A dynamic approach to addressing observation-minus-forecast mean differences in a land surface skin temperature data assimilation system

Clara Draper[1,2], * Rolf Reichle[1], Gabrielle De Lannoy[1,2]

1. Global Modeling and Assimilation Office, NASA GSFC, Greenbelt, MD, USA

2. Universities Space Research Association, Columbia, MD, USA

Benjamin Scarino[3]

3. Science Systems and Applications, Inc., Hampton, VA, USA

* *Corresponding author address:* Clara Draper, Global Modeling and Assimilation Office, NASA GSFC, Code 610.1, Greenbelt, MD, 20771, USA.

E-mail: clara.draper@nasa.gov

ABSTRACT

In land data assimilation, bias in the observation-minus-forecast (O-F) residuals is typically removed from the observations prior to assimilation by rescaling the observations to have the same long-term mean (and higher-order moments) as the corresponding model forecasts. Such observation rescaling approaches require a long record of observed and forecast estimates, and an assumption that the O-F mean differences are stationary. A two-stage observation bias and state estimation filter is presented, as an alternative to observation rescaling that does not require a long data record or assume stationary O-F mean differences. The two-stage filter removes dynamic (nonstationary) estimates of the seasonal scale O-F mean difference from the assimilated observations, allowing the assimilation to correct the model for synoptic-scale errors without adverse effects from observation biases. The two-stage filter is demonstrated by assimilating geostationary skin temperature (T_{skin}) observations into the Catchment land surface model. Global maps of the O-F mean differences are presented, and the two-stage filter is evaluated for one year over the Americas. The two-stage filter effectively removed the T_{skin} O-F mean differences, for example the GOES-West O-F mean difference at 21:00 UTC was reduced from 5.1 K for a bias-blind assimilation to 0.3 K. Compared to independent in situ and remotely sensed T_{skin} observations, the two-stage assimilation reduced the unbiased Root Mean Square Difference (ubRMSD) of the modeled T_{skin} by 10% of the open-loop values.

1. Introduction

Within the context of data assimilation, 'bias' refers to errors in modeled or observed variables that persist over time and/or space. Standard 'bias-blind' data assimilation methods are based on the assumption that neither the forecast model nor the observations are biased, and these methods will produce suboptimal output in the presence of bias (Dee and Da Silva 1998). Unfortunately, the forecast models and observation data sets used in Earth system applications, including for the land surface, typically are biased (Dee and Todling 2000; Reichle et al. 2004). Observation biases can arise from errors in the observing instrument and its calibration, the observation operator, or the retrieval model, as well as representativity errors between the observed state variables and their modeled counterparts. Likewise, forecast biases can arise from errors in the forecast model structure, parameters, initial conditions, and forcing.

Ideally, the cause of observation and forecast biases should be diagnosed and treated at the source. Where this is not possible, these biases can also be addressed in data assimilation by applying an observation bias correction prior to assimilation (e.g., Harris and Kelly, 2001) or by using a 'bias-aware' assimilation system explicitly designed to correct either observation biases (e.g., Auligné et al. 2007; Fertig et al. 2009) or forecast biases (e.g., Dee and Todling 2000; Keppenne et al. 2005). Bias correction methods require that the bias be observable (Dee and Da Silva 1998), and the ocean and atmosphere examples cited above measure the biases against confident estimates of the true mean state, typically obtained with reference to point-based observations (e.g., ocean buoys, radiosondes). However, the land surface is much more heterogeneous than the ocean and atmosphere, and point-based in situ observations are in general not representative of the coarse resolution states estimated by remote sensors and land surface models (Crow et al. 2012). Consequently, for large domains the true mean land surface states are unknown, since there are large systematic differences between the mean (and variance) of different observed and modeled land surface data sets, none of which can in general be identified as having statistics representative of the true state (Reichle et al.

2004).

Since observation and forecast biases cannot be observed for land surface states, it is standard practice to remove the systematic differences between the observed and forecast estimates from land data assimilation, usually by rescaling the observations to be consistent with the long-term mean (and variance, and sometimes higher order moments) of the forecasts (e.g., Reichle and Koster 2004; Drusch et al. 2005; Scipal et al. 2008; Crow et al. 2011). This prevents the systematic differences from adversely impacting the model state, while satisfying the minimum criterion for optimal bias-blind data assimilation that there be no difference between the mean values of the observed and forecast estimates. The assimilation can then correct the model for random errors developing during each forecast, where 'random errors' are errors persisting over time scales much shorter than the assumed bias time scale. Data assimilation with observation rescaling has been shown to yield land surface estimates that are superior to modeled or observed estimates alone (Slater and Clark 2006; Reichle et al. 2007; Ghent et al. 2010; Crow et al. 2011; Draper et al. 2012; De Lannoy et al. 2012; de Rosnay et al. 2013). This rescaling approach is often referred to as 'observation bias correction', although strictly speaking, it is not the observation bias (defined against the true mean state) that is corrected, but the lumped observation-bias-minus-forecast-bias.

The long data record of observed and forecast state estimates required for estimating observation rescaling coefficients has slowed the implementation of land data assimilation in large-scale applications, particularly within atmospheric systems, which are frequently updated and yet prohibitively expensive to replay over long periods. Consequently, Dharssi et al. (2011) and de Rosnay et al. (2013) identify the difficulty in obtaining observation rescaling coefficients as one cause of the limited impact of assimilating remotely sensed soil moisture observations into atmospheric models. The long data record requirement also prevents the assimilation of new remotely sensed data sets, and necessitates costly reprocessing of the rescaling parameters after significant updates to assimilated data sets.

Consequently, this manuscript presents a method for removing the O-F mean difference (i.e., the lumped observation-bias-minus-forecast-bias) in land data assimilation systems without access to a long data record, by using a two-stage observation bias and state update estimation filter. 'Bias' is defined subjectively, in terms of the temporal and spatial scales over which it applies. In seeking a bias correction method that does not require a long data record, the bias is necessarily defined over shorter time scales, and the presented two-stage filter dynamically estimates nonstationary O-F mean differences that evolve at seasonal time scales.

There are typically large systematic differences between remotely sensed and modeled T_{skin} (Ghent et al. 2010; Wang et al. 2014), and if not adequately addressed these differences will result in a sub-optimal assimilation, potentially leading to degraded flux forecasts (e.g., Reichle et al. 2010). Hence, the two-stage observation bias and state estimation scheme has been demonstrated here by assimilating geostationary T_{skin} observations into the Catchment land surface model.

The remainder of this manuscript is outlined as follows. In Section 2, the two-stage observation bias and state estimation scheme is developed, and contrasted to observation rescaling approaches. The two-stage filter is then demonstrated with an example assimilation of remotely sensed skin temperature (T_{skin}) observations into a land surface model. The T_{skin} assimilation experiments are outlined in Section 3, before the results are presented in Section 4. Finally, Section 5 presents a summary and conclusions.

2. The state and bias filter equations

The two-stage observation bias and state estimation approach introduced here is based on the on-line two-stage forecast bias and state estimation approach of Dee and Da Silva (1998), which has been successfully implemented in atmosphere (Dee and Todling 2000), ocean (Chepurin et al. 2005; Keppenne et al. 2005), and land (Bosilovich et al. 2007; De Lannoy et al. 2007; Reichle et al. 2010) data assimilation. Following Friedland (1969), Dee and

Da Silva (1998) decouple the forecast bias estimation from the state update, and use a separate Kalman filter to estimate the forecast bias. The (bias-blind) state update innovations (i.e., the O-F residuals) are used to measure the forecast bias for the bias update, based on the assumption that the observations are unbiased, and persistence is used to predict the forecast bias. Pauwels et al. (2013) recently extended the theory of the two-stage forecast bias and state estimation filter to also estimate the observation bias. In their approach, demonstrated with synthetic experiments, the (bias-blind) state update innovation measures the observation bias plus the forecast bias, and is partitioned into the two separate bias terms by calibration. However, observations of the true mean state are ultimately required to partition the sum of the biases.

In contrast, we derive the two stage filter as if to estimate the observation biases measured using the (bias-blind) state update innovations, based on the assumption that the forecasts are unbiased. However, in the intended land data assimilation applications, it is recognized that the forecasts are almost certainly biased, so that the estimated 'observation bias' really represents the O-F mean difference (the lumped observation-bias-minus-forecast-bias), to be used to adjust the observations to have the same mean value as the forecast estimates, consistent with observation rescaling approaches.

Below, the bias-free EnKF equations are reviewed (Section 2a), before the optimal solution for the two-stage observation bias and state estimation filter is derived (Section 2b). Then, a parameterization of the Kalman gain for the bias update is introduced, to avoid specifying the unknown prior observation bias uncertainty (Section 2c).

a. The bias-free EnKF

The bias-free EnKF, as implemented by Reichle et al. (2013) for land data assimilation, consists of a model forecast step and a state update step. For the ith ensemble member, the state forecast and update at the kth assimilation time are:

$$x^-_{k,i} = f(x^+_{k-1,i}, q_{k,i}) \quad (1)$$

$$x^+_{k,i} = x^-_{k,i} + K_k(y^o_{k,i} + H_k x^-_{k,i}) \quad (2)$$

$$y^o_{k,i} = y^o_k + v_{k,i} \quad (3)$$

where x is the model state vector, $f(.)$ is the forecast model, q represents the model error (or perturbation vector), K is the Kalman gain matrix, y^o is the observation vector, H is the observation operator, and v is an applied (zero mean, normal) perturbation representative of the expected observation errors. For simplicity we assume H to be linear, however the theory is unchanged if this assumption is relaxed. Throughout this manuscript, a super-scripted state vector indicates an estimated value, with the $-$ and $+$ superscripts indicating the prior and posterior estimates, respectively. In contrast, the absence of a superscript for a state variable indicates the true state vector.

