Correcting Land Surface Model Predictions for the Impact of Temporally Sparse Rainfall Rate Measurements Using an Ensemble Kalman Filter and Surface Brightness Temperature Observations

WADE T. CROW

Hydrology and Remote Sensing Laboratory, USDA/ARS, Beltsville, Maryland

(Manuscript received 13 January 2003, in final form 25 April 2003)

ABSTRACT

Current attempts to measure short-term (<1 month) rainfall accumulations using spaceborne radiometers are characterized by large sampling errors associated with low observation frequencies for any single point on the globe (from two to eight measurements per day). This degrades the value of spaceborne rainfall retrievals for the monitoring of surface water and energy balance processes. Here a data assimilation system, based on the assimilation of surface L-band brightness temperature (T_B) observations via the ensemble Kalman filter (EnKF), is introduced to correct for the impact of poorly sampled rainfall on land surface model predictions of root-zone soil moisture and surface energy fluxes. The system is evaluated during the period from 1 April 1997 to 31 March 1998 over two sites within the U.S. Southern Great Plains. This evaluation includes both a data assimilation experiment, based on synthetically generated T_B measurements, and the assimilation of real T_B observations acquired during the 1997 Southern Great Plains Hydrology Experiment (SGP97). Results suggest that the EnKF-based assimilation system is capable of correcting a substantial fraction (>50%) of model error in root-zone (40 cm) soil moisture and latent heat flux predictions associated with the use of temporally sparse rainfall measurements as forcing data. Comparable gains in accuracy are demonstrated when actual T_B measurements made during the SGP97 experiment are assimilated.

1. Introduction

Over large portions of the globe, the accuracy of rainfall data is a major source of error for efforts to monitor and/or predict surface water and energy balance processes. Rainfall rate measurements are based primarily on four observational techniques: ground-based radar, rain gauge observations, satellite-based infrared and visible observations, and satellite-based radiometer observations. With the exception of the United States and Europe, land-based radar systems are not well developed. Gauge-based networks are more extensive but still lacking in many tropical and arid regions of the world (New et al. 2001). Infrared retrievals from geostationary satellites can provide frequent observations over continental-scale regions, but the retrieval is based on the use of cloud height as a surrogate for rainfall rate. As such, rate retrievals are indirect and prone to error—especially over short temporal scales and at high latitudes (Huffman et al. 2001). Spaceborne radiometers offer perhaps the most attractive option for global-scale coverage, but current spaceborne systems [e.g., the Tropical Rainfall Measuring Mission (TRMM) and the Special Sensor Microwave Imager (SSM/I)] make observations of a given point at too infrequent a temporal rate (twice daily) to provide accurate accumulation estimates at timescales shorter than monthly. Providing sufficient temporal sampling to facilitate retrieval of accurate accumulations over finer timescales is a major objective of the proposed Global Precipitation Mission (GPM) satellite constellation scheduled to launch in 2007 (Flaming et al. 2001). As currently envisioned, the system will provide approximately eight retrievals of rainfall rates per day for most areas of the globe within a relative accuracy of about 25% (Adams et al. 2002). Because offline global land surface modeling efforts [see, e.g., the Global Land Data Assimilation System (Rodell et al. 2003, manuscript submitted to *Bull. Amer. Meteor. Soc.*) and the United States Air Force Weather Agency's Agricultural Meteorological Model (AGRMET) project] rely heavily on satellite-based observations of rainfall for forcing data, the improved sampling and accuracy characteristics of the GPM constellation should allow for enhanced global-scale monitoring of surface water and energy balance processes.

Given the long list of observation techniques, and the absence of any currently viable method without at least one shortcoming, it is natural that data-fusion strategies

Corresponding author address: Dr. Wade T. Crow, Hydrology and Remote Sensing Laboratory, USDA/ARS, RM 104, Building 007, BARC-W, Beltsville, MD 20705-2350.
E-mail: wcrow@hydrolab.arsusda.gov

involving one or more observational technique have developed. Examples within the United States include the merging of ground-based radar and rain gauge measurements in Stage-III precipitation products produced by regional National Oceanic and Atmospheric Administration (NOAA) River Forecasting Centers and quilted into a national Stage-IV product by the National Centers for Environmental Prediction (see information online at http://www.emc.ncep.noaa.gov/mmb/stage4/). Over regions of the globe not covered by extensive ground-based radar or gauge networks, the best current source of rainfall measurements are obtained by merging high-frequency, but uncertain, infrared rainfall rate retrievals with more direct, but temporally sparser, spaceborne radiometer observations (Todd et al. 2001). A merged 1° lat $\times$ 1° lon daily (1DD) product based on infrared retrievals from the Television Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS) and the Geostationary Operational Environmental Satellite (GOES) and passive microwave measurements from SSM/I is currently being archived as part of the Global Precipitation Climatology Project (GPCP) (Huffman et al. 2001).

Additional opportunities for the fusion of various observational sources are afforded by the development of data assimilation systems designed to update dynamic land surface models with remote observations. A particularly powerful aspect of these systems is their ability to integrate observations that are only indirectly related to rainfall. For instance, over lightly vegetated regions of the globe, L-band surface brightness temperature (T_B) observations can be inverted to produce estimates of near-surface (5 cm) soil moisture. Potential spaceborne sources of high-frequency (1–3-day repeat time) global L-band T_B observations in the near future include the European Space Agency's (ESA) Soil Moisture and Ocean Salinity Mission (SMOS) scheduled for launch in 2006 and the Hydrospheric States (HYDROS) mission currently slated as a back-up mission in the National Aeronautics and Space Administration's (NASA's) Earth System Science Pathfinder (ESSP) program. Spaceborne T_B observations provide an indirect measure of antecedent rainfall that, if properly interpreted by the land data assimilation system, could correct land surface model predictions for the impact of rainfall inputs derived from the sparse sampling of an intermittent rainfall event. It is important to note that such systems will not correct rainfall estimates themselves. Rather, their aim would be to mitigate the impact of rainfall-forcing errors on a land surface model's representation of surface state and flux variables.

The purpose of this analysis is to examine the potential of a data assimilation system designed around the ensemble Kalman filter (EnKF) to combine temporally sparse rainfall rate measurements—ostensibly from a spaceborne source–with T_B observations driven by near-surface soil moisture conditions. The application of the EnKF to the assimilation of T_B observations into a land surface model was first described in Reichle et al. (2002a). Recent work has successfully extended the application of the EnKF to cases involving real T_B observations acquired during the 1997 Southern Great Plains Hydrology Experiment (SGP97) (Margulis et al. 2002; Crow and Wood 2003). The approach used here is based on the generation of a set of precipitation realizations (conditioned by both climatological expectations and temporally sparse rainfall observations) to force an ensemble of land surface state forecasts. These forecasts are, in turn, updated using a remote observation of L-band brightness temperature and the standard Kalman filter update equation. Preliminary work in Crow and Wood (2003) describes an EnKF-based data assimilation system where the model forecast ensemble is based solely on climatological expectations concerning precipitation. This implies that no observations of antecedent rainfall are available at EnKF update times. Even for remote portions of the globe such an assumption is unrealistically restrictive. The development of global spaceborne retrieval technologies virtually guarantees that, however incomplete or temporally sparse, some type of rainfall observation will be available to condition likelihood concerning recent rainfall accumulations beyond what is possible from purely climatological considerations. Consequently, this paper expands on earlier work by presenting a data assimilation system capable of integrating both surface T_B observations and sparse rainfall accumulations derived from the temporal sampling of rainfall rates at frequencies consistent with expectations for current and next-generation spaceborne radiometer systems (2–12 observations per day). As a first approach for proof of concept, initial results are generated using an identical twin data assimilation methodology where observations are synthetically generated by a model, perturbed with noise, and then assimilated back into the same model. Additional results incorporating real T_B observations made during SGP97 are also presented.

