Assimilation of Satellite-Derived Skin Temperature Observations into Land Surface Models

ROLF H. REICHLE

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

SUJAY V. KUMAR

Science Applications International Corporation, Beltsville, and Hydrological Sciences Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland

SARITH P. P. MAHANAMA

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, and Goddard Earth Sciences and Technology Center, University of Maryland, Baltimore County, Baltimore, Maryland

RANDAL D. KOSTER

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

Q. LIU

Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, and Science Applications International Corporation, Beltsville, Maryland

(Manuscript received 29 December 2009, in final form 13 April 2010)

ABSTRACT

Land surface (or ''skin'') temperature (LST) lies at the heart of the surface energy balance and is a key variable in weather and climate models. In this research LST retrievals from the International Satellite Cloud Climatology Project (ISCCP) are assimilated into the Noah land surface model and Catchment land surface model (CLSM) using an ensemble-based, offline land data assimilation system. LST is described very differently in the two models. A priori scaling and dynamic bias estimation approaches are applied because satellite and model LSTs typically exhibit different mean values and variabilities. Performance is measured against 27 months of in situ measurements from the Coordinated Energy and Water Cycle Observations Project at 48 stations. LST estimates from Noah and CLSM without data assimilation (''open loop'') are comparable to each other and superior to ISCCP retrievals. For LST, the RMSE values are 4.9 K (CLSM), 5.5 K (Noah), and 7.6 K (ISCCP), and the anomaly correlation coefficients (R) are 0.61 (CLSM), 0.63 (Noah), and 0.52 (ISCCP). Assimilation of ISCCP retrievals provides modest yet statistically significant improvements (over an open loop, as indicated by nonoverlapping 95% confidence intervals) of up to 0.7 K in RMSE and 0.05 in the anomaly R. The skill of the latent and sensible heat flux estimates from the assimilation integrations is essentially identical to the corresponding open loop skill. Noah assimilation estimates of ground heat flux, however, can be significantly worse than open loop estimates. Provided the assimilation system is properly adapted to each land model, the benefits from the assimilation of LST retrievals are comparable for both models.

1. Introduction

Land surface conditions are intimately connected with the global climate system and have been associated, through different pathways, with atmospheric

Corresponding author address: Rolf H. Reichle, NASA Goddard Space Flight Center, Mail Code 610.1, 8800 Greenbelt Rd., Greenbelt, MD 20771.
E-mail: rolf.reichle@nasa.gov

DOI: 10.1175/2010JHM1262.1

© 2010 American Meteorological Society

predictability. Land surface temperature (LST) lies at the heart of the surface energy balance and is therefore a key variable in weather and climate models (see section 2 for more on the specific definition of LST used here). LST influences the latent and sensible heat fluxes to the atmosphere through which it affects the planetary boundary layer and atmospheric convection. LST also plays an important role in the assimilation of atmospheric remote sensing observations. Because forward radiative transfer modeling for surface-sensitive (window) channels requires accurate information about land surface conditions, radiance observations from window channels are typically not assimilated. Accurate LST estimation is therefore critical to improving our estimates of the surface water, energy, and radiation balance as well as atmospheric temperature and humidity profiles, which in turn are all critical to improving weather and climate forecast accuracy.

Satellite retrievals of LST (also referred to as the skin temperature) are available from a variety of polar-orbiting and geostationary platforms carrying infrared and microwave sensors (Aires et al. 2004; Jin 2004; Minnis and Khaiyer 2000; Pinheiro et al. 2004; Rossow and Schiffer 1991; 1999; Trigo and Viterbo 2003; Wan and Li 1997). Land surface models (driven by observed meteorological forcing data or coupled to an atmospheric model) offer estimates of global land surface conditions, including LST. Errors in the forcing fields, however, along with the imperfect parameterization of land–atmosphere interactions can lead to considerable drifts in modeled land surface states. Land data assimilation systems combine the complementary information from modeled and observed land surface fields and produce dynamically consistent, spatially complete and temporally continuous estimates of global land surface conditions. Assimilating LST retrievals into a land surface model should, in concept, improve model estimates of land surface conditions.