In a bias-free EnKF, the errors in x^- and y^o are assumed to have vanishing long-term mean errors, and to be uncorrelated with each other. Under these assumptions, x^+ provides an unbiased estimate of x, and the optimal (minimum posterior state error variance) Kalman gain for the kth state update, K_k, is given by:

$$K_k = P^{x-}_k H^T_k (R^o + H_k P^{x-}_k H^T_k)^{-1} \quad (4)$$

where P^{x-} is the prior model state error covariance matrix, and R^o is the observation error covariance matrix. P^{x-} is diagnosed from the ensemble spread, while for land data assimilation R^o is typically assumed to be constant in time and have zero off-diagonal terms (e.g., Draper et al. 2012). Applying the above equations in the presence of (unknown) observation and/or forecast biases is sub-optimal, and is referred to as 'bias-blind' data assimilation (Dee and Da Silva 1998).

b. The two-stage observation bias and state estimation

For an observation-bias-aware assimilation, the observation vector is allowed to have a nonzero mean error persisting over some extended time period (a bias). The biased observations, written $\tilde{y}^o_k$, can be partitioned into the bias term, b_k, and the remaining zero-mean error component, y^o_k:

$$\tilde{y}^o_k = b_k + y^o_k \tag{5}$$

The observations are then bias-corrected within the state update (equation 2) to remove the bias from the innovations, giving an unbiased estimate of x^+:

$$x^+_{k,i} = x^-_{k,i} + \tilde{K}_k(\tilde{y}^o_{k,i} - b_k - H_k x^-_{k,i}) \tag{6}$$

where $\tilde{K}$ is the Kalman gain for the state update based on the bias corrected observation vector.

A separate, discrete Kalman filter is then used to estimate the observation bias. The observation bias is measured using the mean O-F ($< \tilde{y^o}_{k,i} - H_k x^-_{k,i} >$, where $< . >$ is the ensemble mean). The bias is initialized at zero, and persistence is used as the bias prediction model, since the bias is assumed not to change significantly during individual assimilation cycles. The persistence model is recognized as an approximation, since a (potentially desirable) feature of the two-stage filter is the nonstationary nature of the bias estimates. The observation bias forecast and update equations for the kth assimilation time are then written:

$$b^-_k = b^+_{k-1} \tag{7}$$

$$b^+_k = b^-_k + L_k < \tilde{y}^o_{k,i} - b^-_k - H_k x^-_{k,i} > \tag{8}$$

where L_k is the Kalman gain for the bias update. Equations 7 and 8 provide an unbiased estimate of the observation bias, regardless of the selection of L_k. Appendix A shows that if the errors in the observations, the prior bias estimate, and the prior state estimate are not correlated with each other, and if b_k^- provides an unbiased estimate of the observation bias, the optimal (minimum error covariance) posterior bias estimate is obtained with L_k equal to:

$$L_k = P_k^{b-}(R^o + P_k^{b-} + H_k P_k^{x-} H_k^T)^{-1} \tag{9}$$

Here R^o is unchanged from equation 4 and represents the random errors in the observations only, while P_k^{b-} is the random error covariance matrix for the prior observation bias estimate.

Substituting the best estimate of the bias (b_k^+; equation 8) into equation 6 then gives the state update equation with observation bias correction:

$$x_{k,i}^+ = x_{k,i}^- + \tilde{K}_k(\tilde{y}_{k,i}^o - b_k^+ - H_k x_{k,i}^-) \tag{10}$$

Up to this point, the presented derivation of the two-stage observation bias and state estimation equations has followed that of Pauwels et al. (2013), with their forecast bias set to zero. However, we now diverge from their approach. In Appendix B, we show that if the optimal expression for L is used (equation 9), $\tilde{K}_k$ in equation 10 is the same as K_k for the bias-free filter (equation 4). That is, the Kalman gain is unchanged by the inclusion of the two-stage observation bias estimate in the state update equation. This result parallels that of Dee and Todling (2000), who show that for the on-line two-stage forecast bias and state estimation filter the state update Kalman gain is unchanged by the inclusion of the forecast bias estimate in the state update equation.

To summarize the two-stage observation bias and state estimation filter equations presented above, equations 1 and 10 are used for the state forecast and update, respectively, together with the state update Kalman gain of equation 4. Equations 7 and 8 are used for the observation bias forecast and update, respectively, together with the bias update Kalman gain of equation 9 (although equation 9 will be replaced by an empirical function in

Section c). For illustrative purposes, substituting equation 8 into equation 10, then taking the ensemble average gives:

$$x^+_{k,i} = x^-_{k,i} + \tilde{K}_k(\tilde{y}^o_{k,i} - b^-_k - H_k x^-_{k,i}) - \tilde{K}_k L_k < \tilde{y}^o_{k,i} - b^-_k - H_k x^-_{k,i} > \quad (11)$$

and:

$$< x^+_{k,i} > = < x^-_{k,i} > + \tilde{K}_k(I - L_k) < \tilde{y}^o_{k,i} - b^-_k - H_k x^-_{k,i} > \quad (12)$$

Comparing equation 12 to equation 8 for the bias update demonstrates how the two-stage filter partitions the innovations $(\tilde{y}^o_{k,i} - b^-_k - H_k x^-_{k,i})$ into updates to the bias estimate and state estimate.

The presented two-stage observation bias and state estimation filter parallels the on-line two-stage forecast bias and state estimation of Dee and Da Silva (1998) but differs from the original two-stage estimation approach of Friedland (1969) in that the state update equation is optimized with the bias correction terms included (i.e., the Kalman gain is obtained by optimizing equation 10, rather than equation 2). The resulting two-stage filter is optimal if the various assumptions stated above hold. However, in practice the filter is unlikely to be optimal, since, for example, the prior state errors and the prior observation bias errors have been assumed uncorrelated, yet both contain information (and errors) from past observations.

c. Parametrization of the bias gain

The two-stage observation bias correction and state estimation approach outlined above requires the specification of the unknown error covariance matrix P^{b-} for the prior bias estimate to calculate the observation bias update Kalman gain, L, in equation 9. Dee and Da Silva (1998) and Pauwels et al. (2013) assumed that the prior forecast bias error covariances were proportional to the prior forecast error covariances, and Pauwels et al.

(2013) assumed that the prior observation bias error covariances were proportional to the forecast observation error covariances. We instead replace L with an empirical function. This approach is made possible because P^{b-} is not required for the bias-aware state update Kalman gain, due to the equivalence of the bias-free and bias-aware Kalman gains noted in Section b.

For the assimilation of a single observation type at a single location, L_k becomes scalar. For the assimilation of the jth location and observation type, we approximate $L_{j,k}$ with a function designed to approach one as the time since the last assimilated observation increases:

$$\lambda_{j,k} = 1 - e^{-\Delta t_{j,k}/\tau_j} \tag{13}$$

where $\Delta t_{j,k}$ is the number of time steps since the most recent observation of type j was assimilated, and τ_j is a user-defined parameter representing the e-folding time scale of the bias memory for observation type j. This function was chosen since it approximates the expected behavior $L_{j,k}$ under two important scenarios. In the first scenario, no observations have been recently assimilated, relative to the assumed time scale of the bias, and there is little information with which to predict $b^-_{j,k}$. Hence, $L_{j,k}$ is expected to be close to one, as predicted by equation 13 for large $\Delta_{j,k}/\tau_j$. In the second scenario, observations are being assimilated with some regularity, and random errors in $b^-_{j,k}$ will be dominated by random errors in the $(\tilde{y^o_k} - H_k x^-_k)$ sequence used to update $b^-_{j,k}$ (since by definition the persistence model will not introduce significant errors into the bias estimate), however, the bias filter will gradually filter these errors over time. Hence, if $\Delta t_{j,k}$ is assumed to generalize the recent availability of observations, equation 13 will approximate the increased certainty in $b^-_{j,k}$ (and subsequent reduction in $\lambda_{j,k}$) as more observations are assimilated.

The empirical $\lambda_{j,k}$ must adequately account for the first scenario described above, of no recent observations, since from equation 12 a large L_k is necessary in this case to prevent the potentially large $b^-_{j,k}$ errors from being propagated into the model state vector. This situation can occur reasonably regularly, since there are often seasonal-scale gaps in land surface observation records, when atmospheric and/or land surface conditions prevent remote sensing of the land surface. Note the contrast to forecast bias correction, for which one can fall back on a conservative approach of underestimating the forecast bias (Dee and Todling 2000; Reichle et al. 2010) when the bias estimate is highly uncertain, since the model state will still be updated towards the true state (defined by the observations in this case).