2. Data assimilation system

The data assimilation system examined here consists of three parts: a land surface model to forecast land surface model states, an observation model to predict surface brightness temperatures (T_B) based on land surface model state predictions, and an EnKF to update state forecasts with T_B observations. The land surface and observation models are presented in sections 2a and 2b. As described in section 2c, the EnKF is based on the use of a model ensemble to temporally propagate the error covariance information required by the state update equation of the standard Kalman filter. For this particular application of the EnKF, the forecast ensemble is generated using a set of rainfall realizations consistent with the likelihood of various daily rainfall accumulations given climatological considerations and the sparse subsampling of actual rainfall at temporal frequencies

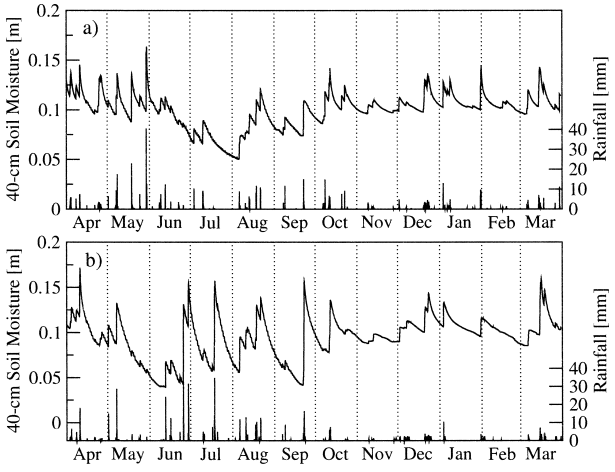

FIG. 1. (a) Times series of TOPLATS simulated 40-cm soil moisture results and daily precipitation at the NOAA ATDD Little Washita site between 1 Apr 1997 and 31 Mar 1998. (b) Same as (a), but for the ARM CART EF site 13.

expected in next-generation spaceborne systems. The statistical procedure used to create these rainfall realizations is detailed in section 2d.

a. Land surface modeling and site descriptions

Numerical modeling of the land surface was based on Topographically-based Land Atmosphere Transfer Scheme (TOPLATS) (Famiglietti and Wood 1994; Peters-Lidard et al. 1997) simulations. The model has been successfully applied to regions within the Southern Great Plains (SGP) by a number of studies (see, e.g., Peters-Lidard et al. 2001; or Crow and Wood 2002). The top soil moisture layer in TOPLATS was set to a depth of 5 cm to reflect the expected vertical depth of sensitivity for remote L-band T_B observations. Simulations were run between 1 April 1997 and 31 March 1998 on an hourly time step at the Department of Energy's Atmospheric Radiation Measurement (ARM) Cloud and Radiation Testbed (CART) extended facility site 13 (EF13; 36°36′N, 97°29′W) near Lamont, Oklahoma, and the NOAA Atmospheric Turbulence and Diffusion Division (ATDD) Little Washita watershed site (34°58′N, 97°57′W) near Chickasha, Oklahoma. Peak leaf area index (LAI) values during the growing season were adjusted to fit local energy flux observations. Simulated soil moisture time series are shown in Fig. 1 and model validation results for both sites are shown in Figs. 2 and 3. Additional details on the application of TOPLATS to these sites can be found in Crow and Wood (2003). Despite their proximity, the two sites demonstrate substantially different root-zone soil water dynamics during the 1997 growing season. The ARM CART EF13 site was usually wet due to the impact of several large precipitation events in June and July. Conditions at the NOAA ATDD Little Washita site are more typical of the region's climatology, with an absence of sustained precipitation during June and July leading to a strong dry down during the midsummer months. In general, denser vegetation at the ARM CART EF13 site produces relatively more transpiration and faster root-zone dry-down dynamics.

b. Microwave emission modeling

Surface (5 cm) soil moisture and soil temperature predictions made by TOPLATS were processed through the Land Surface Microwave Emission Model (LSMEM) (Crow et al. 2001) to produce corresponding estimates of L-band surface brightness temperature (T_B). Microwave radiative transfer parameters for both sites were taken from Jackson et al. (1999). Brightness temperature validation results in Fig. 2 and 3 are based on airborne electronically scanning thinned array radiometer (ESTAR) measurements made during the SGP97 experimental period (17 June to 16 July 1997). The ESTAR measurements have a pixel resolution of 800 m. However, to account for georegistration uncertainties, ESTAR T_B measurements used here were derived from averaging observations within a 3×3 pixel window centered on each site. It is worth noting that comparisons between 5-cm soil moisture gravimetric soil moisture samples in the immediate vicinity of the ARM CART EF13 site and TOPLATS predictions are quite good. This suggests that the pronounced negative bias in LSMEM T_B predictions shown in Fig. 3 is at least partially a scale effect reflecting the difference in support between the plot-scale model simulation and the field-scale ESTAR retrieval (Crow and Wood 2003).

c. The ensemble Kalman filter

The EnKF uses an ensemble-based Monte Carlo approach to temporally propagate error covariance information required by the standard Kalman filter (KF) for updating of model predictions with observations (Evenson 1994; Reichle et al. 2002a). Using notation presented in Reichle et al. (2002a), this section describes the state space representation of the EnKF. Take $\mathbf{Y}(t)$ to be a vector of model state variables at time t. The potentially nonlinear prognostic equation $\mathbf{f}$ describing the temporal evolution of these states can be represented as

$$\frac{d\mathbf{Y}}{dt} = \mathbf{f}(\mathbf{Y}, \mathbf{w}). \tag{1}$$

The error term $\mathbf{w}$ originates from inadequacies in model physics, poor parameter selection, and/or errors in model forcing data. Let the operator $\mathbf{M}$ represent the measurement operator that converts the state variables in $\mathbf{Y}$ into a measurements taken at time t_k:

$$\mathbf{Z}_k = \mathbf{M}[\mathbf{Y}(t_k)] + \mathbf{v}_k, \tag{2}$$

where $\mathbf{v}_k$ represents additive Gaussian measurement noise with covariance $\mathbf{C}_{vk}$. The EnKF is based on the Monte Carlo generation of an ensemble of $\mathbf{Y}$ predictions

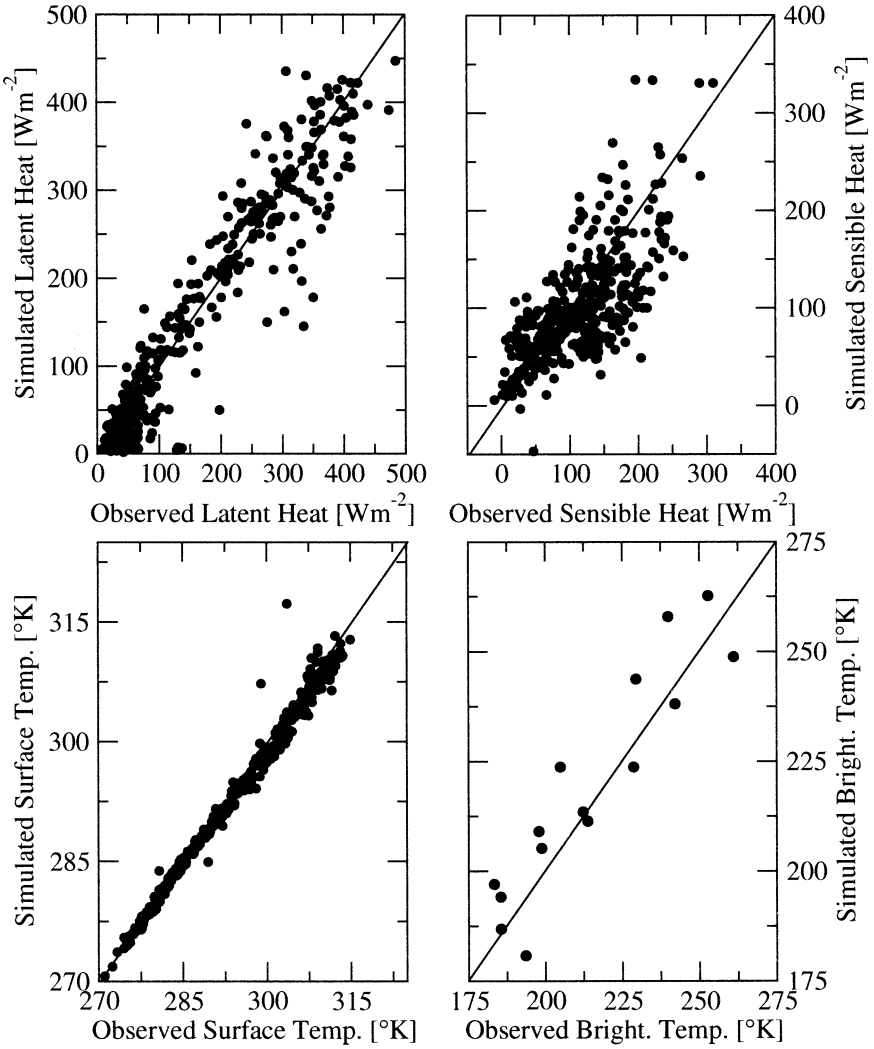

FIG. 2. TOPLATS and LSMEM validation results for the ARM CART EF13 site near Lamont, OK. Plotted energy flux and surface temperature results are averaged values between 1000 and 1600 local time. Observed brightness temperatures are taken from L-band ESTAR retrievals during SGP97.