There has been considerable progress in the methodological development and application of land data assimilation algorithms (Andreadis and Lettenmaier 2006; Balsamo et al. 2007; Crow and Wood 2003; Drusch 2007; Dunne and Entekhabi 2006; Mahfouf et al. 2009; Margulis et al. 2002; Pan and Wood 2006; Reichle et al. 2009; Seuffert et al. 2003; Slater and Clark 2006; Walker et al. 2001; Zaitchik and Rodell 2009; Zhou et al. 2006), with ensemble-based Kalman filtering and smoothing algorithms emerging as a promising method for land data assimilation. Development and applications of land data assimilation for global land surface and land–atmosphere models, however, have largely focused on assimilating observations of surface soil moisture, snow cover, and snow water equivalent, with less effort devoted to LST assimilation.

The goal of this study is to investigate the potential for assimilating satellite retrievals of LST within a state-of-the-art land surface data assimilation system. Specifically, LST retrievals from the International Satellite Cloud Climatology Project (ISCCP) are assimilated into the National Aeronautics and Space Administration's (NASA) Catchment land surface model (CLSM; Koster et al. 2000) and into the Noah land surface model (Ek et al. 2003) with the ensemble Kalman filter (EnKF) developed at the NASA Global Modeling and Assimilation Office (Reichle et al. 2009). For validation of the assimilation products we use in situ observations from the Coordinated Energy and Water Cycle Observations Project (CEOP). We pay particular attention to the bias between the observed and modeled LSTs and have fitted the EnKF with several bias estimation algorithms designed specifically to address LST biases (section 4c). It will be shown that the assimilation algorithm must be customized for the model-specific representation of LST.

2. Background

Dating back to the early 1980s, infrared and microwave sensors on geostationary and polar-orbiting platforms have been used to retrieve LSTs. Given the extensive global and multidecadal record of these satellite-based LST retrievals, and given the importance of accurate LST estimation in particular for global atmospheric data assimilation systems, it is telling that the challenge of operational LST assimilation has been largely unmet.

The difficulties of LST data assimilation are rooted in the nature of LST retrievals and modeling. LST data from retrievals and land surface models typically exhibit strong biases that depend on the observation time and location and that have been well documented (see, e.g., Jin et al. 1997; Trigo and Viterbo 2003; Jin 2004; see also section 6 for examples). Biases arise for a variety of reasons. For instance, LST modeling is fraught with numerical stability problems because in nature the effective heat capacity associated with LST is very small. Land modelers are thus forced to approximate the corresponding heat capacity as zero or to use a surface temperature prognostic variable that represents more than just a very thin layer. The first approach, used for example in Noah, derives LST as a diagnostic variable from the surface energy balance. The second approach, used for example in CLSM, lumps the vegetation and top few centimeters of soil matter into a single model prognostic variable with a small but nonzero heat capacity. The latter approach is obviously at odds with satellite retrievals of LST, which describe the temperature in a much shallower layer at the land–atmosphere

interface (vegetation or soil, as viewed from the satellite sensor). On the other hand, the zero-heat capacity approach requires an additional connection between the diagnostic model LST variable and a model prognostic variable to which data assimilation increments can be applied (so that they can alter the model forecast). Additional discrepancies between the LST observed by the satellite and that computed by the land model stem from the inability of global land models to resolve the land surface at the same fine horizontal resolution as infrared satellite sensors.

Satellite-based LST retrievals suffer from their own set of disadvantages and problems. Infrared LST retrievals are only available under clear-sky conditions and are notoriously prone to cloud contamination (Jin 2004). Microwave LST retrievals are available under cloudy conditions (Aires et al. 2004) but depend on uncertain estimates of microwave land surface emissivity. Both infrared and microwave LST retrievals depend on knowledge of the atmospheric conditions above the LST footprint. LST retrievals also depend on the look angle, which is particularly important for retrievals at high viewing angles (Minnis and Khaiyer 2000; Pinheiro et al. 2004). The benefit of having a great variety and long record of different platforms from which LST can be retrieved is partly negated by the corresponding variety of sensor characteristics and sensor-specific LST retrieval algorithms that make it difficult to achieve a homogeneous satellite LST record.

Additional complications arise when model and satellite LSTs are combined in a data assimilation system. The strong seasonal and diurnal cycles of LST must be considered because error characteristics may depend on the time of day and season. Moreover, obvious problems result when (clear sky) LST retrievals are assimilated into a model at a time and location for which the model state or forcing indicates cloudy conditions. Because the infrared and microwave emissivities of the land surface are not well known, it is difficult to compare the radiometric temperature observed by the satellite with the physical temperature of the land model. Last but not least, the dearth of validating in situ observations of LST and land surface fluxes is a severe impediment to the validation of data products from satellite observations, modeling, and data assimilation.

