distribution. This is achieved by comparing the cumulative distribution functions of the sample and of a continuous normal distribution having the same mean and variance. The KS statistic is defined as the maximum distance between the two functions across the range of the random variable. Assuming that the sample is sufficiently large, higher KS statistics generally communicate more confidence that the sample distribution does not sample from the reference distribution. Here, we test every ensemble prior that was used as input for an analysis at each location and use a significance level of 5%. We take the null hypothesis to be that the ensemble prior is indeed Gaussian. This analysis was carried out for each of the 18 locations for every ensemble prior during JJA 2009, which is representative of the 10-year experiment period.

The KS values across all locations ranged from ~0.5-0.65, which are all quite large. Every corresponding p-value was orders of magnitude smaller than the target significance level. This suggests that the evidence allows for us to reject the null hypothesis at each of the analysis times in JJA 2009; in other words, we can be confident that none of the prior distributions were Gaussian. When comparing the average KS value for each filter, the RHF had slightly higher values than the EnKF ensemble at all 18 sites. The difference does not depend on whether inflation was used (this distinction is present when making the same comparisons for the posterior ensembles). Also noted was that the mean KS values for the open loop were usually significantly higher than those of the data assimilation experiments. This result highlights how the update procedures used by the two filters lead to the destruction of non-Gaussian information in the ensemble priors, with the RHF maintaining more non-Gaussianity in the ensemble than the EnKF. In Figure 7, Little Washita is used as a representative example of the relative relationships of the KS time series for the open loop and for the ensemble priors corresponding to the four different data assimilation experiments. Here, the KS values for the RHF tend to be somewhat higher than those for the EnKF. The inflated EnKF shows the lowest KS readings. Note the relatively higher KS values for the open loop compared to all the assimilation cases.

Chart, histogram

Description automatically generated

Fig. 7. Time series of Kolmogorov-Smirnov values for the ensemble prior at the Little Washita site during JJA 2009.

**5. Summary and conclusion**

In this paper, we presented results highlighting the differences between the EnKF and RHF when surface soil moisture observations were assimilated in a synthetic experiment. There were four cases: each filter was used with and without an adaptive inflation technique. The synthetic framework allows for more control of the experimental details, namely the better contextualization of model error, and the complete avoidance of representativity issues associated with using complex forward operators. In all the metrics observed, the RHF typically performs on par with the EnKF, but often slightly better. This consistent improvement suggests that the non-Gaussianity present in the problem was significant enough to be exploited by the RHF. Since the adaptive inflation was shown beneficial for both filters, the deficiencies in ensemble spread were indeed relevant to performance and therefore requiring treatment. Further testing in a real-observation context would aid in understanding how the quality of the system error specifications affects the performance gap between the two filters.

The RHF allows for the non-Gaussian representation of the forecast and observation distributions. In our experiments, the observation error distributions were deliberately kept Gaussian to isolate the effects of the forecast-error parameterization. The RHF performed better than the open loop for all the  metrics examined here. This was not the case for the EnKF, which showed near-zero  values for prmc. Evaluation of the KS statistics suggests that the RHF, inflated or not, allows for improved conservation of the ensemble’s non-Gaussian information. Given the results here, we propose that it is worthwhile to further examine the RHF as a potential alternative to the EnKF for more complex land assimilation problems (including using bias corrected filters, modifying a filter for spatially distributed application, or assimilating multiple types of observations).

Acknowledgments.

Emmanuel Dibia was supported by the NASA Fellowship Activity NNH18ZHA003N, Rolf Reichle was supported by the NASA SMAP mission, and Xin-Zhong Liang by NOAA Center for Atmospheric Science and Meteorology (NCAS-M) under Grant NA16SEC481006. Computational resources were provided by the NASA High-End Computing program through the NASA Center for Climate Simulation. The GEOS source code is available under the NASA Open‐Source Agreement (https://github.com/GEOS-ESM).

Data Availability Statement.

MERRA-2 data can be obtained from the Goddard Earth Sciences Data and Information Services Center at https://earthdata.nasa.gov/eosdis/daacs/gesdisc.

APPENDIX

Appendix A. SMAP L4\_SM validation sites

As follows is a figure and a table describing the locations of the 18 SMAP L4\_SM validation sites that were tested in the study.