via (1). This requires an a priori assumption concerning the nature and structure of model error represented by **w** and the perturbation of individual model realizations (or forecasts) with an appropriate statistical representation of this error. At each measurement time, state predictions made by the *i*th model realization within the ensemble are referred to as the forecast $\mathbf{Y}^{\mathbf{i}}_{-}$. If **f** is linear and all errors are additive, independent, and Gaussian, the optimal updating of $\mathbf{Y}^{\mathbf{i}}_{-}$ by the actual measurement $\mathbf{Z}_k$ is given by

$$\mathbf{Y}^{\mathbf{i}}_{+} = \mathbf{Y}^{\mathbf{i}}_{-} + \mathbf{K}_k[\mathbf{Z}^{\mathbf{i}}_k - M_k(\mathbf{Y}^{\mathbf{i}}_{-})], \quad \text{and} \tag{3}$$

$$\mathbf{K}_k = [\mathbf{C}_{YM}(\mathbf{C}_M + \mathbf{C}_v)^{-1}]_{t=t_k}, \tag{4}$$

where $\mathbf{C}_M$ is the error covariance matrix of the measurement forecasts $\mathbf{M}_k(\mathbf{Y}^{\mathbf{i}}_{-})$, $\mathbf{C}_{YM}$ is the cross-covariance matrix between the predicted measurements and state variables contained $\mathbf{Y}^{\mathbf{i}}_{-}$, and $\mathbf{Y}^{\mathbf{i}}_{+}$ is the updated (or analysis) state representation. For each ensemble realization, an additive noise term consistent with $\mathbf{v}_k$ must be synthetically generated and added to $\mathbf{Z}_k$ to produce the perturbed observation $\mathbf{Z}^{\mathbf{i}}_k$. This is required to ensure that the spread of the updated state ensemble accurately reflects the true state error covariance (Burgers et al. 1998). From the ensemble, $\mathbf{C}_M$ and $\mathbf{C}_{YM}$ are statistically estimated by calculating covariances around the ensemble mean. Each ensemble replicate is updated using (3) and (4) and then allowed to evolve via (1) until the next measurement time. Filter predictions are typically derived from averaging model forecasts across the ensemble.

d. Conditioned rainfall likelihood distributions

A statistical representation of error in daily rainfall accumulations associated with the sparse temporal sam-

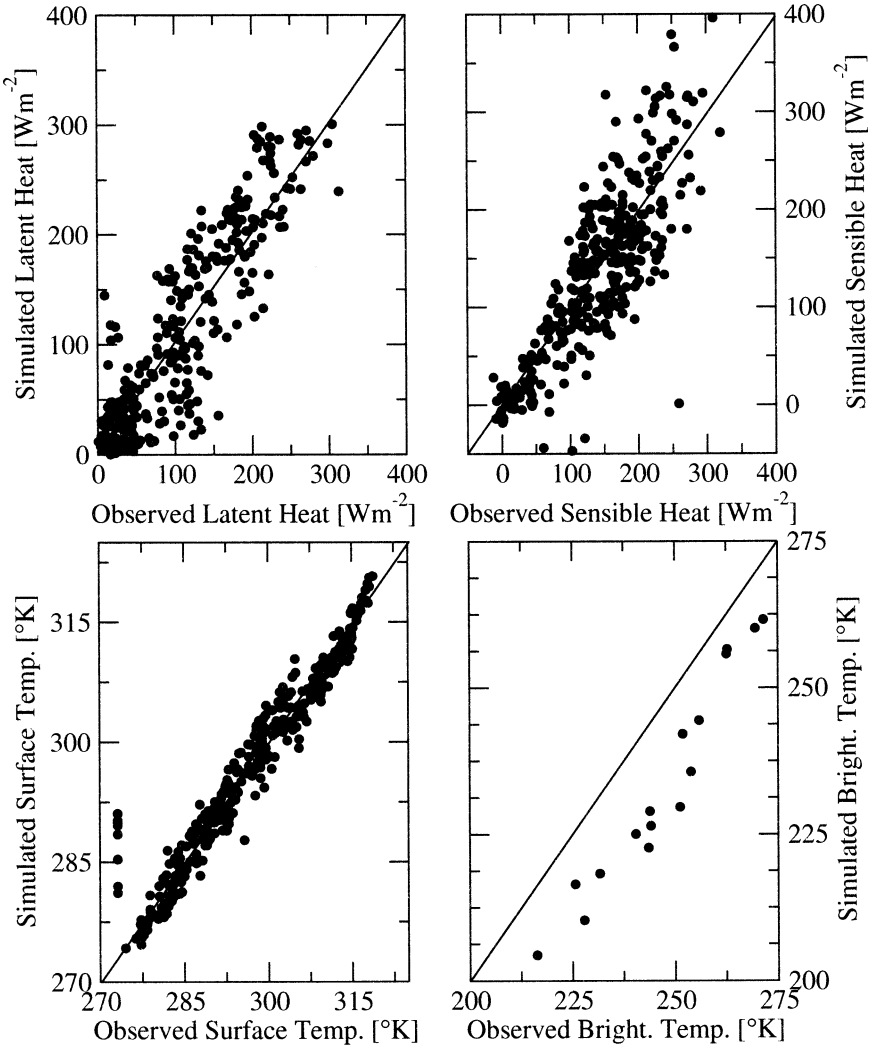

FIG. 3. TOPLATS and LSMEM validation results for the NOAA ATDD Little Washita site near Chickasha, OK. Plotted energy flux and surface temperature results are averaged values between 1000 and 1600 local time. Observed brightness temperatures are taken from L-band ESTAR retrievals during SGP97.

pling of rainfall was constructed by applying a direct subsampling approach to 15-min rain gauge data collected at all Oklahoma Mesonet stations during the calendar years 1997 and 1999. For a daily sampling frequency of ν, rainfall rates were assumed constant and equal to the observed 15-min gauge-derived rate for the $24\nu^{-1}$ h period centered on each observation. Rainfall rate estimates for the each of the $24\nu^{-1}$ h periods in a single day were then averaged and used to construct a daily rainfall accumulation estimate $\hat{R}$. Each daily rainfall estimate $\hat{R}$ was paired with the corresponding true daily accumulation rainfall value R derived from summing all observations within a given day. This pairing was repeated for every day at every Oklahoma Mesonet station during 1997 and 1999. The resulting set of ($\hat{R}$, R) pairs were then discretely binned according to both $\hat{R}$ and R and used to construct a series of conditional likelihood distributions $f_\nu(R|\hat{R})$ for various discrete ranges of $\hat{R}$. This set of distributions included the case of zero measured rainfall $f_\nu(R|\hat{R} = 0)$ describing the probability of falsely negative rainfall observations.

The conditioning of accumulation likelihoods depends on the frequency of sampling that supports the accumulation estimate. Sets of conditional distributions were constructed for rainfall sampling rates (ν) of 2, 4, 6, 8, and 12 day^{-1}. The special case of unconditioned rainfall intensities $f_0(R)$ was also constructed to represent daily rainfall likelihood in the absence of any rainfall measurements (i.e., $\nu = 0$). Figure 4 shows a set of $f_\nu(R|\hat{R})$ histograms for $\nu = 8$ day^{-1}, constructed using seasonally and geographically pooled observations within the entire state of Oklahoma. No attempt was made to parameterize or model the likelihood distributions. Instead, individual rainfall realizations were