One major development path is to assimilate LST retrievals into a simple representation of the land surface energy balance using an adjoint-based variational assimilation approach (Castelli et al. 1999; Boni et al. 2001; Caparrini et al. 2004; Sini et al. 2008). This elegant method requires only a minimal amount of ancillary data and provides robust estimates of the evaporative fraction. It is not, however, easily applicable to existing global atmospheric or land data assimilation systems because it is very difficult to develop and maintain adjoint models for the complex land surface model components in such systems. Recently, Meng et al. (2009) developed the adjoint model of just the surface energy balance component of the Common Land Model. Using the variational method, the authors assimilated in situ LST observations from four AmeriFlux sites for up to 20 days and report improvements in evapotranspiration estimates when verified against coincident in situ observations.

Other offline surface temperature assimilation studies used filtering techniques. Kumar and Kaleita (2003) used the extended Kalman filter to assimilate in situ observations of surface soil temperature from a single site for 1 month into a soil heat transfer model based on the discretized diffusion equation. Lakshmi (2000) merged 1 yr of satellite retrievals of LST over the Red River–Arkansas River basin into a simple two-layer model of the surface water and energy balance.

A few attempts have been made to adjust terms in the surface energy balance of atmospheric models in response to satellite LST retrievals. McNider et al. (1994) describe a technique to assimilate satellite LST into the surface energy budget of a regional-scale atmospheric model. Their method assumes that surface soil moisture is the least known parameter in the early morning surface energy budget. Van den Hurk et al. (2002) assimilated satellite LST and near-surface relative humidity measurements into a regional weather forecast model. By adjusting the root zone soil moisture and the roughness length for heat, the authors find small improvements in the surface energy balance. Garand (2003) outlines a variational method for a unified land and ocean surface skin temperature analysis, including a linear a priori bias correction of the assimilated radiances.

Bosilovich et al. (2007) developed an algorithm for LST assimilation into a global model that introduces an incremental bias correction term into the model's surface energy budget. In contrast to the McNider et al. (1994) approach, all temperature-dependent terms in the surface energy budget respond directly to the LST retrievals. In its simplest form, the Bosilovich et al. (2007) algorithm estimates and corrects a constant time mean bias for each grid point; additional benefits are attained with a refined version of the algorithm that allows for a correction of the mean diurnal cycle. The results of Bosilovich et al. (2007) indicate that LST assimilation improves estimates of 2-m air temperature, both in mean and variability, in a coupled land–atmosphere model. Neglecting the diurnal cycle of the LST bias causes degradation of the diurnal amplitude of the background model air temperature in many regions.

In situ measurements of energy fluxes at several locations were used to inspect the surface energy budget more closely. LST assimilation generally improves the sensible heat flux and, in some cases, it improves the Bowen ratio. At many stations, however, LST assimilation slightly increases the bias in the monthly latent heat flux. A critical limitation of the method of Bosilovich et al. (2007) is the assumption of unbiased LST retrievals.

In this paper, we restrict ourselves to uncoupled land data assimilation and test several variants of the bias estimation strategy of Bosilovich et al. (2007). We also explore an alternative strategy of scaling the LST retrievals into the climatology of the modeled LST. As we will show in section 6, not scaling the LST retrievals prior to (uncoupled land) data assimilation can create serious imbalances in the model-generated mass and energy fluxes and can lead to entirely unrealistic land surface fluxes. We test these approaches with two land surface models that represent LST very differently: CLSM and Noah (section 4). Our results are directly linked to weather and climate prediction applications because these two land models are used in the atmospheric data assimilation systems of the NASA Global Modeling and Assimilation Office (GMAO) and the National Oceanic and Atmospheric Administration/National Centers for Environmental Prediction (NOAA/NCEP), respectively.