For the assimilation of multiple observation types and locations, $\lambda_{j,k}$ can be extended in the obvious way to a matrix, Λ_k, by setting the jth diagonal element of Λ_k to $\lambda_{j,k}$, and setting the off-diagonal terms to zero (i.e, disregarding potential spatial correlation, or cross-correlation between observation types, in the bias updates). A potential weakness of the above parameterization of $\lambda_{j,k}$ is that a $b^-_{j,k}$ estimate based on a single recent observation would be assigned high confidence. Consequently, observations are excluded from the state update when the bias estimate is based on less than two observations within the last $\tau_j/2$ time steps (although these observations are still used to update $b^-_{j,k}$).

d. Comparison to observation rescaling

The two-stage observation bias and state estimation method presented above treats the systematic differences between observations and forecasts quite differently compared to the observation rescaling methods currently used in many land data assimilation systems. Observation rescaling (Reichle and Koster 2004; Drusch et al. 2005; Scipal et al. 2008; Crow et al. 2011) is designed to remove the long-term systematic differences in the mean and variance (and possibly higher order moments) of the observed and forecast state estimates, where 'long-term' is defined by the length of the data record used to calculate the rescaling parameters. These systematic differences are typically assumed to be stationary, and a static set of bias correction parameters is used. Consequently, a (bias-free) data assimilation with observation rescaling will then adjust the model states to reduce residual differences between the observations and model forecasts. Such differences include those occurring at sub-seasonal time-scales, differences in the phase of the seasonal cycle, and also differences in the intra-annual seasonal cycle, if the data record used to estimate the rescaling coefficients was sufficiently long to sample the climatological inter-annual variability.

In contrast, the two-stage observation bias and state estimation method presented here is designed to remove only the systematic difference in the mean of the observed and forecast state estimates, and this mean difference is not restricted to being stationary. The filter dynamically estimates the O-F mean differences based only on measurements up to the current assimilation cycle, with greater weight placed on more recent measurements. The resulting estimates are then nonstationary, and will evolve at a time scale determined by the τ parameter in equation 13. Specifying τ to represent seasonal time scales will result in the observations being adjusted to match the seasonal cycle of the forecast estimates. The assimilation will then adjust the model state vector to reduce differences between the observations and forecasts at sub-seasonal time scales, somewhat consistent with the observation rescaling approach. Although systematic differences in the variance of the observations and forecasts are not explicitly removed, as they are in observation rescaling, the component of variance due to seasonal, or longer, time scale dynamics will be addressed.

For a given data assimilation experiment, the suitability of the two-stage filter depends on the distribution of the systematic differences between the observed and forecast estimates. For T_{skin}, there can be large differences between the mean values of different model forecast and observed estimates (Wang et al. 2014), however T_{skin} variability is reasonably well constrained, due in part to the tight coupling between T_{skin} and the (comparatively well observed) low-level atmospheric temperature. Hence, using the two-stage observation bias and state estimation to adjust the seasonal cycle of the mean observed T_{skin} to match that of the forecast estimates is expected to effectively address the systematic differences between observed and forecast T_{skin} in an assimilation. However, for many other land surface variables this approach may not be sufficient. Most notably, for near-surface soil moisture there are large systematic differences between the variability of different data sets, including the sub-seasonal-scale variability (e.g., see Draper et al. (2013), their Figure 2). This is due in part to the absence of global data sets constraining the possible soil moisture range, and the subsequent uncertainty in the parameters controlling the soil moisture response to atmospheric forcing (specifically controlling the total volume of pore space available for water storage in the soil column).

3. Skin temperature assimilation

The two-stage observation bias and state estimation scheme has been demonstrated by assimilating geostationary T_{skin} observations into the Catchment land surface model. Two separate assimilation experiments were performed. First, the T_{skin} data were assimilated over the Americas at 0.3125°x0.25° longitude by latitude resolution, from 1 June, 2012 to 31 May, 2013. Second, to obtain example global maps of the mean differences between the observed and forecast T_{skin}, the T_{skin} data were assimilated globally, at a coarser resolution of 0.625°x0.50°, from 1 May, to 1 August, 2012.

a. Catchment land surface model

Catchment (Koster et al. 2000) is the land surface modeling component of the Goddard Earth Observing System Model, version 5 (GEOS-5; Rienecker et al. 2008). The catchment model equivalent variable to remotely sensed T_{skin} is the surface temperature (T_{surf}), defined as the average temperature of the canopy and soil surface, and representative of an arbitrarily thin layer separating the canopy and soil surface from the atmosphere. While the Catchment T_{surf} is prognostic, it has a very short memory over most land surface types due to its very low surface specific heat capacity (200 $\mathrm{JK^{-1}m^{-2}}$, except for broadleaf evergreen vegetation). The assimilation experiments were performed off-line (i.e., decoupled from the atmospheric model), using meteorological forcing data from the NASA Modern-Era Retrospective analysis for Research and Applications (MERRA) (Rienecker et al. 2011) and Catchment model parameters from the routine GEOS-5 system. The initial land surface state was spun-up from an archived GEOS-5 restart file on 1 January, 2000, by integrating the model forward (without perturbations) to 1 January 2012, and the model ensemble was then spun up from 1 January, 2012 to the start of the assimilation on 1 June, 2012.

b. Geostationary skin temperature data

The assimilated T_{skin} observations are retrieved from geostationary Thermal Infrared (TIR) brightness temperature observations at the NASA Langley Research Center (Scarino et al. 2013). The T_{skin} data are retrieved every three hours, and reported on the 0.3125°x0.25° GEOS-5 model grid. The geostationary data have been produced in near-real time since 2011, from a constellation of satellites providing global (53° S to 53° N, after quality control) coverage: Geostationary Operational Environmental Satellites (GOES)-East, GOES-West, the second Multifunctional Transport Satellite (MTSAT-2), Feng Yun-2E (FY-2E), and Meteosat-9 (Met-9). However, for the assimilation experiment over the Americas domain, an updated data set from the GOES-East and GOES-West satellites, produced with the latest retrieval model, has been used. Where observations are available from more than one geostationary satellite, only the observations from the closest satellite were assimilated. The observation quality control discards observations with a viewing zenith angle greater than 60°, a solar zenith angle between 83° and 90°, a grid-cell cloud fraction above 20%, or if the land modeling system indicates precipitation or a snow-covered surface.

Figure 1 shows the coverage of the observation-quality controlled (GOES-West and GOES-East) T_{skin} observations assimilated in the Americas experiment, as a fraction of the total number of possible observation times (eight 3-hourly observation times per day). There are few observations available during colder periods, due mostly to increased cloudiness. Hence, the coverage is very low ($<$ 15% of the maximum possible coverage) at higher latitudes. The coverage is also low over the Amazon, again due to cloudiness. There is some diurnal variation in the coverage, with slightly more observations available during the daytime hours (10% more than nighttime). In Section 4 evaluation statistics are only reported at locations where observations were assimilated for at least 7.5% of the possible observation times at each time of day ($\sim$ 30 observations).

c. Assimilation system

The state update component of the two-stage filter uses the EnKF (Reichle et al. 2013), with 12 ensemble members and 3-hourly assimilation of the T_{skin} observations. The assimilation update vector consists of T_{surf} and the ground heat content (GHT1) associated with the near-surface (0-10 cm) soil temperature. The ensemble was generated using the forcing and model state perturbations in Table 1, which were adapted from Reichle et al. (2010) to account for the inclusion of GHT1 in the state update vector. Note that the Catchment model version used in Reichle et al. (2010) had a much higher specific heat capacity for T_{surf} (70,000 JK^{-1}m^{-2}) than is currently used, and T_{surf} represented a 5 cm layer depth (hence Reichle et al. (2010) updated only T_{surf}). The observation error standard deviations for the T_{skin} retrievals were set at 1.3 K and 2.1 K during the nighttime and daytime, respectively, which implies that the model and observations have roughly similar skill.

The Catchment model divides each model grid cell into multiple computational elements, and a 3-D filter (with non-zero horizontal model and observation error correlations, Reichle and Koster 2003) was used to spread the observations to all model computational surface elements within each grid cell. For both the observation errors and the (forcing and state vector) ensemble perturbations in Table 1, relatively short horizontal error correlation scales with an e-folding distance of 0.17° were applied. Note that preliminary experiments with increased horizontal error correlation scales (between 0.5° and 1.0°) degraded the assimilation results, likely because the strong dependence on cloud cover limits the horizontal error correlations of estimated T_{skin}.

The observation bias update was performed independently at each model grid cell (i.e., using a 1-D filter). Since there is a strong diurnal cycle in the observations-minus-forecast mean difference (as will be shown in Section 4), the observation bias was modeled separately at each of the eight diurnal observation times, following Reichle et al. (2010).

d. Evaluation of assimilation output

The results of the assimilation experiment over the Americas have been evaluated by comparison to independent observations of clear sky T_{skin}, from the in situ SURFRAD network (Augustine et al. 2005), and from remotely sensed MODIS TIR observations. The six SURFRAD sites shown in Figure 1 were used (Fort Peck was excluded since the geostationary satellite viewing zenith angle exceeds 60° there). For each of the validation sites, 3-hourly T_{skin} were calculated from the SURFRAD up-welling and down-welling TIR radiation observations using the Stefan-Boltzmann equation, and broad-band emissivity calculated from MODIS Terra monthly narrow-band emissivity observations (MOD11C3), using Wang et al. (2005). For MODIS, Aqua (MYD11C1) and Terra (MOD11C1) daily clear-sky T_{skin} data on the 0.05° Climate Modeling Grid have been averaged up to the GEOS-5 model grid, and assumed to occur at the geostationary observation time closest to the median MODIS observation time over the domain for each satellite orbit direction.