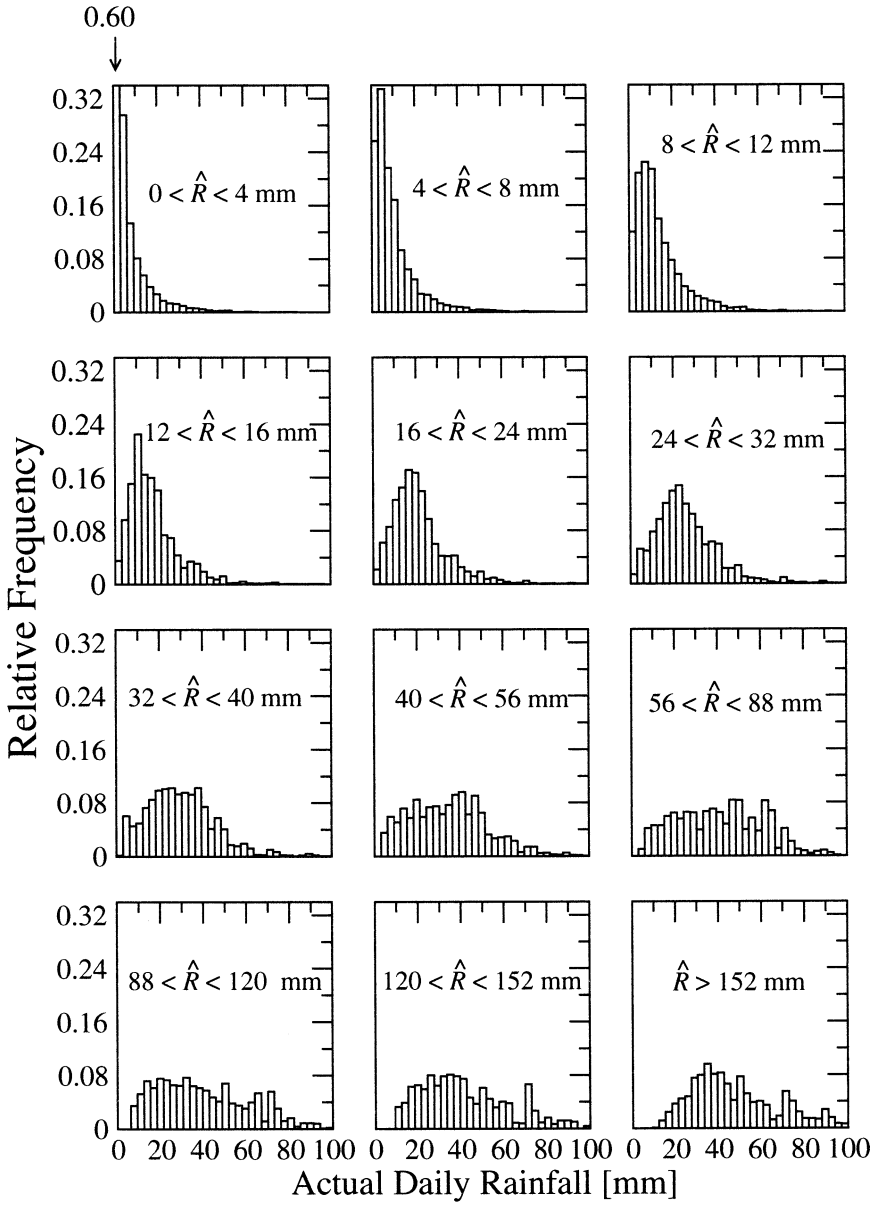

FIG. 4. Set of normalized conditional rainfall likelihood distributions for actual daily rainfall accumulations given a rainfall estimate $\hat{R}$ derived from the subsampling of eight observations a day. Results are derived from 15-min OK mesonet rain gauge measurements collected during calendar years 1997 and 1999.

randomly sampled from the exact set of observations used to construct the likelihood distributions.

3. Methodology

Two separate methodologies were utilized to evaluate the data assimilation system. The first approach was based on an identical twin data assimilation experiment design. In this approach, TOPLATS/LSMEM simulations forced by complete (i.e., not subsampled) local rain gauge observations were designated as truth. These simulations, shown in Fig. 1 and validated in Figs. 2 and 3, will be referred to as ''benchmark'' simulations. LSMEM T_B predictions from benchmark simulations were perturbed with random error ($C_v = 9$ K^2) to form a set of daily synthetic observations $\mathbf{Z}_k$ and reassimilated back into TOPLATS via the EnKF. The second methodology utilized real ESTAR observations of T_B acquired during SGP97. During the experiment (17 June 1997 to 16 July 1997), ESTAR imaging was performed on 18, 19, 20, 25, 26, 27, 29, and 30 June, and on 1, 2, 3, 11, 12, 13, 14, and 16 July.

For the 24-h period preceding 1600 UTC (the approximate measurement time for ESTAR measurements

during SGP97), daily estimated rainfall amounts ($\hat{R}$) were derived at each site through the subsampling of ν 15-min rain gauge observations per day. This subsampled estimate was then used to select the proper $f_\nu(R|\hat{R})$ histogram (Fig. 4) from which to sample a set of daily rainfall amounts. Sampled daily rainfall amounts were distributed equally among $24c\nu^{-1}$ consecutive hours of rainfall randomly located within the 24-h period, where c is the number of 15-min rainfall gauge measurements where rainfall was observed. This ensemble of hourly rainfall time series was used to generate an ensemble of 25 TOPLATS state ($\mathbf{Y}$) and LSMEM observation [$\mathbf{M}(\mathbf{Y})$] forecasts from which forecast estimates of model error ($\mathbf{C}_{YM}$ and $\mathbf{C}_M$) could be obtained. Model ensembles were created for every 24-h period regardless of whether T_B observations were available for assimilation. At times with observations, each ensemble member was updated using the T_B observations ($\mathbf{Z}_k$) and the EnKF update equation given in (3). Four TOPLATS soil moisture states (0–5, 5–15, 15–40, and 40 cm to water table) and one soil temperature state (7.5 cm) were updated in this way. These simulations will be referred to as ''EnKF'' or ''EnKF-based assimilation'' results. Likewise, simulations where rainfall ensembles were generated but not updated with T_B observations will be referred to as ''open loop'' simulations.

Additional TOPLATS simulations were constructed that relied directly on subsampled rainfall data to provide hourly precipitation forcing. Like open loop results, these simulations did not assimilate T_B observations. However, unlike open loop results, they did not employ the rainfall ensemble generation procedure described in section 2d. Instead, for a given daily sampling frequency of ν, hourly rainfall rates were assumed constant and equal to the observed 15-min gauge-derived rate for the $24\nu^{-1}$ h period centered on each observation. TOPLATS output for these simulations will be referred to as ''subsampled precipitation'' results. The accuracy of TOPLATS and LSMEM predictions derived from the assimilation of T_B observations via the EnKF will be evaluated based on their ability to improve surface state and flux predictions relative to subsampled precipitation simulations that utilize neither T_B observations nor the EnKF methodology.

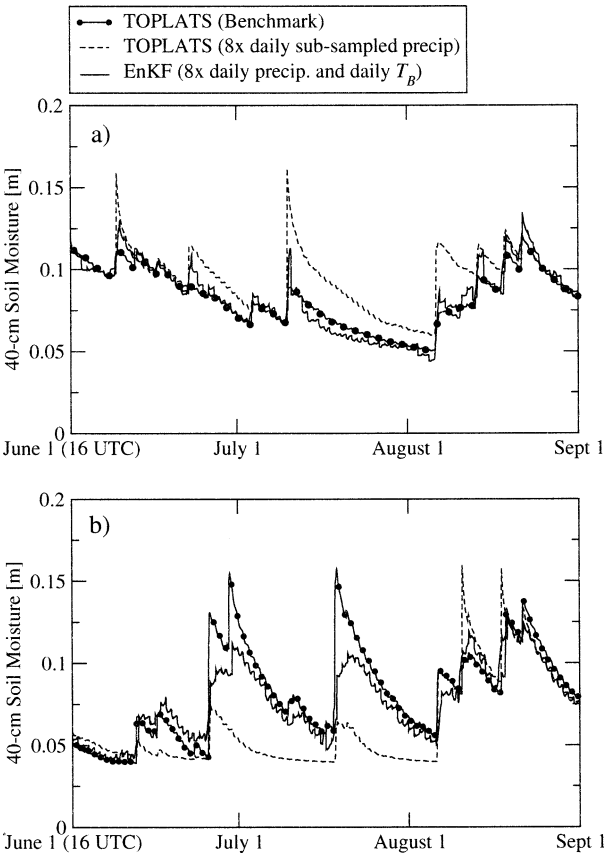

FIG. 5. (a) At the NOAA ATDD Little Washita site, time series of 40-cm soil moisture results for: the benchmark TOPLATS simulation, TOPLATS simulations forced by eight-times-daily subsampling of precipitation, and EnKF results (eight-times-daily precipitation and daily T_B observations) during the 1997 growing season. (b) Same as (a), but for the ARM CART EF13 site.

4. Results: Identical twin experiment

Figure 5 shows root-zone soil moisture results derived from the EnKF-based assimilation of L-band surface brightness temperature (T_B) into TOPLATS simulations forced using sparse observations of rainfall. The T_B measurements assimilated to produce EnKF results in Fig. 5 were synthetically generated every day at 1600 UTC (1000 CST) using the identical twin experiment methodology described in section 3. Rainfall observations were taken from the subsampling of eight 15-min rain gauge observations per day. This series of eight daily measurements were spaced evenly throughout the day with the first measurement occurring at 0000 UTC. Also shown are TOPLATS results derived from the direct use of subsampled precipitation and benchmark TOPLATS predictions based on all available rainfall data at both sites. At the NOAA ATDD Little Washita site, the assimilation of T_B observations allows for a more accurate representation of a midsummer dry down by compensating for the consistently high bias of precipitation estimates derived from sparse subsampling of rainfall events. Conversely, at the ARM CART EF13 site, large precipitation events in late June and early July are underestimated by temporally sparse rainfall sampling, leading to anomalously dry conditions during the middle portion of the growing season. By interpreting surface L-band T_B observations and accurately replenishing root-zone soil water at deeper levels, the EnKF-based assimilation of T_B is able to compensate for error in rainfall forcing data. It should be noted that the root-zone depth assumed in Fig. 5 (40 cm) is considerably deeper than the shallow vertical depth of sensitivity for T_B observations (5 cm). Consequently, results do not