3. Data

In this study, we assimilate LST retrievals from ISCCP (information online at http://isccp.giss.nasa.gov; see also Rossow and Schiffer 1991, 1999). The ISCCP archive contains satellite-based estimates of global cloud cover and radiative properties from 1983 through the present (recent data are added with a latency of about 1 yr), and is based on observations from an international network of meteorological satellites. Specifically, the ISCCP 30-km Pixel Level Cloud product (DX) includes global, 3-hourly, clear-sky LST retrievals from infrared radiances. For this study, we extracted LST retrievals from the DX archive for the geostationary platforms, including the U.S. Geostationary Operational Environmental Satellites (GOES) series, the European Meteosat series, and the Japanese Geostationary Meteorological Satellite (GMS) series. The data were then aggregated to a global latitude–longitude grid with 1° resolution for assimilation into our system.

The availability of validating land surface temperature and flux data is very limited. In this study, we use the comparably large collection of such data provided by the Coordinated Energy and Water Cycle Observations Project (CEOP; information online at http://www.ceop.net) to validate the data assimilation products. Specifically, we obtained hourly data from the network of CEOP reference sites from 1 October 2002 through 31 December 2004 (Fig. 1, Table 1). Sufficient data for validation are available at 48 distinct sites, of which 19 sites have LST data, 30 have latent heat (LH) and sensible heat (SH) flux data, and 20 sites have ground heat (GH) flux data. Only four stations have LST as well as LH, SH, and GH observations sufficient for validation (Cabauw, the Netherlands; Bondville, Illinois; Lindenberg Falkenberg, Germany; and Lindenberg Forest, Germany; see also section 5). The hourly CEOP data were aggregated to 3-hourly averages for comparison with the 3-hourly retrieval, model, and assimilation products.

The surface meteorological forcing data for the two land models are from the Global Land Data Assimilation Systems (GLDAS) project (Rodell et al. 2004; information online http://ldas.gsfc.nasa.gov) and were provided at 3-hourly time steps and at 2° and 2.5° resolutions in latitude and longitude, respectively. The GLDAS data stream is based on output from the global atmospheric data assimilation system at the NASA Global Modeling and Assimilation Office (Bloom et al. 2005) adjusted with pentad precipitation observations from the Climate Prediction Center Merged Analysis of Precipitation (CMAP; information online at http://www.cdc.noaa.gov/cdc/data.cmap.html) and daily estimates of surface radiation from the Air Force Weather Agency's (AFWA) Agricultural Meteorology (AGRMET) system. The observation-based corrections ensure that the forcing data and hence the land model output are as close to reality as is possible (without the benefit of assimilating the LST retrievals).

4. Data assimilation system

a. The Catchment and Noah land surface models

Modeled LST and land surface fluxes are from integrations of CLSM (Koster et al. 2000) and Noah (Ek et al. 2003). Again, these two models are the land model components of the atmospheric data assimilation and forecasting systems at GMAO and NCEP, respectively. Both models dynamically predict land surface water and energy fluxes in response to surface meteorological forcing but follow markedly different approaches to modeling soil moisture and LST.

CLSM's basic computational unit is the hydrological catchment (or watershed). The global land surface is divided into catchments (excluding inland water and ice-covered areas) with a mean linear scale of around 50 km (ranging from a just a few to 250 km). Unlike traditional,