The skill of the T_{skin} assimilation experiment in predicting each of the independent data sets has been compared to the skill of an open-loop (no data assimilation) ensemble, generated with the same model perturbations as used for the assimilation experiment. For both cases, instantaneous model T_{surf} is compared to the independent T_{skin} observations at times for which geostationary T_{skin} observations are available (i.e., for the assimilation experiment the posterior T_{surf} is evaluated). There are systematic differences between the mean values of the T_{skin} data sets used here, and these differences cannot be attributed to biases in any particular data set. Hence, the evaluation statistic is the unbiased Root Mean Square Difference (ubRMSD), calculated at each model grid cell after removing the mean difference over the full time period (separately at each time of day) between the data sets.

4. Results

a. O-F mean differences

Without bias correction there is a strong diurnal cycle in the mean difference between the observed and forecast T_{skin}. For example, Figure 2 shows the diurnal cycle in the spatial mean O-F mean difference over the Americas for a bias-blind assimilation of the GOES-East and GOES-West geostationary T_{skin} observations (using the same observation error covariances and forecast ensemble perturbations as for the bias-aware assimilation experiments). For both GOES-East and GOES-West, the O-F mean differences are more positive after solar noon. The GOES-West O-F mean differences are consistently positive, and larger than those for GOES-East throughout the diurnal cycle, with a maximum value of 5.1 K at 21:00 UTC, compared to values < 2 K during the nighttime. In contrast, the GOES-East O-F mean differences are negative during the nighttime, and positive during the daytime, but with magnitude consistently < 1 K in both cases, except for the 2.8 K maximum at 18:00 UTC. The T_{skin} data retrieved from the different geostationary satellite are reasonably well calibrated (Minnis et al. 2002), and the differences between the GOES-East and GOES-West O-F mean differences in Figure 2 are almost certainly not related to the sensors themselves, but to the contrasting land covers observed by each. The small spatial mean O-F mean differences for GOES-East are due to cancellation between regions of positive and negative O-F mean differences in the spatial means.

While the effectiveness of the observation bias correction has been analyzed throughout the diurnal cycle, for brevity the focus here is on the results at 21:00 UTC, when the largest O-F mean differences occurred in Figure 2. To demonstrate the influence of τ (the time scale of the bias estimate) on the O-F mean differences estimated by the filter (i.e., the b^+), Figure 3 compares the b^+ time series at 21:00 UTC, estimated using τ values between 10 and 30 days, at the three SURFRAD locations with the greatest observation coverage. The SURFRAD locations are used only for convenience, and no SURFRAD data were used in these plots.

For comparison, each panel also includes a smoothed O-F time series, estimated using the first two annual Fourier harmonics, following Vinnikov et al. (2008). Recall from Section 2d, that selecting τ to represent seasonal time scales will allow the assimilation to correct for sub-seasonal-scale (e.g., synoptic-scale) errors. The bias filter tracks the expected seasonal-scale O-F mean differences, while filtering out the higher-frequency noise in the observed and forecast T_{skin}. As expected, the filtered b^+ time series lag the smoothed time series, with the lag increasing as τ increases in Figure 3. The minimum time scale of the features resolved by the b^+ time series also increases as τ increases, and for shorter τ values there is more noise around the seasonal cycle (particularly for 10 days). The greatest differences between the b^+ time series with different τ (and between the filtered and smoothed time series) occurred at Sioux Falls, where the O-F seasonal cycle had the steepest temporal gradient. In particular, during the 2012 summer when the O-F decreased rapidly, the b^+ time series are much higher than the smoothed time series (likely due to the first two Fourier harmonics in the smoothed time series being insufficient to capture the sharp gradient).

For a given application the best choice of τ for estimating the seasonal-scale O-F mean differences will depend on the relative variability of the innovations at seasonal and sub-seasonal time scales. For geostationary T_{skin} assimilation, a compromise value of $\tau = 20$ days has been selected, since this produced b^+ time series with reasonably smooth seasonal cycles that did not lag the O-F time series by too much (Figure 3).

With a τ of 20 days, Figure 4 compares histograms of the state update innovations at 21:00 UTC at the same three locations plotted in Figure 3, for both the bias-blind assimilation experiments and the two-stage observation bias and state estimation scheme. As expected, the innovation distributions for the bias-blind assimilation are biased, with mean values between 1.3 K and 8.0 K (Figures 4a-c). The inclusion of the observation bias correction reduced the mean innovations to magnitudes less than 0.5 K at each location (Figures 4d-f). The observation bias correction also changed the shape of the innovation distributions in Figure 4, reducing their spread and skew. Consequently, the standard deviation at each site is reduced, with the greatest reductions occurring at Sioux Falls, from 4.0 K for the bias-blind assimilation to 2.5 K with the observation bias correction. The altered shape of the innovation distribution is a consequence of the nonstationary bias estimation method accounting for seasonal-scale evolution of the O-F mean difference. In contrast, if a single (stationary) correction were applied to the mean over the full time period, the higher order moments of the innovation distribution would have been unchanged.

The histograms in Figure 4 are representative of the performance of the observation bias correction across the full domain, and throughout the diurnal cycle. For example, for both satellites in Figure 2, the two-stage filter reduced the spatial mean O-F mean difference to magnitudes between 0.0 - 0.3 K throughout the day, compared to bias-blind maxima of 5.1 K and 2.8 K, for GOES-West and GOES-East, respectively. Likewise the mean standard deviation of the innovations across the domain was also reduced throughout the diurnal cycle (not shown), for example from 3.8 K to 3.1 K for GOES-West, and from 2.7 K to 2.1 K for GOES-East, both at 21:00 UTC.

Finally, in Section 2d it was hypothesized that for the assimilation of T_{skin}, the variability of modeled and observed estimates is reasonably well constrained so that adjusting the mean seasonal cycle of the observations (with the two-stage filter) would be sufficient to address the systematic differences between the observed and forecast estimates. Comparing the variance of the observed and forecast T_{skin} confirms that this was the case in the assimilation experiments performed here. For example, over the Americas at 21:00 UTC, the spatially averaged temporal standard deviation of the GOES-West observations was 8.0 K, compared to 7.3 K for the model forecasts over the same domain, with a spatial mean absolute difference between their standard deviations of 1.1 K. Likewise, for GOES-East at 21:00 UTC the mean standard deviation was 5.1 K, compared to 4.9 K for the forecasts, with a spatial mean absolute difference of 0.9 K. The two-stage observation bias correction reduced the differences in the variance, and the 'bias corrected' observations had spatially averaged standard deviations very close to the model forecasts, of 7.6 K for GOES-West, with a spatial mean absolute difference of just 0.4 K, and of 5.1 K for GOES-East, giving a spatial mean absolute difference of 0.3 K.

b. Global O-F mean difference maps

Figure 5 shows maps of the estimated b^+ at 9:00 UTC on June 1, July 1, and August 1, 2012. There is substantial spatial variation in the b^+, with a clear signal of land surface conditions. There are no obvious discontinuities between the b^+ estimated for adjacent satellites in Figure 5, although the limited regions of overlapping observations from neighboring satellites (at sufficiently small viewing angles) makes the direct assessment of such discontinuities difficult. At 9:00 UTC it is daytime over Africa and Europe, and this region has the largest estimated b^+ in Figure 5, with distinct regions of large positive values ($>$ 10 K) in the drier regions of Africa, the Arabian peninsula, and western Asia, with a band of negative values (< -5 K) over equatorial Africa. In contrast, the regions experiencing nighttime generally have smaller b^+ (magnitude $<$5 K), except for the drier regions of western North America and Australia, with mean differences of 5-10 K. This tendency for very large positive daytime b^+ over dry regions occurs consistently across the globe, particularly in the summer hemisphere; the same pattern was evident in Figure 2 for GOES-West, which observes the arid southwest of the US. In terms of the temporal evolution of the b^+, the large-scale spatial patterns are consistent between the three months plotted in Figure 5, although the gradual evolution of the b^+ estimates is evident.

c. Evaluation against independent T_{skin} observations

Figures 6 and 7 demonstrate that the two-stage observation bias and state estimation filter improved the modeled T_{surf} sub-seasonal-scale variability, compared to independent observations, albeit by a modest amount. In Figure 6 the mean ubRMSD of the assimilation estimates versus SURFRAD observations is reduced at each time of day by between 0.05 K

- 0.31 K (∼5-10%), with the greatest improvements (>0.2 K) occurring during the first half of the day (09:00-15:00 UTC). The ubRMSD across all times of day is significantly reduced (at a 5% level) from 2.1 K to 1.9 K.