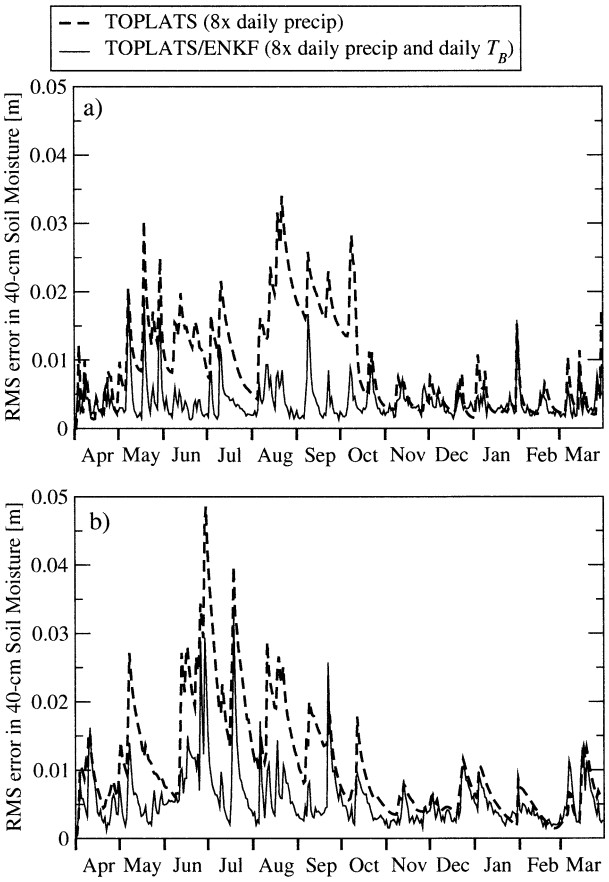

FIG. 6. (a) At the NOAA ATDD Little Washita site, time series of rms error in EnKF 40-cm soil moisture predictions. Errors are pooled values based on all possible start times for precipitation observations. Also shown are rms errors associated with the direct forcing of TOPLATS using subsampled precipitation data. (b) Same as (a), but for the ARM CART EF13 site.

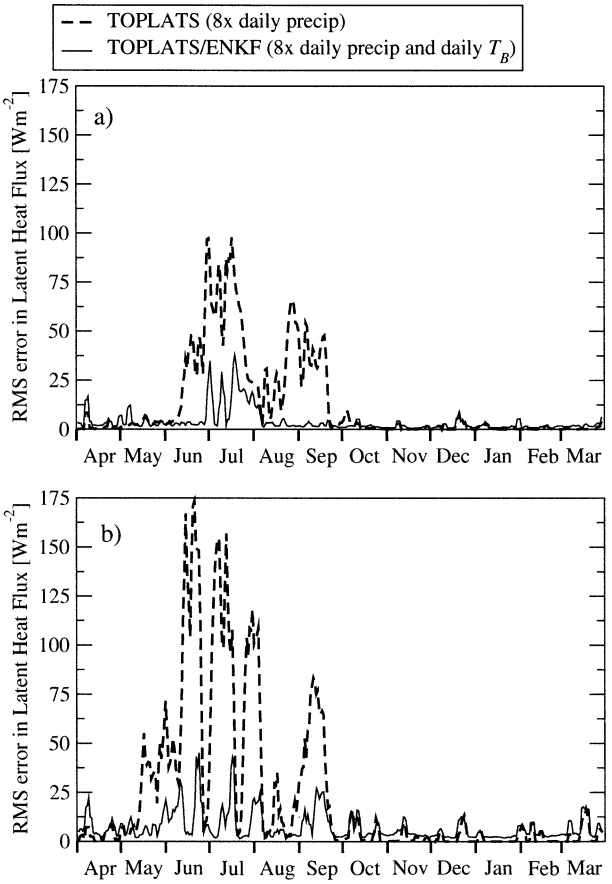

FIG. 7. (a) At the NOAA ATDD Little Washita site, time series of rms error in daily averaged (1000 to 1600 local time) EnKF latent heat flux predictions. Errors are pooled values based on all possible start times for precipitation observations. Also shown are rms errors associated with the direct forcing of TOPLAS using subsampled precipitation data. To improve readability, all errors are temporally smoothed within a 3-day moving average window. (b) Same as (a), but for the ARM CART EF13 site.

simply reflect the direct replacement of model-generated soil moistures with remote observations.

Results in Fig. 5 are based on the designation of just one of the 12 possible 15-min intervals between 0000 and 0300 UTC as the start time for the cyclic sequence of eight rainfall observations per day. Instead of plotting results for all 12 possible start times, Figs. 6 and 7 summarize them by plotting pooled root-mean-square (rms) differences between the benchmark TOPLATS simulation and EnKF root-zone soil moisture results. Also plotted are pooled rms errors associated with the direct use of eight-times-daily subsampled precipitation to force TOPLATS. EnKF root-zone soil moisture results in Fig. 6 show substantial improvements versus subsampled precipitation results during the growing season. Spikes in rms error, coinciding with large precipitation events, suggest that the EnKF does not immediately incorporate the impact of intense precipitation; however, the assimilation of T_B observations in the aftermath of rainfall events allows for a sharp correction during the early portion of dry-down events. Relative to the growing season, less improvement is observed during fall and winter months. Raising the observational error used to create synthetic T_B observations from 3 to 6 K increases rms error in EnKF predictions by about 10% at both sites (not shown).

Analogous latent heat flux results are shown in Fig. 7. Large errors in subsampled precipitation results are due to the inability of sparse rainfall rate sampling to accurately capture alternating periods of water- and energy-limited evapotranspiration during the growing season. For instance, subsampled precipitation results for the NOAA ATDD Little Washita site in Fig. 7a show large errors during a dry period in July (see Fig. 5). Much of this error is due to instances where the subsampling of modest precipitation events in late June and early July overestimates their intensity and underestimates the severity of the subsequent water-limited evapotranspiration period. As seen in Fig. 5a, the EnKF-based assimilation of T_B observations removes water

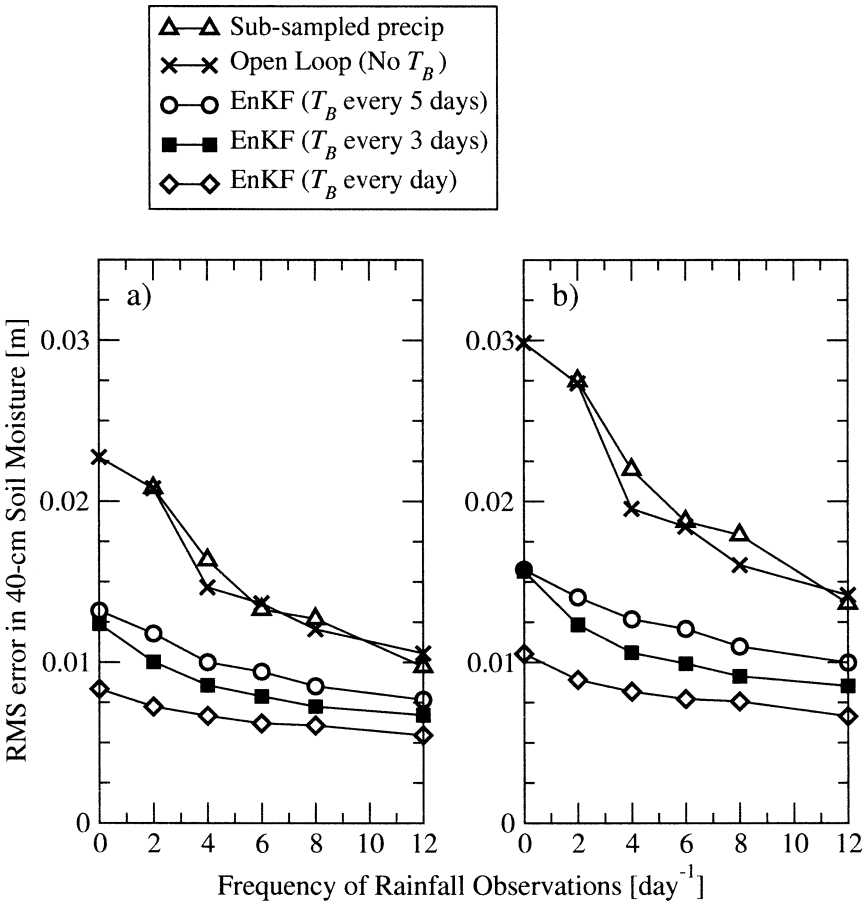

FIG. 8. (a) Pooled rms error for EnKF 40-cm soil moisture results over entire simulation period (1 Apr 1997 to 31 Mar 1998) for both study sites at a range of rainfall and T_B sampling frequencies. Errors are calculated relative to benchmark TOPLATS simulations. (b) Same as (a), but only for the growing season (1 May 1997 to 31 Sep 1997).

from the root-zone and compensates for the overestimation of antecedent rainfall amounts. This leads to an improved representation of soil water limitation on evapotranspiration at both sites. The absence of evapotranspiration errors from October to May is due to the near uniformity of energy-controlled evapotranspiration conditions at both sites during the winter and spring.

a. Impact of observation frequency

Figure 8a plots 40-cm soil moisture rms error results for a range of rainfall and T_B sampling frequencies during the entire simulation period (from 1 April 1997 to 31 March 31 1998). Error statistics were derived from pooling soil moisture results at both study sites and using all possible start times for daily rainfall observations. Figure 8b is identical, except the values are derived from pooling results during the growing season only (from 1 May 1997 to 31 September 1997). The lowest possible T_B observation frequency (i.e., never) corresponds to an open loop case where individual rainfall realizations are generated from rainfall likelihood distributions but not updated via (3). Comparison between open loop and subsampled precipitation cases provides an important check that the rainfall ensemble procedure described in section 2d is generating errors consistent with actual uncertainties associated with the direct use of subsampled rainfall.