Map

Description automatically generated

Fig. A1. Map of SMAP L4\_SM validation sites.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Site Name** | **Latitude** | **Longitude** | **Climate Regime** | **MODIS IGBP Land Cover** |
| *1. Remedhus* | 41.42 | -5.37 | Temperate | Croplands |
| *2. Reynolds Creek* | 43.19 | -116.72 | Arid | Grassland |
| *3. Yanco* | -34.72 | 146.13 | Semi-Arid | Croplands (Crop-grass mix) |
| *4. Carman* | 49.67 | -97.98 | Cold | Croplands |
| *5. Walnut Gulch* | 31.72 | -110.09 | Arid | Shrub open |
| *6. Little Washita* | 34.92 | -98.23 | Temperate | Grasslands |
| *7. Fort Cobb* | 35.38 | -98.57 | Temperate | Grasslands (Crop-grass mix) |
| *8. Little River* | 31.72 | -83.73 | Temperate | Cropland/natural mosaic |
| *9. St Josephs* | 41.45 | -84.97 | Cold | Croplands |
| *10. South Fork* | 42.42 | -93.53 | Cold | Croplands |
| *11. Monte Buey* | -33.01 | -62.49 | Arid | Croplands |
| *12. Tonzi Ranch* | 38.43 | -120.95 | Temperate | Savannas woody |
| *13. Kenaston* | 51.39 | -106.51 | Cold | Croplands |
| *14. Valencia* | 39.57 | -1.26 | Arid | Savannas woody |
| *15. Niger* | 13.57 | 2.66 | Arid | Grasslands |
| *16. Benin* | 9.77 | 1.68 | Arid | Savannas |
| *17. TxSON* | 30.43 | -98.61 | Temperate | Grasslands |
| *18. HOBE* | 55.97 | 9.10 | Temperate | Croplands (Crop-natural mix) |

Table A1. Latitude and longitude of SMAP L4\_SM validation sites, along with corresponding climate regime and MODIS IGBP land cover type.

Appendix B. Ensemble perturbations

As follows is a table describing the types of perturbations used to maintain the ensemble during the forecast period.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Perturbed Variable** | surface excess  [additive] | catchment deficit [additive] | precipitation  [multiplicative] | shortwave radiation [multiplicative] | longwave radiation (downwelling) [additive] |
| **Standard Deviation (S)** | 0.16 | 0.15 | 0.5 | 0.3 | 20 |
| **Timescale [hours]** | 3 | 24 | 72 | 72 | 72 |

Table B1. Meteorological forcing variables that are perturbed to maintain

ensemble spread along with their corresponding perturbation standard deviations (S values) and timescales.

REFERENCES

Anderson, J. L., and S. L. Anderson, 1999: A Monte Carlo Implementation of the Nonlinear Filtering Problem to Produce Ensemble Assimilations and Forecasts. Mon. Wea. Rev., 127, 2741–2758, [https://doi.org/10.1175/1520-0493(1999)127<2741:amciot>2.0.co;2](https://doi.org/10.1175/1520-0493(1999)127%3c2741:amciot%3e2.0.co;2).

Anderson, J. L., 2007: An adaptive covariance inflation error correction algorithm for ensemble filters. Tellus A: Dynamic Meteorology and Oceanography, 59, 210–224, <https://doi.org/10.1111/j.1600-0870.2006.00216.x>.

Anderson, J. L., 2009a: Spatially and temporally varying adaptive covariance inflation for ensemble filters, Tellus A: Dynamic Meteorology and Oceanography, 61:1, 72-83, https://www.tandfonline.com/doi/abs/10.1111/j.1600-0870.2007.00361.x.

Anderson, J., T. Hoar, K. Raeder, H. Liu, N. Collins, R. Torn, and A. Avellano, 2009b: The Data Assimilation Research Testbed: A Community Facility. Bull. Amer. Meteor. Soc., 90, 1283–1296, <https://doi.org/10.1175/2009bams2618.1>.

Anderson, J. L., 2010: A Non-Gaussian Ensemble Filter Update for Data Assimilation. Monthly Weather Review, 138, 4186–4198, <https://doi.org/10.1175/2010mwr3253.1>.

Anderson, J. L., 2020: A Marginal Adjustment Rank Histogram Filter for Non-Gaussian Ensemble Data Assimilation. Monthly Weather Review, 148, 3361–3378, <https://doi.org/10.1175/mwr-d-19-0307.1>.