Similar results were obtained by comparison to Terra and Aqua MODIS T_{skin} observations over the Americas, as shown in Figure 7. During the night, the open-loop ubRMSD was already very small, with a spatial mean of 1.9 K for both Terra and Aqua. During the day, the open-loop ubRMSD was much larger, except over the Amazon, with a spatial mean of 3.6 K for both Terra and Aqua. For all MODIS overpasses, the assimilation consistently improved the model fit to the MODIS data across the domain, except over the Amazon where the open-loop ubRMSD was already very low and the improvement from the assimilation was weaker, and even slightly negative in places. While the consistency of the positive improvements in Figure 7 is encouraging, these improvements were significant (at the 5% level) over only a small fraction (<10%) of the domain. For each MODIS orbit direction the spatial mean ubRMSD across the domain is shown in Table 2, and in each case the assimilation reduced the spatial mean ubRMSD by around 10% of the open-loop value, with ubRMSD reductions of 0.1 - 0.2 K during the nighttime, and 0.2-0.3 K during the daytime.

While the above evaluation consistently indicates that the T_{skin} assimilation has improved the model T_{surf}, the improvements are rather modest. This is despite the use of only assimilation update times in the evaluation, which will have exaggerated the assimilation impact. There are several reasons for the modest results. Most importantly, the skill of the model T_{surf}, in terms of the anomaly behavior assessed here, is already very good. Also, the Catchment model T_{surf} has an extremely short memory, associated with its very low heat capacity, hence the analysis updates do not persist into the subsequent model time step, and the model has very little memory of improvements previously gained from the assimilation. Including GHT1 in the state update vector did not increase the T_{surf} memory of previous analysis updates, since the T_{surf} dynamics are dominated by the radiation budget. Finally, the lack of memory is compounded by the low data volume associated with the lack of

TIR observations under cloudy conditions. The modest impact of the assimilation is not related to the observation bias correction method, since similar results were obtained using cumulative distribution functions (Reichle and Koster 2004) to rescale the observations (not shown).

5. Summary and conclusions

A two-stage observation bias and state estimation scheme has been developed for use in land data assimilation. In this scheme, the observation-minus-forecast (O-F) mean differences are estimated and removed from the innovations prior to updating the model state. In applications where the model predictions are bias-free, the two-stage filter could also be used to correct the observations towards the true mean state. The presented method is computationally affordable, straightforward to implement in an existing assimilation, requires specification of only a single additional parameter, and can be used to assimilate satellite radiances or retrieved geophysical variables. In contrast to the observation rescaling methods currently used in land data assimilation systems, the two-stage filter does not require a long data record. Hence, it has the potential to facilitate the use and success of land data assimilation, particularly in atmospheric modeling systems for which long records of consistently forecast land surface estimates are typically not available.

The two-stage filter includes a parameterization of the Kalman gain for the bias update that introduces an explicit specification of the time scale of the O-F mean differences. Defining the O-F mean difference over seasonal time scales allows the assimilation to update the model state vector in response to sub-seasonal-scale (e.g., synoptic scale) differences between observed and forecast estimates.

In experiments assimilating geostationary T_{skin} observations into the Catchment land surface model, the two-stage filter effectively removed the O-F mean difference from the observations, and consequently improved synoptic-scale dynamics in the model T_{surf} (the model equivalent variable to T_{skin}). These improvements were measured using the ubRMSD with independent estimates of T_{skin} from the SURFRAD network (at six sites in the US), and from MODIS satellite observations over the Americas. While modest, the improvements highlight the potential value of the geostationary T_{skin} for future modeling efforts.

Global maps of the O-F mean differences estimated by the two-stage filter show clear spatial coherence, with a signal of local land surface conditions. Most prominently, there is a strong tendency for large positive O-F differences in dry regions during the daytime. In this study, the O-F mean difference was estimated independently at each model grid cell. However, the spatial cohesion of the estimates suggests the potential to improve the two-stage filter design by incorporating horizontal information into the observation bias estimates. This could be achieved by either including spatial smoothing in the bias forecast model (assuming correlations between the O-F mean difference in adjacent areas), or by implementing the bias update using a 3-D filter (assuming correlations between the errors in the O-F mean difference estimates).

In addition to the difficulty of obtaining suitable data records for observation rescaling, several studies have highlighted other shortcomings arising from the stationary nature of the observation rescaling approaches for bias correction. For example, the inability of a stationary approach (CDF-matching) to distinguish between near-surface soil moisture variability over seasonal and sub-seasonal time scales can result in inadequate matching of the seasonal cycles between forecast estimates and CDF-matched observations (Draper et al. 2009). Also Drusch et al. (2005) argues that uncertainty in the inter-annual variability of the vegetation characteristics used in both soil moisture retrieval and land surface modeling may necessitate nonstationary observation bias correction methods, based on either frequent updates of observation rescaling coefficients, or the use of more sophisticated methods. More recently, Crow et al. (2011) showed that results from the assimilation of remotely sensed soil moisture into a simple water balance model were improved by using seasonally variable observation rescaling coefficients for adjusting the mean. The nonstationary nature of filtering may also have practical advantages for the estimation of O-F mean differences, in that the estimates can respond to step changes, caused for example, by changes in the forecast model, remote sensor, or retrieval model. Hence, in atmospheric assimilation the ability of variational observation bias correction schemes to respond to temporal changes in the bias has proven beneficial (Auligné et al. 2007; Dee and Uppala 2009).

Unlike observation rescaling, the two-stage filter presented here does not explicitly address systematic differences between higher-order moments of the observations and the model estimates. For the T_{skin} assimilation experiments presented here, the two-stage filter proved sufficient. However, other land surface variables, including near-surface soil moisture, can have large systematic differences in the sub-seasonal-scale variability of observed and forecast estimates. Work is underway to expand the two-stage filter to also account for systematic differences in the higher order moments, thus providing an alternative to observation rescaling for soil moisture data assimilation.

Acknowledgments.

The research was supported by the NASA Modeling, Analysis, and Prediction program, the NASA High-End Computing program, and the NASA Satellite Calibration Interconsistency program. MODIS land surface data were provided by NASA's Earth Observing System Data and Information System, and the SURFRAD data were provided by the NOAA Earth System Research Laboratory.

APPENDIX

Appendix A. Derivation of L_k.

In the bias state update equation (equation 8), the model state, observation bias, and observation estimates can each be partitioned into their true value, a random (zero-mean) error, and for the observations a long term mean error (bias): $x_k^- = x_k + e^{x-}$, and $b_k^- = b_k + e^{b-}$, and $\tilde{y}_k^o = \tilde{y}_k + e_k^o = y_k + b_k + e_k^o$, where e represents the random error in the superscripted variable. To derive L_k, minimize the expected error in b_k^+, $P_k^{b+} = E[e^{b+}(e^{b+})^T]$, where E is the expectation over time. Substituting equation 8 into P_k^{b+}, then partitioning each variable into its constituent parts gives:

$$
\begin{aligned}
P_k^{b+} &= E[(b_k^+ - b_k)(b_k^+ - b_k)^T] && \text{(A1)}\\
&= E[(b_k^- + L_k < \tilde{y}_k{}^o - b_k^- - H_k x_k^- > - b_k)(b_k^- + L_k < \tilde{y}_k{}^o - b_k^- - H_k x_k^- > - b_k)^T] && \text{(A2)}\\
&= E[(e_k^{b-} + L_k < e_k^o - e_k^{b-} - H_k e_k^{x-} >)(e_k^{b-} + L_k < e_k^o - e_k^{b-} - H_k e_k^{x-} >)^T] && \text{(A3)}
\end{aligned}
$$

The derivative of P_k^{b+} w.r.t L_k is:

$$\frac{\delta P_k^{b+}}{\delta L_k} = 2E[(e_k^{b-} + L_k < e_k^o - e_k^{b-} - H_k e_k^{x-} >)(< e_k^o - e_k^{b-} - H_k e_k^{x-} >)^T)] \quad \text{(A4)}$$

Setting the derivative to zero, and solving for L, gives the P_k^{b+} minimum:

$$L_k = E[-e_k^{b-}(< e_k^o - e_k^{b-} - H_k e_k^{x_k-} >)^T(< e_k^o - e_k^{b-} - H_k e_k^{x-} > (< e_k^o - e_k^{b-} - H_k e_k^{x-} >)^T)^{-1}] \quad \text{(A5)}$$

If e_k^o, e_k^{b-}, and e_k^{x-} are not cross-correlated with each other, the expectation is:

$$L_k = P_k^{b-}(R^o + P_k^{b-} + H_k P_k^{x-} H_k^T)^{-1} \tag{A6}$$

Appendix B. Derivation of $\tilde{K}$, and equivalence to K.