In contrast to instantaneous measurements of precipitation fluxes, T_B observations measure a surface state—soil moisture—containing integrated information about past rainfall events. Figures 8 and 9 demonstrate the potential value of this memory for filtering errors associated with sparsely subsampled and memoryless precipitation flux retrievals. For instance, during the growing season (Fig. 8b) the assimilation of daily surface T_B measurements reduces rms error in 40-cm soil moisture predictions by about 50% for 12 rainfall samples per day and up to 65% for twice-daily sampling. Relative to open loop results, the EnKF-based assimilation of T_B also reduces the sensitivity of root-zone soil moisture errors to the frequency of rainfall observations. This loss of sensitivity is evident for T_B measurement rates as infrequent as once every 5 days. One consequence

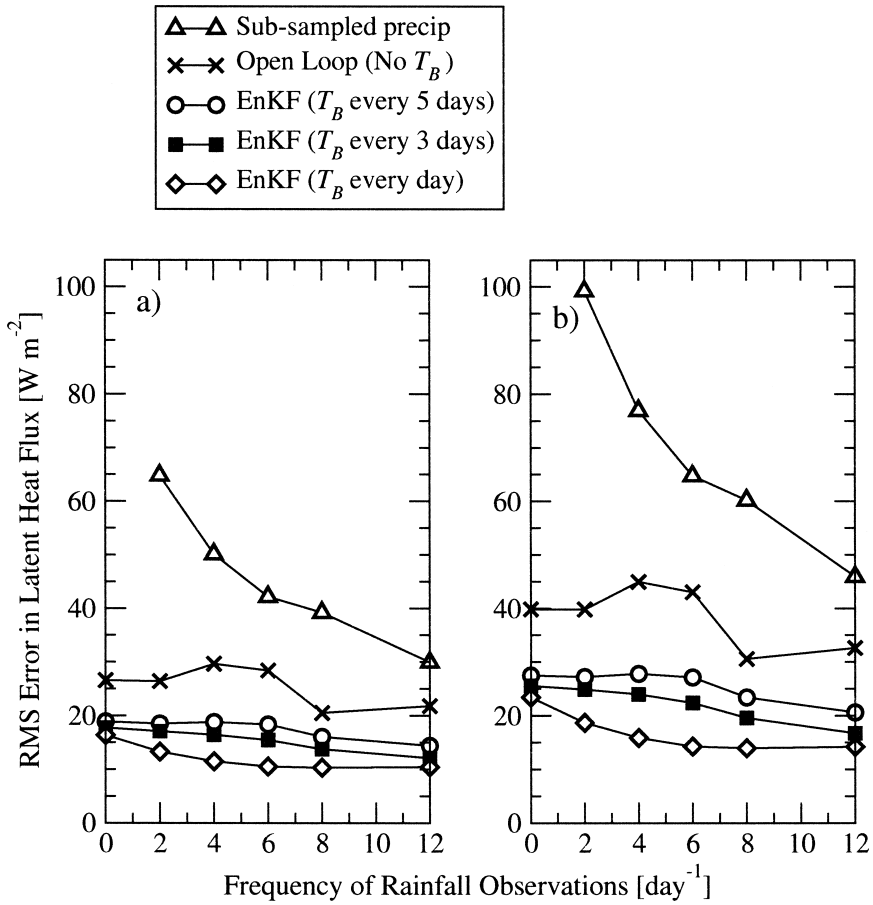

FIG. 9. (a) Pooled rms error for daily averaged (1000 to 1600 local time) EnKF latent heat flux results over entire simulation period (1 Apr 1997 to 31 Mar 1998) for both study sites at a range of rainfall and T_B sampling frequencies. Errors are calculated relative to benchmark TOPLATS simulations. (b) Same as (a), but only for the growing season (1 May 1997 to 31 Sep 1997).

of this insensitivity is the relative superiority of a combination of sparse L-band T_B observations and rainfall sampling rates versus much more frequent rainfall observations in isolation. For instance, during the growing season twice-daily rainfall sampling and one T_B observation every five days leads to errors that are comparable to 12 rainfall samples per day and no T_B observations. Likewise, results for twice-daily retrievals of rainfall rate and daily T_B observations are as accurate as those for any combination of more frequent rainfall observations (up to 12 samples per day) and sparser T_B retrieval rates.

Figure 9 is analogous to Fig. 8, except that it demonstrates latent heat flux errors. One clear contrast between Figs. 8 and 9 is the degree to which open loop and subsampled precipitation results differ. For latent heat flux, open loop TOPLATS results, where estimated $\hat{R}$ values are used to sample an ensemble of precipitation intensities from the appropriate $f(R|\hat{R})$ likelihood distribution, are more accurate than subsampled precipitation results based on the direct use of $\hat{R}$. Sparse subsampling of precipitation tends to enhance precipitation intermittency and strongly overpredict both the length and severity of soil-controlled evapotranspiration periods. In contrast, using $\hat{R}$ to generate a likelihood-based ensemble of precipitation events temporally smoothes precipitation and moderates the severity of soil water content extremes. This keeps evapotranspiration rates for open loop results at or near potential evapotranspiration levels. Because the 1997 growing season is relatively wet, a blanket assumption of potential evapotranspiration is acceptable for many periods of time and open loop latent heat flux predictions do not exhibit large errors. Nevertheless, daily assimilation of T_B measurements allows for improved representation of soil-controlled periods and reduces rms errors in TOPLATS latent heat flux predictions by up to 50% relative to the open loop case.

b. Filter performance

EnKF results can also be evaluated based on their consistency with regards to the assumptions that underlie the optimality of the Kalman filter update equa-

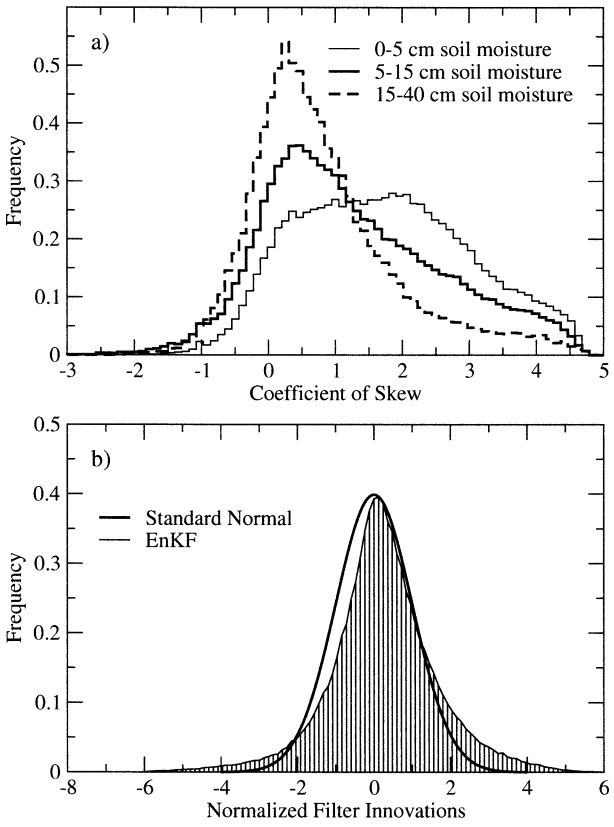

FIG. 10. (a) Distribution of coefficient of skew results for soil moisture ensemble forecasts at a range of depths. (b) Distribution of normalized filter innovations and a reference standard normal distribution.

tion. For instance, previous work has noted that uncertainty in rainfall does not necessarily lead to Gaussian error forecasts for T_B and surface soil moisture (Crow and Wood 2003). Figure 10a plots histograms for coefficient of skew results calculated from pooled model forecasts of soil moisture at both study sites. Surface (0–5 cm) soil moisture results are generally associated with large positive skew. Even when no rainfall is detected (i.e., $\hat{R} = 0$) some positive rainfall realizations are required for the rainfall ensemble to properly account for falsely negative precipitation observations. Therefore, in contrast to the majority of ensemble realizations capturing a drying tendency, isolated ensemble members with positive rainfall forcing will become sharply wetter. This leads to positively skewed forecast ensembles for surface soil moisture and potentially nonoptimal updating using (3). One influence on forecast skew is the frequency of rainfall observations. Increasing the frequency of rainfall rate observations from 2 to 12 day^{-1} reduces the average coefficient of skew in 5-cm forecasts, but only slightly (from 1.87 to 1.58). A much stronger factor is the depth of the forecasted soil moisture state. As demonstrated in Fig. 9a, forecast skew is progressively reduced within deeper soil layers.