Burgers, G., P. Jan van Leeuwen, and G. Evensen, 1998: Analysis Scheme in the Ensemble Kalman Filter. Mon. Wea. Rev., 126, 1719–1724, [https://doi.org/10.1175/1520-0493(1998)126<1719:asitek>2.0.co;2](https://doi.org/10.1175/1520-0493(1998)126%3c1719:asitek%3e2.0.co;2).

Colliander, A., and Coauthors, 2022: Validation of Soil Moisture Data Products From the NASA SMAP Mission. IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing, 15, 364–392, <https://doi.org/10.1109/jstars.2021.3124743>. Draper, C. S., R. H. Reichle, G. J. M. De Lannoy, and Q. Liu, 2012: Assimilation of passive and active microwave soil moisture retrievals. Geophys. Res. Lett., 39, n/a-n/a, <https://doi.org/10.1029/2011gl050655>.

Draper, C., and R. Reichle, 2015: The impact of near-surface soil moisture assimilation at subseasonal, seasonal, and inter-annual timescales. Hydrol. Earth Syst. Sci., 19, 4831–4844, https://doi.org/10.5194/hess-19-4831-2015.

Evensen, G., 1992: Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model. J. Geophys. Res., 97, 17905–17924, <https://doi.org/10.1029/92jc01972>.

Ford, T. W., S. M. Quiring, and O. W. Frauenfeld, 2016: Multi‐decadal variability of soil moisture–temperature coupling over the contiguous United States modulated by Pacific and Atlantic sea surface temperatures. Int. J. Climatol., 37, 1400–1415, <https://doi.org/10.1002/joc.4785>.

Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). J. Climate, 30, 5419–5454, <https://doi.org/10.1175/jcli-d-16-0758.1>.

Gharamti, M., 2018: Enhanced Adaptive Inflation Algorithm for Ensemble Filters. Mon. Wea. Rev., 146, 623–640, <https://doi.org/10.1175/mwr-d-17-0187.1>.

Gordon, N. J., D. J. Salmond, and A. F. M. Smith, 1993: Novel approach to nonlinear/non-Gaussian Bayesian state estimation. IEE Proc. F Radar Signal Process. UK, 140, 107, <https://doi.org/10.1049/ip-f-2.1993.0015>.

Houtekamer, P. L., and F. Zhang, 2016: Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation. Monthly Weather Review, 144, 4489–4532, <https://doi.org/10.1175/mwr-d-15-0440.1>.

Ide, K., P. Courtier, M. Ghil, and A. C. Lorenc, 1997: Unified Notation for Data Assimilation : Operational, Sequential and Variational (Special Issue Data Assimilation in Meteorology and Oceanography: Theory and Practice). Journal of the Meteorological Society of Japan, 75, 181–189, <https://doi.org/10.2151/jmsj1965.75.1b_181>.

Kalman, R. E., 1960: A New Approach to Linear Filtering and Prediction Problems. Journal of Basic Engineering, 82, 35–45, <https://doi.org/10.1115/1.3662552>.

Keppenne, C. L., 2000: Data Assimilation into a Primitive-Equation Model with a Parallel Ensemble Kalman Filter. Mon. Wea. Rev., 128, 1971–1981, [https://doi.org/10.1175/1520-0493(2000)128<1971:daiape>2.0.co;2](https://doi.org/10.1175/1520-0493(2000)128%3c1971:daiape%3e2.0.co;2).

Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar, 2000: A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure. J. Geophys. Res., 105, 24809–24822, <https://doi.org/10.1029/2000jd900327>.

Koster, R. D., P. A. Dirmeyer, A. N. Hahmann, R. Ijpelaar, L. Tyahla, P. Cox, and M. J. Suarez, 2002: Comparing the Degree of Land–Atmosphere Interaction in Four Atmospheric General Circulation Models. J. Hydrometeor, 3, 363–375, [https://doi.org/10.1175/1525-7541(2002)003<0363:ctdola>2.0.co;2](https://doi.org/10.1175/1525-7541(2002)003%3c0363:ctdola%3e2.0.co;2).

Koster, R. D., and M. J. Suarez, 2003: Impact of Land Surface Initialization on Seasonal Precipitation and Temperature Prediction. J. Hydrometeor, 4, 408–423, [https://doi.org/10.1175/1525-7541(2003)4<408:iolsio>2.0.co;2](https://doi.org/10.1175/1525-7541(2003)4%3c408:iolsio%3e2.0.co;2).