To derive $\tilde{K}$ minimize the expected error $x_{k,i}^+$, $P^{x+} = E[(e_k^{x+})(e_k^{x+})^T]$. Substituting equation 11 into P_k^{x+}, and as in Appendix A, partitioning each variable into its constituent terms, gives:

$$P^{x+} = E[(x_k^+ - x_k)(x_k^+ - x_k)^T] \tag{A7}$$

$$\begin{aligned} &= E[(x_k^- + \tilde{K}_k(\tilde{y}_k^o - b_k^- - H_k x_k^-) - \tilde{K}_k L_k < \tilde{y}_k^o - b_k^- - H_k x_k^- > -x_k) \\ &\qquad (x_k^- + \tilde{K}_k(\tilde{y}_k^o - b_k^- - H_k x_k^-) - \tilde{K}_k L_k < \tilde{y}_k^o - b_k^- - H_k x_k^- > -x_k)^T] \end{aligned} \tag{A8}$$

$$\begin{aligned} &= E[(e_k^{x-} + \tilde{K}_k(e_k^o - e_k^{b-} - H_k e_k^{x-}) - \tilde{K}_k L_k < e_k^o - e_k^{b-} - H_k e_k^{x-} >) \\ &\qquad (e_k^{x-} + \tilde{K}_k(e_k^o - e_k^{b-} - H_k e_k^{x-}) - \tilde{K}_k L_k < e_k^o - e_k^{b-} - H_k e_k^{x-} >)^T] \end{aligned} \tag{A9}$$

The derivative of P_k^{x+} w.r.t $\tilde{K}_k$ is:

$$\begin{aligned} \frac{\delta P_k^{x+}}{\delta \tilde{K}_k} &= 2E[(e_k^{x-} + \tilde{K}_k(e_k^o - e_k^{b-} - H_k e_k^{x-}) - \tilde{K}_k L_k < e_k^o - e_k^{b-} - H_k e_k^{x-} >) \\ &\qquad (e_k^o - e_k^{b-} - H_k e_k^{x-} - L_k < e_k^o - e_k^{b-} - H_k e_k^{x-} >)^T] \end{aligned} \tag{A10}$$

If e^o, e^{x-}, and e^{b-} are not cross-correlated with each other, setting the derivatives to zero to minimize P_k^{x+}, and taking the expectation gives:

$$\tilde{K}_k(I - L_k) = P_k^{x-} H_k^T (R^o + P_k^{b-} + H_k P_k^{x-} H_k^T)^{-1} \tag{A11}$$

Substituting equation 9 into A11 gives:

$$\tilde{K}_k(I - P_k^{b-}(R^o + P_k^{b-} + H_k P_k^{x-} H_k^T)^{-1}) = P_k^{x-} H_k^T (R^o + P_k^{b-} + H_k P_k^{x-} H_k^T)^{-1} \tag{A12}$$

$$\tilde{K}_k(R^o + P_k^{b-} + H_k P_k^{x-} H_k^T - P_k^{b-}) = P_k^{x-} H_k^T \tag{A13}$$

$$\tilde{K} = P_k^{x-} H_k^T (R^o + H_k P_k^{x-} H_k^T)^{-1} \tag{A14}$$

which is the same as equation 4 for the Kalman gain for the bias-free EnKF state update. This demonstrates that the inclusion of the observation bias estimate from the two-stage state and bias estimation does not change the expression of the solution for the Kalman gain for the state update in equation 10 (assuming that equation 9 is used for L_k).

REFERENCES

Augustine, J., G. Hodges, C. Cornwall, J. Michalsky, and C. Medina, 2005: An update on SURFRAD: The GCOS Surface Radiation Budget Network for the continental United States. *Journal of Atmospheric Oceanic Technology*, **22**, 1460–1472, doi: 10.1175/JTECH1806.1.

Auligné, T., A. McNally, and D. Dee, 2007: Adaptive bias correction for satellite data in a numerical weather prediction system. *Q. J. R. Meteorol. Soc.*, **133**, 631–642, doi: 10.1002/qj.56.

Bosilovich, M., J. Radakovich, A. da Silva, R. Todling, and F. Verter, 2007: Skin temperature analysis and bias correction in a coupled land-atmosphere data assimilation system. *Journal of the Meteorological Society of Japan*, **85A**, 205–228, doi:10.2151/jmsj.85A.205.

Chepurin, G., J. Carton, and D. Dee, 2005: Forecast model bias correction in ocean data assimilation. *Mon. Wea. Rev.*, **133**, 1328–1342, doi:10.1175/MWR2920.1.

Crow, W., M. van den Berg, G. Huffman, and T. Pellarin, 2011: Correcting rainfall using satellite-based surface soil moisture retrievals: The Soil Moisture Analysis Rainfall Tool (SMART). *Water Resources Research*, **47**, W08 521, doi:10.1029/2011WR010576.

Crow, W., et al., 2012: Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products. *Reviews of Geophysics*, **50**, RG2002, doi:10.1029/2011RG000372.

De Lannoy, G., R. Reichle, K. Arsenault, P. Houser, S. Kumar, N. Verhoest, and V. Pauwels, 2012: Multi-scale assimilation of AMSR-E snow water equivalent and MODIS snow cover fraction in northern Colorado. *Water Resources Research*, **48**, W01 522.

De Lannoy, G., R. Reichle, P. Houser, V. Pauwels, and N. Verhoest, 2007: Correcting for forecast bias in soil moisture assimilation with the ensemble Kalman filter. *Water Resources Research*, **43**, W09 410, doi:10.1029/2006WR005449.

de Rosnay, P., M. Drusch, D. Vasiljevic, G. Balsamo, C. Albergel, and L. Isaksen, 2013: A simplified Extended Kalman Filter for the global operational soil moisture analysis at ECMWF. *Quarterly Journal of the Royal Meteorological Society*, **139**, 1199–1213, doi: 10.1002/qj.2023.

Dee, D. and A. Da Silva, 1998: Data assimilation in the presence of forecast bias. *Q.J.R. Meteorol. Soc.*, **124**, 269–295, doi:10.1002/qj.49712454512.

Dee, D. and R. Todling, 2000: Data assimilation in the presence of forecast bias: The GEOS moisture analysis. *Mon. Wea. Rev.*, **128**, 3268–3282, doi:{10.1175/1520-0493(2000)128⟨3268:DAITPO⟩2.0.CO;2}.

Dee, D. and S. Uppala, 2009: Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Q. J. R. Meteorol. Soc.*, **135**, 1830–1841, doi:10.1002/qj.493.

Dharssi, I., K. Bovis, B. Macpherson, and C. Jones, 2011: Operational assimilation of ASCAT surface soil wetness at the Met Office. *Hydrology and Earth System Sciences*, **15**, 2729–2746, doi:0.5194/hess-15-2729-2011.

Draper, C., J.-F. Mahfouf, and J. Walker, 2009: An EKF assimilation of AMSR-E soil moisture into the ISBA land surface scheme. *Journal of Geophysical Research*, **114**, D20 104, doi:10.1029/2008JD011650.

Draper, C., R. Reichle, R. de Jeu, V. Naeimi, R. Parinussa, and W. Wagner, 2013: Estimating root mean square errors in remotely sensed soil moisture over continental scale domains. *Remote Sensing of Environment*, **137**, 288–298, doi:10.1016/j.rse.2013.06.013.

Draper, C., R. Reichle, G. De Lannoy, and Q. Liu, 2012: Assimilation of passive and active microwave soil moisture retrievals. *Geophysical Research Letters*, **39**, L04 401, doi:10.1029/2011GL050655.

Drusch, M., E. Wood, and H. Gao, 2005: Observation operators for the direct assimilation of TRMM microwave imager retrieved soil moisture. *Geophysical Research Letters*, **32**, L15 403, doi:10.1029/2005GL023623.

Fertig, E., et al., 2009: Observation bias correction with an ensemble Kalman filter. *Tellus A*, **61**, 210–226, doi:10.1111/j.1600-0870.2008.00378.x.

Friedland, B., 1969: Treatment of bias in recursive filtering. *IEEE Transactions on Automatic Control,*, **14**, 359–367, doi:10.1109/TAC.1969.1099223.

Ghent, D., J. Kaduk, J. Remedios, J. Ardö, and H. Balzter, 2010: Assimilation of land surface temperature into the land surface model JULES with an ensemble Kalman filter. *Journal of Geophysical Research*, **115**, D19 112, doi:10.1029/2010JD014392.

Harris, B. and G. Kelly, 2001: A satellite radiance-bias correction scheme for data assimilation. *Q.J.R. Meteorol. Soc.*, **127**, 1453–1468, doi:10.1002/qj.49712757418.

Keppenne, C., M. Rienecker, N. Kurkowski, and D. Adamec, 2005: Ensemble Kalman filter assimilation of temperature and altimeter data with bias correction and application to seasonal prediction. *Nonlinear Processes in Geophysics*, **12**, 491–503.

Koster, R., M. 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. *Journal of Geophysical Research*, **105**, 24 809–24 822, doi:10.1029/2000JD900327.

Minnis, P., L. Nguyen, D. Doelling, D. Young, W. Miller, and D. Kratz, 2002: Rapid Calibration of Operational and Research Meteorological Satellite Imagers. Part II: Comparison of Infrared Channels. *J. Atmos. Oceanic Technol.*, **19**, 1250–1266, doi:10.1175/1520-0426(2002)019⟨1250:RCOOAR⟩2.0.CO;2.

Pauwels, V., G. De Lannoy, H.-J. Hendricks Franssen, and H. Vereecken, 2013: Simultaneous estimation of model state variables and observation and forecast biases using a two-stage hybrid kalman filter. *Hydrol. Earth Syst. Sci.*, **17**, 3499–3521, doi:10.5194/hess-17-3499-2013.