The forecast skew observed in Fig. 10a can be reduced through ancillary measurements, which eliminate the possibility of falsely negative rainfall observations. For instance, in cases where sparse rainfall radiometer subsampling detects no rainfall (i.e., $\hat{R} = 0$), visible (VIS) and thermal infrared (IR) precipitation indices from geostationary satellites could be used to make a rain/no-rain determination for the entire day. During days with no rain, a reliable no-rain determination eliminates the need to generate the small number of positive rainfall realizations that tend to skew the entire forecast ensemble. If the VIS–IR observation system detects rain and the sparse radiometer sampling does not, rainfall rates can be sampled from a conditioned likelihood distribution reflecting positive rainfall, which is undetected by sparse subsampling at a rate of ν day^{-1} {i.e., $f_\nu[R|(R > 0$ and $\hat{R} = 0)]$}. Incorporating this modification, EnKF results for eight rainfall observations per day and daily T_B measurements were repeated at the NOAA ATDD site using a modified assimilation system that assumed perfect rain/no-rain determinations were available. Eliminating the possibility of falsely negative rainfall observations reduces the coefficient of skew in surface (0–5 cm) soil moisture forecasts by 71% (1.71–0.50), yet improves model EnKF predictions of root-zone soil moisture by only 12% (0.0047–0.0041 m). Such a modest improvement in accuracy implies that skew in model forecasts of T_B and surface soil moisture is not a major source of overall error in results for the original filter. One possible explanation for this is the confinement of large skew to state forecasts within surface layers that constitute a minor fraction of overall root-zone volume. Deeper, more substantial soil moisture states exhibit considerably less skew and are, therefore, more amenable to updating using the EnKF. This result is consistent with earlier work by Reichle et al. (2002b) who demonstrate the adequacy of the EnKF in updating skewed forecast ensembles arising from nonlinearites in land surface models.

The statistical properties of the filter innovations provide an additional diagnostic tool for assessing filter performance. A filter's innovations are defined as the sequence of differences between forecasted and actual observations $\mathbf{Z}_k - \mathbf{M}_k(\mathbf{Y}^i_-)$. If the underlying assumptions of the Kalman filter are fully met (i.e., linear models and uncorrelated Gaussian errors), observed innovations should be temporally uncorrelated, mean zero, and Gaussian with a covariance equal to $\mathbf{C}_M + \mathbf{C}_v$. The degree to which actual innovations depart from this ideal gives a sense as to the applicability of the Kalman filter update equation to a particular filtering problem. Figure 10b plots a histogram of normalized innovations—$[\mathbf{Z}_k - \mathbf{M}_k(\mathbf{Y}^i_-)]/(\mathbf{C}_M + \mathbf{C}_v)$—observed at both study sites during assimilation of daily T_B observations and eight-times-daily subsampling of rainfall. A standard normal distribution is also plotted for reference. The positive shift of the normalized innovation distribution relative to the standard normal reflects the continual readjustment of excessively wet (i.e., low T_B) ensemble mem-

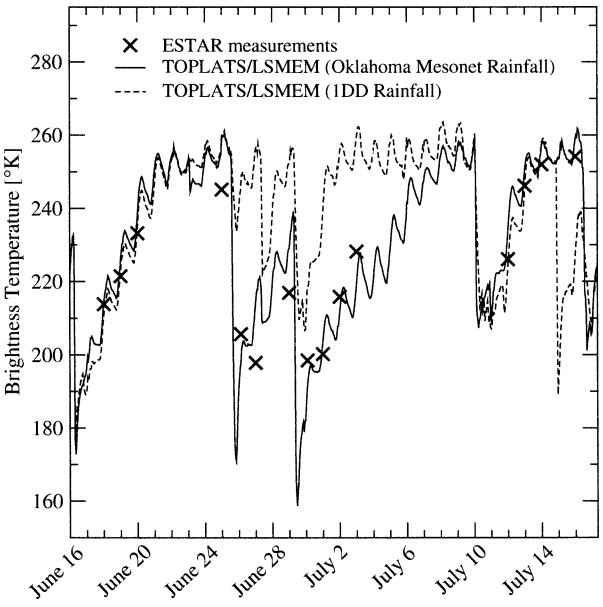

FIG. 11. Time series of TOPLATS LSMEM T_B predictions derived from both OK Mesonet rainfall gauge measurements and precipitation retrievals extracted from the GPCP 1DD dataset within the 1° lat × 1° lon grid cell surrounding the ARM CART EF13 site (36° to 37°N and −97° to −98°W). Also plotted are spatially averaged ESTAR T_B observations within the same degree box.

bers produced to capture the possibility of falsely positive rainfall observations. Frequent low T_B forecasts relative to actual measurements lead, in turn, to an excessive amount of positive innovations. Likewise, the non-Gaussian tail at large negative normalized innovations is due to the inability of the climatologically based ensembling procedure to capture large precipitation events during the modeling period. Because rainfall ensembles are based on climatological likelihoods, large events are not well represented. In a Bayesian sense, this is because unusually intense rainfall observations are more likely associated with large sampling errors than an actual rare event. Consequently, for large $\hat{R}$, few, if any, rainfall realizations generated from the $f(R|\hat{R})$ likelihood distribution will be greater than $\hat{R}$ (see Fig. 4). However, several large precipitation observations at the ARM CART EF13 site are, in fact, accurate representations of unusually intense precipitation events. The inability of the rainfall ensemble generation procedure to reflect such extreme events leads to the strong overprediction of T_B and the non-Gaussian negative tail on the innovation histogram shown in Fig. 9.

5. Results: Real observations

Datasets collected during SGP97 provide an opportunity to evaluate the proposed data assimilation system using real remote sensing observations of T_B. Figure 11 shows T_B predictions for a TOPLATS LSMEM simulation of the ARM CART EF13 site forced by both rain gauge observations and rainfall data extracted from the satellite-derived GPCP 1DD dataset. Also plotted are L-band T_B observations derived from airborne ESTAR measurements during SGP97. To ensure compatibility with the 1° resolution of the 1DD rainfall observations, ESTAR observations, LAI forcing for TOPLATS, and gauge-based precipitation observations were averaged within the 1° lat × 1° lon grid box roughly centered on the ARM CART EF13 site. Due to apparent errors in the 1DD rainfall, the two simulations diverge on 25 June, 29 June, and 15 July. While TOPLATS LSMEM T_B predictions derived from local rain gauge observations are quite accurate, results with spaceborne 1DD rainfall observations are clearly not consistent with independent ESTAR T_B observations. It is this lack of consistency that forms the observational basis for the detection, and eventual correction, of land surface model errors associated with poor rainfall forcing.

While Fig. 11 demonstrates the basic detectability assumption at the heart of the data assimilation system, it does not validate the system itself. As a first step toward this, Table 1 presents results for the assimilation of actual ESTAR T_B measurements during the SGP97 period (17 June to 16 July 1997). Results are shown for rainfall forcing consistent with both twice and eight-times-daily subsampling of rainfall rates. ''ESTAR'' results are based on updating with real L-band T_B obser-

TABLE 1. Comparison between rms errors in EnKF 40-cm soil moisture ($\theta_{40\,\mathrm{cm}}$) and latent heat flux (λE) results utilizing both synthetic and actual ESTAR T_B observations. Given rms errors are calculated by pooling results between 17 Jun and 16 Jul 1997. Results are shown for rainfall rate subsampling frequencies (ν) of both 2 and 8 day^{-1}.

Site	ν (day^{-1})	T_B source	$\theta_{40\,\mathrm{cm}}$ (m)	λE (W m^{-2})
NOAA ATDD Little Washita	2	Synthetic	0.0079	26.5
	2	ESTAR	0.0114	63.2
	2	ESTAR (bias corrected)	0.0074	21.7
	8	Synthetic	0.0061	21.4
	8	ESTAR	0.0107	58.6
	8	ESTAR (bias corrected)	0.0061	19.6
ARM CART EF13	2	Synthetic	0.0167	24.4
	2	ESTAR	0.0188	29.7
	8	Synthetic	0.0138	27.3
	8	ESTAR	0.0154	28.0

vations at each of the 16 ESTAR observations times (see Fig. 11). Results for ''synthetic T_B'' observations are based on updating at the same 16 times, but with T_B values synthetically generated by the identical twin experiment. As in the identical twin experiment, errors for both are based on comparisons to the benchmark TOPLATS simulation described in section 3. Consequently, comparisons in Table 1 give a sense as to the feasibility of obtaining the accuracy improvements seen in Figs. 6, 7, 8, and 9 with real T_B observations. At the ARM CART EF13 site, real T_B observations are nearly as effective as synthetically generated observations. However, the assimilation of T_B observations at the NOAA ATDD Little Washita site does not lead to comparable accuracies until after the T_B bias observed in Fig. 3 is removed. In a truly operational setting, such an adjustment would require the successful application of a bias detection and correction strategy (Dee and DaSilva 1998).