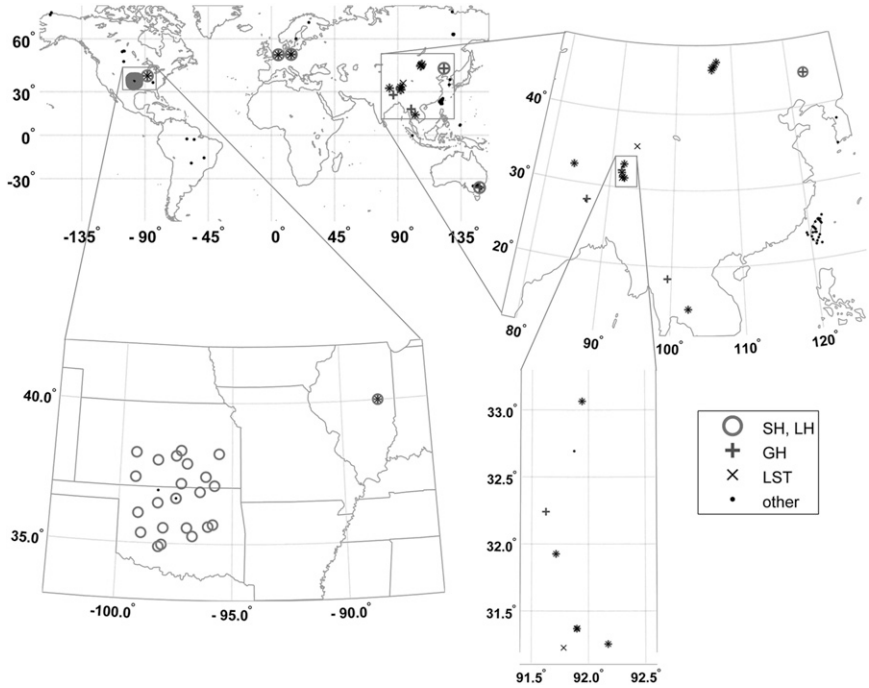

FIG. 1. Locations of CEOP stations. Stations suitable for validation are indicated with circles (SH and LH, 30 stations), plus signs (GH, 20 stations), and crosses (LST, 19 stations).

layer-based models, CLSM includes an explicit treatment of the spatial variation of soil water and water table depth within each hydrological catchment based on the statistics of the catchment topography. The surface energy balance is computed separately for the (dynamically varying) saturated, transpiring, and wilting subareas of each catchment. In each of these three subareas, the bulk temperature of the vegetation canopy and the top 5 cm of the soil column is modeled with a ''surface temperature'' (TSURF) prognostic variable that is specific to the soil moisture regime. The three TSURF prognostic variables interact with an underlying heat diffusion model for soil temperature (consisting of six layers with depths equal to about 10, 20, 40, 75, 150, and 1000 cm from top to bottom) that is common to the three subareas (see Fig. 5 in Koster et al. 2000). In the absence of snow, the area-weighted average of the three prognostic TSURF variables (hereinafter also referred to as the ''surface temperature'' in CLSM) is the most appropriate quantity to compare to satellite-based LST retrievals (Fig. 2). CLSM integrations were conducted using the GMAO land data assimilation system (Reichle et al. 2009) with a model time step of 20 min.

Noah is a more traditional, layer-based model. Four soil layers of increasing thicknesses of 10, 30, 60, and 100 cm are used to model soil temperature and moisture dynamics with layer-based formulations of the heat diffusion equation (for energy) and of the standard diffusion and gravity drainage equations (for unsaturated water flow). LST in Noah is diagnosed from the surface energy balance equation and, unlike in CLSM, is *not* a prognostic variable and has no associated heat capacity (Fig. 2). In this study, we use Noah version 2.7.1 on a 0.5° grid with a 30-min time step. Noah integrations were carried out with the Land Information System (Kumar et al. 2008) fitted with the GMAO ensemble data assimilation and bias estimation modules (Reichle et al. 2009).

b. Data assimilation method and parameters

In a data assimilation system, the model-generated land surface estimates are corrected toward observational estimates, with the degree of correction determined by the levels of error associated with each. The assimilation system used here is based on the ensemble Kalman filter (EnKF), which is well suited to the nonlinear and intermittent character of land surface processes (Reichle et al. 2002a,b). The EnKF works sequentially by performing in turn a model forecast step and a filter update step. Formally, the forecast step for ensemble member i can be written as

$$\mathbf{x}_{t,i}^{-} = f(\mathbf{x}_{t-1,i}^{+}, \mathbf{q}_{t,i}), \tag{1}$$

where $\mathbf{x}_{t,i}^{-}$ and $\mathbf{x}_{t-1,i}^{+}$ are the forecast (denoted with a superscript minus sign) and analysis (denoted with a superscript plus sign) state vectors at times t and $t-1$,

TABLE 1. CEOP stations with sufficient in situ LST, LH, SH, or GH observations for validation. The final column groups stations within each Continental-Scale Experiment (CSE) into categories with similar data availability (a = LH+SH; b,c = LST+LH+SH+GH; d,h = LH+SH+GH; e = LST+GH; f = LST; g = GH; see Fig. 3).