Reichle, R., G. De Lannoy, B. Forman, C. Draper, and Q. Liu, 2013: Connecting satellite observations with water cycle variables through land data assimilation: Examples using the NASA GEOS-5 LDAS. *Surveys in Geophysics*, **35**, 577–606, doi:0.1007/s10712-013-9220-8.

Reichle, R. and R. Koster, 2003: Assessing the impact of horizontal error correlations in background fields on soil moisture estimation. *Journal of Hydrometeorology*, **4**, 1229–1242, doi:10.1175/1525-7541(2003)004⟨1229:ATIOHE⟩2.0.CO;2.

Reichle, R. and R. Koster, 2004: Bias reduction in short records of satellite soil moisture. *Geophysical Research Letters*, **31**, L19 501, doi:10.1029/2004GL020938.

Reichle, R., R. Koster, J. Dong, and A. Berg, 2004: Global soil moisture from satellite observations, land surface models, and ground data: Implications for data assimilation. *Journal of Hydrometeorology*, **5**, 430 – 442, doi:10.1175/1525-7541(2004)005⟨0430:GSMFSO⟩2.0.CO;2.

Reichle, R., R. Koster, P. Liu, S. Mahanama, E. Njoku, and M. Owe, 2007: Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR). *Journal of Geophysical Research* , **112**, D09 108, doi:{10.1029/2006JD008033}.

Reichle, R., S. Kumar, S. Mahanama, R. Koster, and Q. Liu, 2010: Assimilation of satellite-derived skin temperature observations into land surface models. *Journal of Hydrometeorology*, **11**, 1103–1122, doi:10.1175/2010JHM1262.1.

Rienecker, M., et al., 2008: The GEOS-5 Data Assimilation System - Documentation of Versions 5.0.1, 5.1.0, and 5.2.0. *Technical Report Series on Global Modeling and Data Assimilation*, **27**.

Rienecker, M., et al., 2011: MERRA - NASA's Modern-Era Retrospective Analysis for Research and Applications. *Journal of Climate*, **24**, 3624–3648, doi:10.1175/JCLI-D-11-00015.1.

Scarino, B., P. Minnis, R. Palikonda, R. Reichle, D. Morstad, C. Yost, B. Shan, and Q. Liu, 2013: Retrieving clear-sky surface skin temperature for numerical weather prediction applications from geostationary satellite data. *Remote Sensing*, **5**, 342–366, doi:10.3390/rs5010342.

Scipal, K., M. Drusch, and W. Wagner, 2008: Assimilation of a ERS scatterometer derived soil moisture index in the ECMWF numerical weather prediction system. *Advances in Water Resources*, **31**, 1101–1112, doi:10.1016/j.advwatres.2008.04.013.

Slater, A. and M. Clark, 2006: Snow Data Assimilation via an Ensemble Kalman Filter. *J. Hydrometeor.*, **7**, 478–493, doi:10.1175/JHM505.1.

Vinnikov, K., Y. Yu, M. Rama Varma Raja, D. Tarpley, and M. Goldberg, 2008: Diurnal-seasonal and weather-related variations of land surface temperature observed from geostationary satellites. *Geophysical Research Letters*, **35**, L22 708, doi:10.1029/2008gl035759.

Wang, A., M. Barlage, X. Zeng, and C. Draper, 2014: Comparison of land skin temperature from a land model, remote sensing, and in-situ measurement. *Journal of Geophysical Research*, **119**, doi:10.1002/2013JD021026.

Wang, K., Z. Wan, P. Wang, M. Sparrow, J. Liu, X. Zhou, and et al., 2005: Estimation of surface long wave radiation and broadband emissivity using Moderate Resolution Imaging Spectroradiometer (MODIS) land surface temperature/emissivity products. *Journal of Geophysical Research*, **110**, D11 109, doi:10.1029/2004JD005566.

6. Figures

7. Tables

List of Tables

TABLE 1. Ensemble Generation Perturbation Parameters for Forcing and Model Prognostic Variables.

	(A)dditive, or (M)ultiplicative	Standard Deviation	AR(1) Time Scale	Perturbation cross-correlation			
				GHT1	T2m	SW	LW
T_{surf}	A	0.2 K	12 hours	0.7	0	0	0
GHT1	A	50,000 J	12 hours	-	0	0	0
2m air temp (T2m)	A	1 K	24 hours		-	0.4	0.4
SW radiation	M	0.3	24 hours			-	-0.6
LW radiation	A	20 Wm^{-2}	24 hours				-

Table 2. Spatial Mean of the ubRMSD (K) with MODIS T_{skin} Reported in Figure 7.

Experiment	MODIS overpass			
	Nighttime		Daytime	
	Terra	Aqua	Terra	Aqua
Open-loop	1.89	1.94	3.62	3.60
T_{skin} assimilation	1.70	1.79	3.36	3.42
Difference	0.19	0.15	0.27	0.18

List of Figures

5 Observation-minus-forecast T_{skin} mean difference, estimated at 09:00 UTC on first a) June, b) July, and c) August, 2012. Values are shown only where the observation bias estimate is considered valid for use in the state update equation. The plotted meridians demark the domain of each satellite: [-175°,-105°] GOES-West, [-105°,-37°] GOES-East,[-37°, 54°] MTSAT-2, [54°,90°] FY-2E, and [90°,-175°] Met-9. 44

6 ubRMSD with SURFRAD T_{skin}, calculated separately for each SURFRAD site and each observation time, for the assimilation of geostationary observations with the two-stage filter (filled circles), and the open-loop (unfilled circles). The mean ubRMSD at each time of day for the assimilation (open-loop) is indicated by the solid (dashed) line. 45

7 ubRMSD with MODIS T_{skin} for the open-loop (upper), and the improvement in the ubRMSD gained from the assimilating geostationary T_{skin} with the two stage filter (lower: Δ ubRMSD=ubRMSD of open-loop - ubRMSD of assimilation), separately for each Terra and Aqua overpass direction. Grey indicates < 30 coincident geostationary and MODIS T_{skin} observations. The plotted meridians demark the GOES-West and GOES-East domains. 46

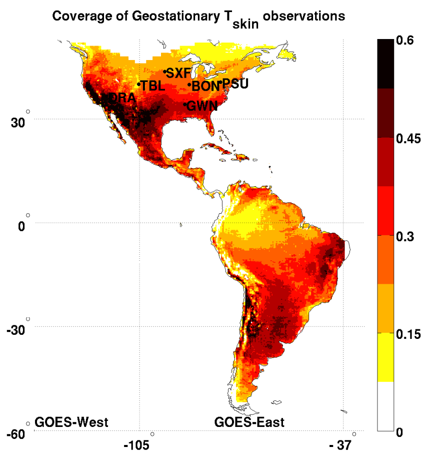

FIG. 1. Coverage of the assimilated GOES-West and GOES-East T_{skin} observations from 1 June, 2012 to 31 May, 2013, as a fraction of the maximum possible coverage (eight observations every day). The locations of the SURFRAD measurement stations are marked as DRA (Desert Rock), TBL (Table Mountain), SXF (Sioux Falls), GWN (Goodwin Creek), BON (Bondville), and PSU (Penn State). The plotted meridians demark the GOES-West and GOES-East domains.

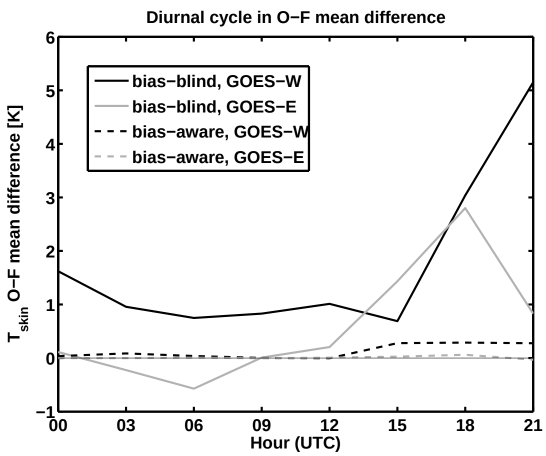

FIG. 2. Diurnal cycle of the T_{skin} O-F mean difference, averaged over the Americas, for a bias-blind assimilation (solid) and the two-stage observation bias and state estimation bias-aware assimilation with τ =20 days (dashed), for GOES-West (black) and GOES-East (grey).