6. Summary and conclusions

Random sampling errors due to temporally sparse observation frequencies (2–12 per day) are expected to comprise a major fraction of retrieval error for current and next generation radiometer systems designed to estimate daily rainfall accumulations from space (e.g., Bell et al. 1990, or Steiner 1996). The presence of random errors in precipitation forcing data for land surface models can lead to large uncertainty in predictions of root-zone (40 cm) soil moisture and surface energy fluxes (Figs. 5 and 6). However, surface T_B observations appear to offer a viable observational basis to detect land surface model errors associated with poorly retrieved rain rates (Fig. 11). This analysis exploits this potential by developing an EnKF-based data assimilation system designed to update land surface model predictions, made uncertain by the impact of random sampling error in rainfall observations, with surface T_B observations supported only by near-surface (0–5 cm) microwave emission.

The EnKF-based land surface data assimilation system presented in section 2 is based on the use of temporally sparse rainfall observations—ostensibly from a spaceborne source—to condition expectations concerning daily rainfall accumulations. Using a Monte Carlo approach, rainfall realizations sampled from histograms of conditional rainfall likelihoods are run through a land surface and forward microwave emission model to create an ensemble of land surface state and surface microwave brightness temperature (T_B) predictions. This ensemble provides the necessary error covariance information to update individual members of the ensemble with T_B observations via Eq. (3). A synthetic identical twin experiment demonstrates that the approach is capable of substantially reducing errors in both root-zone soil moisture and surface energy flux predictions relative to results derived from the direct use of sparsely subsampled rainfall to force TOPLATS (Figs. 6, 7, 8, and 9). Preliminary results with real T_B predictions obtained during SGP97 suggest that comparable levels of correction are obtainable using actual L-band T_B observations (Table 1), although in some cases, reaching this potential may require the application of an effective bias detection strategy.

Results in Figs. 8 and 9 can also be used to gauge the relative value of spaceborne brightness temperature and rainfall rate observations for land surface modeling. A generalization emerging from the figures is the degree to which a combination of relatively sparse rainfall rate and T_B measurements leads to more accurate state and flux predictions than open loop (i.e., no T_B assimilation) results driven by more frequent rainfall observations. Given twice-daily observations of rainfall, L-band spaceborne retrievals of T_B every 3 days are more valuable for root-zone soil moisture and surface energy flux prediction than an increase in the observational frequency of rainfall rate retrievals from 2 to 8 day^{-1}. There are also indications that adequate T_B sampling may preclude the need to improve the temporal sampling characteristics of spaceborne rainfall systems. Updating with daily observations T_B and the EnKF substantially reduces the observed sensitivity of results to the frequency of rainfall observations.

A central issue in the design of any data assimilation system is the choice of a particular assimilation technique. As a noted in Reichle et al. (2002a) the EnKF has a number of attributes that make it well suited for the assimilation of T_B observations into a land surface model. Of particular interest here is the EnKF's flexibility with regard to model error type and ability to temporally propagate uncertainty associated with poor rainfall forcing. This flexibility distinguishes the EnKF from other variants of the Kalman filter—most notably the extended Kalman filter (Reichle et al. 2002b). One potential downside for the EnKF is the presence of skew in ensemble model forecasts (Fig. 10), which will prevent optimal updating of ensembles via the Kalman filter. Because alternative assimilation strategies, particularly the fully nonlinear Monte Carlo Bayesian filter presented by Anderson and Anderson (1999) and the variational smoother incorporating model error presented by Reichle et al. (2001), are better suited for dealing with non-Gaussian errors than the EnKF, measuring the impact of this forecast skew on eventual EnKF update accuracy is critical for assessing the relative suitability of the EnKF versus other assimilation techniques. Section 4b demonstrates that assuming the availability of accurate rain/no-rain determinations eliminates a majority of the forecast skew. Crucially, this sharp reduction is not associated with substantial improvements in assimilation results. This implies that forecast skew was not a major source of error in the original EnKF-updated state predictions, which, in turn, supports the EnKf's overall suitability for this application.

Several caveats should be noted concerning overall results presented here. First, the U.S. SGP region is widely regarded as nearly optimal for soil moisture remote sensing. The value of L-band T_B observations for the inference of surface state conditions will almost certainly be reduced for more densely vegetated regions. Finally, a more thorough analysis would include additional causes of uncertainty in spaceborne rainfall rate retrievals (e.g., ambiguities surrounding the selection of an appropriate rain rate–reflectivity relationship and beam-filling problems for spaceborne radiometers) as well as a realistic spatial resolution for spaceborne T_B observations.

Acknowledgments. Data support from the DOE's Atmosphere Radiation Measurement Program, Tilden Meyers (NOAA ATDD), Thomas Jackson (USDA ARS), the Oklahoma Mesonet, and the Global Precipitation Climatology Project is gratefully acknowledged.

REFERENCES

Adams, W. J., and Coauthors, 2002: Global precipitation measurement—Report 8 white paper. NASA Tech. Rep. NASA/TM-2002-211609, 33 pp.

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.

Bell, T. L., A. Abdullah, R. L. Martin, and G. R. North, 1990: Sampling errors for satellite-derived tropical rainfall: Monte Carlo study using a space–time stochastic model. *J. Geophys. Res.,* **95,** 2195–2205.

Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.,* **126,** 1719–1724.

Crow, W. T., and E. F. Wood, 2002: The value of coarse-scale soil moisture observations for regional surface energy balance modeling. *J. Hydrometeor.,* **3,** 467–482.

——, and ——, 2003: The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using ensemble Kalman filtering: A case study based on ESTAR measurements during SGP97. *Adv. Water Resour.,* **26,** 137–149.

——, M. Drusch, and E. F. Wood, 2001: An observation system simulation experiment for the impact of land surface heterogeneity on AMSR-E soil moisture retrieval. *IEEE Trans. Geosci. Remote Sens.,* **39,** 1622–1631.

Dee, D., and A. M. DaSilva, 1998: Data assimilation in the presence of forecast bias. *Quart. J. Roy. Meteor. Soc.,* **124,** 269–295.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.,* **99,** 10 143–10 162.

Famiglietti, J. F., and E. F. Wood, 1994: Multiscale modeling of spatially variable water and energy balance processes. *Water Resour. Res.,* **30,** 3061–3078.

Flaming, G. M., W. J. Adams, S. P. Neeck, and E. A. Smith, 2001: Planning for global precipitation measurement. *Proc. 2002 Int. Geoscience and Remote Sensing Symp.,* Sydney, Australia, IEEE, 685–687.

Huffman, G. J., R. F. Adler, M. M. Morrissey, D. T. Bolvin, S. Curtis, R. Joyce, B. McGavock, and J. Susskind, 2001: Global precipitation at one-degree daily resolution from multisatellite observations. *J. Hydrometeor.,* **2,** 36–50.

Jackson, T. J., D. M. LeVine, A. V. Hsu, A. Oldak, P. J. Starks, C. T. Swift, J. D. Isham, and M. Haken, 1999: Soil moisture mapping at regional scales using microwave radiometry: The Southern Great Plains Hydrology Experiment. *IEEE Trans. Geosci. Remote Sens.,* **37,** 2136–2150.

Margulis, S. A., D. McLaughlin, D. Entekhabi, and S. Dunne, 2002: Land data assimilation of soil moisture using measurements from the Southern Great Plains 1997 Field Experiment. *Water Resour. Res.,* **38,** 1299, doi:10.1029/2001WR001114.

New, M., M. Todd, M. Hulme, and P. Jones, 2001: Precipitation measurements and trends in the twentieth century. *Int. J. Climatol.,* **21,** 1899–1922.

Peters-Lidard, C. D., M. S. Zion, and E. F. Wood, 1997: A soil–vegetation–atmosphere transfer scheme for modeling spatially variable water and energy balance processes. *J. Geophys. Res.,* **102,** 4303–4324.

——, F. Pan, and E. F. Wood, 2001: A re-examination of modeled and measured soil moisture spatial variability and its implications for land surface modeling. *Adv. Water Resour.,* **24,** 1069–1083.

Reichle, R. H., D. B. McLaughlin, and D. Entekhabi, 2001: Variational data assimilation of microwave radiobrightness observations for land surface hydrology applications. *IEEE Trans. Geosci. Remote Sens.,* **39,** 1708–1719.

——, ——, and ——, 2002a: Hydrologic data assimilation with the ensemble Kalman filter. *Mon. Wea. Rev.,* **130,** 103–114.

——, J. P. Walker, R. D. Koster, and P. R. Houser, 2002b: Extended versus ensemble Kalman filtering for land data assimilation. *J. Hydrometeor.,* **3,** 728–740.

Steiner, M., 1996: Uncertainty of estimates of monthly area rainfall for temporally sparse remote observations. *Water Resour. Res.,* **32,** 373–388.

Todd, M. C., C. Kidd, D. Kniveton, and T. J. Bellerby, 2001: A combined satellite infrared and passive microwave technique for estimation of small-scale rainfall. *J. Atmos. Oceanic Technol.,* **18,** 742–755.