CEOP identifier			Coordinates (°)		
CSE	Reference site	Station	Lat	Lon	Data availability
Baltic Sea Experiment (BALTEX)	Cabauw	Cabauw	51.97	4.93	b
BALTEX	Lindenberg	Falkenberg	52.17	14.12	b
BALTEX	Lindenberg	Forest	52.18	13.95	b
CEOP Asia–Australia Monsoon Project (CAMP)	Chao Phraya River, Thailand	Lampang	18.40	99.47	g
CAMP	Himalayas	Pyramid	27.96	86.81	g
CAMP	Mongolia	BTS	46.78	107.14	e
CAMP	Mongolia	DGS	46.13	106.37	e
CAMP	Mongolia	DRS	46.21	106.71	e
CAMP	Mongolia	MGS	45.74	106.26	e
CAMP	NE Thailand	Nakhonrachasima	14.47	102.38	e
CAMP	Tibet	Amdo-Tower	32.24	91.62	g
CAMP	Tibet	ANNI-AWS	31.25	92.17	e
CAMP	Tibet	BJ-SAWS1	31.37	91.90	f
CAMP	Tibet	BJ-SAWS2	31.37	91.90	f
CAMP	Tibet	BJ-SAWS3	31.37	91.90	f
CAMP	Tibet	BJ-Tower	31.37	91.90	e
CAMP	Tibet	D105-AWS	33.06	91.94	e
CAMP	Tibet	D66-AWS	35.52	93.78	f
CAMP	Tibet	Gaize	32.30	84.05	e
CAMP	Tibet	MS3478-AWS	31.93	91.71	e
CAMP	Tibet	MS3608-AWS	31.23	91.78	f
CAMP	Tongyu, China	Cropland	44.42	122.87	h
CAMP	Tongyu	Grassland	44.42	122.87	h
GEWEX American Prediction Project (GAPP)	Bondville	Bondville	40.01	−88.29	c
GAPP	Southern Great Plains (SGP) site, Atmospheric Radiation Measurement (ARM) program	E 1 Larned	38.20	−99.32	a
GAPP	SGP	E 2 Hillsboro	38.31	−97.30	a
GAPP	SGP	E 3 Le_Roy	38.20	−95.60	a
GAPP	SGP	E 4 Plevna	37.95	−98.33	a
GAPP	SGP	E 5 Halstead	38.11	−97.51	a
GAPP	SGP	E 6 Towanda	37.84	−97.02	a
GAPP	SGP	E 7 Elk_Falls	37.38	−96.18	a
GAPP	SGP	E 8 Coldwater	37.33	−99.31	a
GAPP	SGP	E 9 Ashton	37.13	−97.27	a
GAPP	SGP	E 10 Tyro	37.07	−95.79	a
GAPP	SGP	E 12 Pawhuska	36.84	−96.43	a
GAPP	SGP	E 13 Lamont	36.60	−97.48	a
GAPP	SGP	E 14 Lamont	36.61	−97.49	a
GAPP	SGP	E 15 Ringwood	36.43	−98.28	a
GAPP	SGP	E16 Vici	36.06	−99.13	a
GAPP	SGP	E 18 Morris	35.69	−95.86	a
GAPP	SGP	E 19 El_Reno	35.55	−98.02	a
GAPP	SGP	E 20 Meeker	35.56	−96.99	a
GAPP	SGP	E 21 Okmulgee	35.62	−96.06	a
GAPP	SGP	E 22 Cordell	35.35	−98.98	a
GAPP	SGP	E 24 Cyril	34.88	−98.20	a
GAPP	SGP	E 26 Cement	34.96	−98.08	a
GAPP	SGP	E 27 Earlsboro	35.27	−96.74	a
Murray–Darling Basin Water Budget Project (MDB)	Tumbarumba, NSW, Australia	Tumbarumba	−35.65	148.15	d

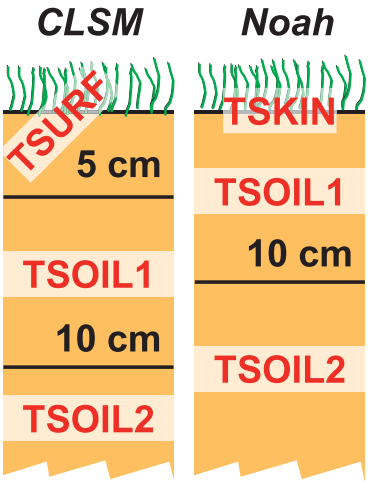

FIG. 2. LST is described in (left) CLSM with a prognostic variable (TSURF) and (right) Noah with a diagnostic variable (TSKIN). LST increments are applied to TSURF in CLSM and to TSOIL1 in Noah (section 4b).