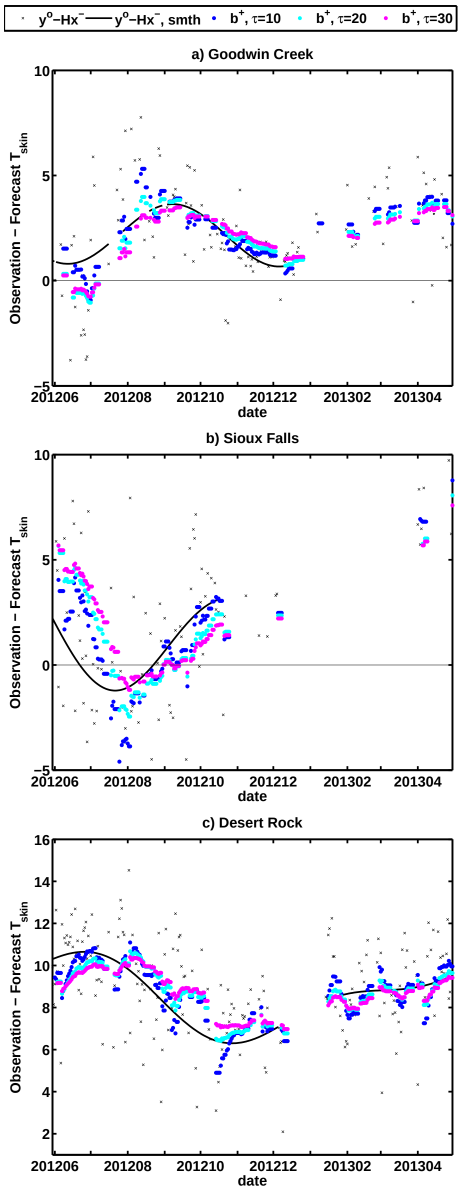

FIG. 3. The T_{skin} O-F residuals [K] at 21:00 UTC (black crosses) at the a) Goodwin Creek, b) Sioux Falls, and c) Desert Rock SURFRAD sites. Black lines show the smoothed O-F time series using the first two annual Fourier harmonics. Dots show the bias estimates from the two-stage observation bias correction scheme using (dark blue) τ=10 days, (light blue) $\tau = 20$ days, and (pink) τ=30 days.

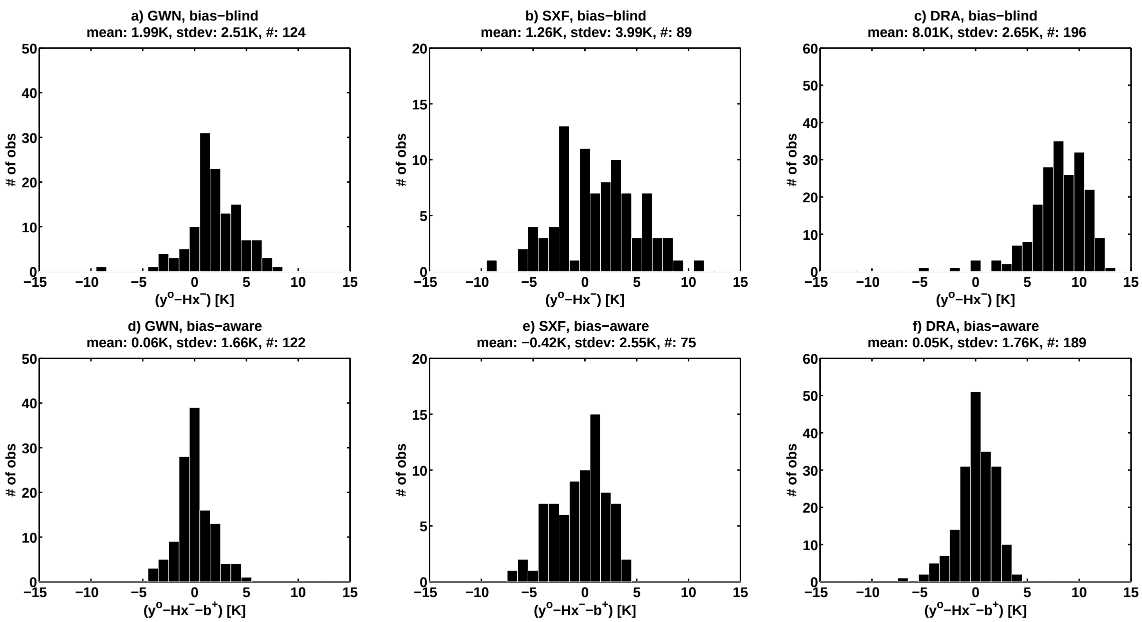

FIG. 4. Histograms of the state update innovations at 21:00 UTC, for the assimilation of geostationary T_{skin}, at the Goodwin Creek (GWN), Sioux Falls (SXF), and Desert Rock (DRA) SURFRAD sites, for a bias-blind assimilation (upper), and for the two-stage observation bias and state estimation bias-aware assimilation with τ=20 days (lower).

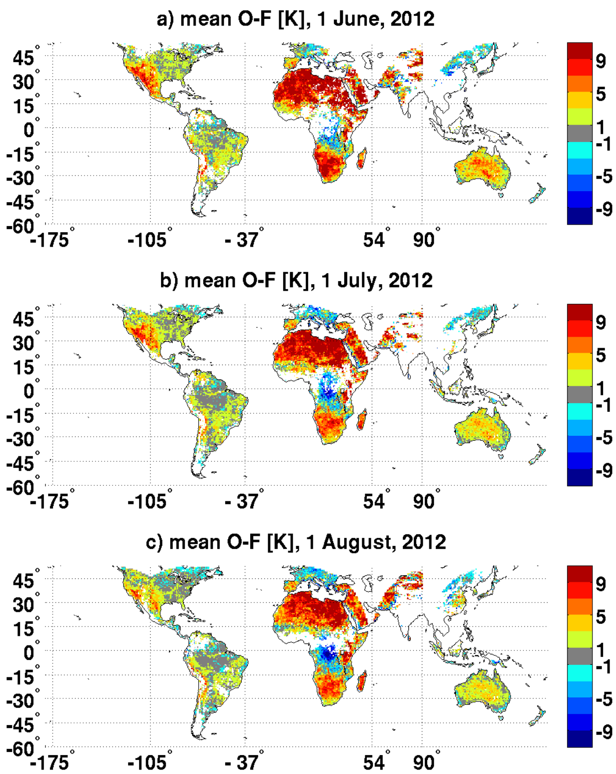

FIG. 5. Observation-minus-forecast T_{skin} mean difference, estimated at 09:00 UTC on first a) June, b) July, and c) August, 2012. Values are shown only where the observation bias estimate is considered valid for use in the state update equation. The plotted meridians demark the domain of each satellite: [-175°,-105°] GOES-West, [-105°,-37°] GOES-East,[-37°, 54°] MTSAT-2, [54°,90°] FY-2E, and [90°,-175°] Met-9.

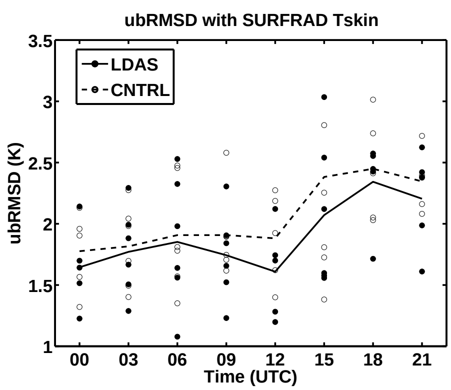

FIG. 6. ubRMSD with SURFRAD T_{skin}, calculated separately for each SURFRAD site and each observation time, for the assimilation of geostationary observations with the two-stage filter (filled circles), and the open-loop (unfilled circles). The mean ubRMSD at each time of day for the assimilation (open-loop) is indicated by the solid (dashed) line.

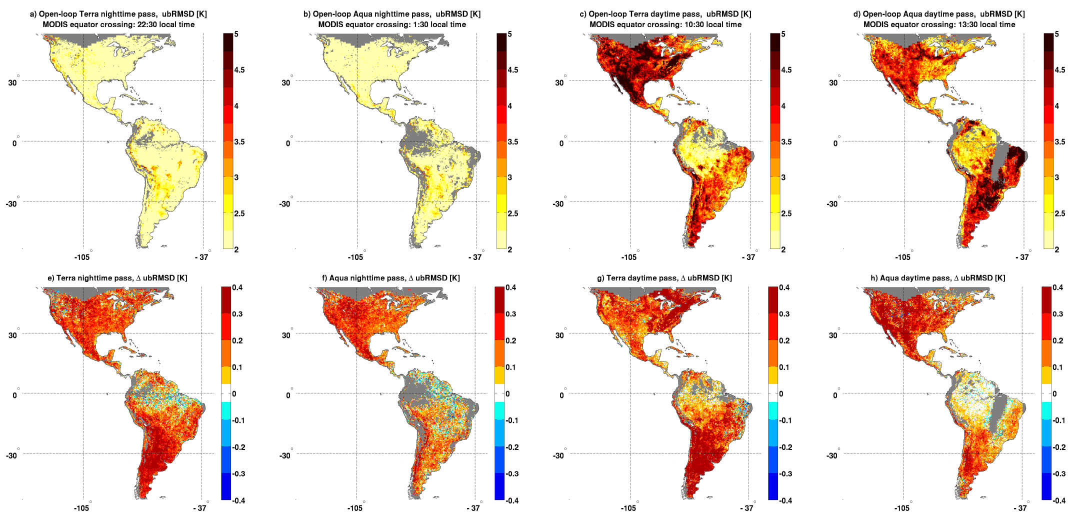

FIG. 7. ubRMSD with MODIS T_{skin} for the open-loop (upper), and the improvement in the ubRMSD gained from the assimilating geostationary T_{skin} with the two stage filter (lower: Δ ubRMSD=ubRMSD of open-loop - ubRMSD of assimilation), separately for each Terra and Aqua overpass direction. Grey indicates $<$ 30 coincident geostationary and MODIS T_{skin} observations. The plotted meridians demark the GOES-West and GOES-East domains.