Kuenzer, C., and Coauthors, 2009: El Niño southern oscillation influences represented in ERS scatterometer-derived soil moisture data. Applied Geography, 29, 463–477, <https://doi.org/10.1016/j.apgeog.2009.04.004>.

Kumar, S. V., R. H. Reichle, R. D. Koster, W. T. Crow, and C. D. Peters-Lidard, 2009: Role of Subsurface Physics in the Assimilation of Surface Soil Moisture Observations. Journal of Hydrometeorology, 10, 1534–1547, <https://doi.org/10.1175/2009jhm1134.1>.

Miller, R. N., M. Ghil, and F. Gauthiez, 1994: Advanced Data Assimilation in Strongly Nonlinear Dynamical Systems. J. Atmos. Sci., 51, 1037–1056, [https://doi.org/10.1175/1520-0469(1994)051<1037:adaisn>2.0.co;2](https://doi.org/10.1175/1520-0469(1994)051%3c1037:adaisn%3e2.0.co;2).

Nakano, S., G. Ueno, and T. Higuchi, 2007: Merging particle filter for sequential data assimilation. Nonlin. Processes Geophys., 14, 395–408, <https://doi.org/10.5194/npg-14-395-2007>.

Reichle, R. H., J. P. Walker, R. D. Koster, and P. R. Houser, 2002a: Extended versus Ensemble Kalman Filtering for Land Data Assimilation. J. Hydrometeor, 3, 728–740, [https://doi.org/10.1175/1525-7541(2002)003<0728:evekff>2.0.co;2](https://doi.org/10.1175/1525-7541(2002)003%3c0728:evekff%3e2.0.co;2).

Reichle, R. H., D. B. McLaughlin, and D. Entekhabi, 2002b: Hydrologic Data Assimilation with the Ensemble Kalman Filter. Mon. Wea. Rev., 130, 103–114, [https://doi.org/10.1175/1520-0493(2002)130<0103:hdawte>2.0.co;2](https://doi.org/10.1175/1520-0493(2002)130%3c0103:hdawte%3e2.0.co;2).

Reichle, R. H., and R. D. Koster, 2003: Assessing the Impact of Horizontal Error Correlations in Background Fields on Soil Moisture Estimation. J. Hydrometeor, 4, 1229–1242, [https://doi.org/10.1175/1525-7541(2003)004<1229:atiohe>2.0.co;2](https://doi.org/10.1175/1525-7541(2003)004%3c1229:atiohe%3e2.0.co;2).

Reichle, R. H., W. T. Crow, and C. L. Keppenne, 2008: An adaptive ensemble Kalman filter for soil moisture data assimilation. Water Resour. Res., 44, <https://doi.org/10.1029/2007wr006357>.

Reichle, R. H., and Coauthors, 2019: Version 4 of the SMAP Level‐4 Soil Moisture Algorithm and Data Product. J. Adv. Model. Earth Syst., 11, 3106–3130, <https://doi.org/10.1029/2019ms001729>.

Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling, 2010: Investigating soil moisture–climate interactions in a changing climate: A review. Earth-Science Reviews, 99, 125–161, <https://doi.org/10.1016/j.earscirev.2010.02.004>.

Verlaan, M., and A. W. Heemink, 1997: Tidal flow forecasting using reduced rank square root filters. Stochastic Hydrology and Hydraulics, 11, 349–368, <https://doi.org/10.1007/bf02427924>.

Whitaker, J. S., and T. M. Hamill, 2012: Evaluating Methods to Account for System Errors in Ensemble Data Assimilation. Monthly Weather Review, 140, 3078–3089, <https://doi.org/10.1175/mwr-d-11-00276.1>.

Xiong, X., I. M. Navon, and B. Uzunoglu, 2006: A Note on the Particle Filter with Posterior Gaussian Resampling. Tellus A: Dynamic Meteorology and Oceanography, 58, 456–460, <https://doi.org/10.1111/j.1600-0870.2006.00185.x>. ­

Zhang, F., M. Zhang, and J. A. Hansen, 2009: Coupling ensemble Kalman filter with four-dimensional variational data assimilation. Adv. Atmos. Sci., 26, 1–8, <https://doi.org/10.1007/s00376-009-0001-8>.