respectively. The model error (or perturbation vector) is denoted with $\mathbf{q}_{t,i}$. The filter update produces the analyzed state vector $\mathbf{x}_{t,i}^{+}$ at time t and can be written as

$$\mathbf{x}_{t,i}^{+} = \mathbf{x}_{t,i}^{-} + \mathbf{K}_{\mathbf{x},t}(\mathbf{y}_{t,i} - \mathbf{H}_t\mathbf{x}_{t,i}^{-}), \tag{2}$$

where $\mathbf{y}_{t,i}$ denotes the observation vector (suitably perturbed) and $\mathbf{H}_t$ is the observation operator [which is written as if it were linear for ease of notation, but in practice the update is solved without explicitly computing $\mathbf{H}_t$; Keppenne (2000)]. The Kalman gain matrix $\mathbf{K}_{\mathbf{x},t}$ is given by

$$\mathbf{K}_{\mathbf{x},t} = \mathbf{P}_{\mathbf{x},t}\mathbf{H}_t^{\mathrm{T}}(\mathbf{H}_t\mathbf{P}_{\mathbf{x},t}\mathbf{H}_t^{\mathrm{T}} + \mathbf{R}_t)^{-1}, \tag{3}$$

where $\mathbf{P}_{\mathbf{x},t}$ is the state forecast error covariance (diagnosed from the ensemble $\mathbf{x}_{t,i}^{-}$), $\mathbf{R}_t$ is the observation error covariance, and a superscript T denotes the matrix transpose. Simply put, the Kalman gain $\mathbf{K}_{\mathbf{x},t}$ represents the relative weights given to the model forecast and the observations, based on their respective uncertainties and based on the error correlations between the elements of the state vector and the model prediction of the observed variable. In this paper, we use 12 ensemble members with a "one-dimensional" (1D) EnKF that processes each location independently of all other locations (see, e.g., Reichle and Koster (2003) for 1D versus 3D assimilation). The key feature of the EnKF is that error estimates of the model-generated results are dynamically derived from an ensemble of nonlinear model integrations.

Perturbation fields were generated and applied to the surface air temperature and radiation, which represent the dominant forcing inputs for LST. Perturbations were also applied to model prognostic variables (the Catchment surface temperature, TSURF, and the Noah top layer soil temperature, TSOIL1), reflecting errors in the modeling of the surface energy balance. Collectively, the perturbations allow us to maintain an ensemble of land surface conditions that represents the uncertainty in modeled LSTs. An overview of the perturbation parameters is given in Table 2. Depending on the variable, normally distributed additive perturbations or lognormally distributed multiplicative perturbations were applied. The ensemble mean for all perturbations was constrained to zero for additive perturbations and to one for multiplicative perturbations. Moreover, time series correlations were imposed via a first-order autoregressive model [AR(1)] for all fields. Since we used a one-dimensional EnKF in this study, the perturbation fields were not spatially correlated. At hourly and daily time scales, the meteorological forcing fields are ultimately based on output from atmospheric modeling and analysis systems and not on direct observations of surface forcings. We imposed error cross correlations that are motivated by the assumption that the atmospheric forcing fields represent a realistic balance between radiation, clouds, and air temperature. Under that assumption, for example, a positive perturbation to the downward shortwave radiation tends to be associated with negative perturbations to the longwave radiation and with a positive perturbation to air temperature.

Model errors are difficult to quantify at the global scale. The parameter values listed in Table 2 are largely based on experience and are partially supported by earlier studies (Reichle et al. 2002b; Reichle and Koster 2003; Reichle et al. 2007). The success of the assimilation presented here (section 6) suggests that these values are acceptable. In any case, further calibration of the filter parameters would, in theory, only improve the assimilation results. Clearly, more research is needed on the exact nature of the model and forcing errors. Recently developed adaptive filtering methods for land assimilation may assist with error parameter estimation (Reichle et al. 2008; Crow and Reichle 2008).

The mapping of the satellite information from the observation space into the space of the model states is accomplished through the Kalman gain during the EnKF update step. Equation (2) linearly relates "innovations" (observations minus corresponding model estimates before EnKF update, i.e., $\mathbf{y}_{t,i} - \mathbf{H}_t\mathbf{x}_{t,i}^{-}$) to "increments" (difference in model states after and before EnKF update, i.e., $\mathbf{x}_{t,i}^{+} - \mathbf{x}_{t,i}^{-}$). In this study, we use CLSM's area-average TSURF variable and Noah's

TABLE 2. Parameters for perturbations to meteorological forcing inputs and model prognostic variables.

Perturbation	Additive (A) or multiplicative (M)?	Std dev	AR(1) time series correlation scale	Cross correlation with perturbations in		
				T2M	SW	LW
Air temp (T2M)	A	1 K	1 day	N/A	0.4	0.4
Downward SW radiation	M	0.3	1 day	0.4	N/A	−0.6
Downward LW radiation	A	20 W m^{-2}	1 day	0.4	−0.6	N/A
Soil temp prognostic variables (Catchment, TSURF; Noah, TSOIL1)	A	0.2 K	12 h	0	0	0

diagnostic TSKIN variable to compute the (observation space) innovations (Fig. 2 and section 4a). The EnKF state vector for CLSM consists of the three TSURF prognostic variables (specific to each soil moisture regime; see section 4a), while the EnKF state vector for Noah consists of the top-layer soil temperature TSOIL1. The key ingredients to the Kalman gain are the error correlations between the LST variables in observation space and the EnKF state variables (Reichle et al. 2002b). For Noah, the relevant error correlation is between the diagnostic TSKIN variable (which has no associated heat capacity) and the temperature in the top-10-cm soil layer. Therefore, in the case of Noah the error correlation is affected by a small phase shift between the diurnal cycle of the diagnostic TSKIN and the top-layer soil temperature TSOIL1. Observations are not assimilated when precipitation is falling or when the ground is covered with snow (as indicated by the land model integration).

c. Bias estimation

For the reasons outlined in section 2 there are considerable differences between the temporal moments of the satellite and model LST (see section 6 for examples). Such biases need to be addressed in the data assimilation system. For this study, we implemented two different strategies. The first strategy is to scale the satellite observations to the model's climatology so that the first and second moments of the satellite LST and the model LST match. This strategy is a simplified version of the cumulative distribution function matching (Reichle and Koster 2004) that has been used successfully for soil moisture assimilation (Reichle et al. 2009). Because of the strong diurnal and seasonal cycles of LSTs, we estimated the multiyear LST mean and variance separately for each calendar month and for eight different times of day (0000, 0300, . . . , 2100 UTC). This strategy is very easy to implement through preprocessing of the LST retrievals and makes no assumptions regarding whether the model's or the observations' climatology is more correct. Although the assimilation estimates are by design produced in the model's climatology, they could be scaled back to the observational climatology if desired. The scaling strategy can be applied to the assimilation of retrievals from a variety of satellite datasets with different climatologies. An obvious disadvantage is the fact that the a priori scaling is static and cannot automatically adjust to dynamic changes in bias.

The second strategy is to dynamically estimate bias parameters along the lines of the algorithm developed by Dee (2005), which was used for LST by Bosilovich et al. (2007) and for soil moisture by De Lannoy et al. (2007). This dynamic bias estimation approach is based on a second Kalman filter for bias estimation (in addition to the Kalman filter for state estimation). Assume that we have a bias estimate $\mathbf{b}_{t-1}^{+}$ at time $t-1$. Furthermore, assume that this bias estimate can be propagated to time t with a simple bias evolution model:

$$\mathbf{b}_t^{-} = \alpha \mathbf{b}_{t-1}^{+}, \tag{4}$$

which relaxes the bias estimates to zero ($0 < \alpha < 1$). In our experiments, we chose α to correspond to an e-folding scale of 1 day. The use of a relaxation factor is different from the implementation of Bosilovich et al. (2007) and is a prudent strategy for experiments that cover many seasons. Because observations may not be available for extended periods, relaxing the bias estimate to zero is safer than keeping the latest bias estimate through seasons for which it may not be appropriate.

Next, we compute a bias-corrected model forecast:

$$\boldsymbol{\xi}_{t,i}^{-} = \mathbf{x}_{t,i}^{-} - \mathbf{b}_t^{-}, \tag{5}$$

which is used in the state update Eq. (2) (instead of the biased model forecast $\mathbf{x}_{t,i}^{-}$). From ensemble average innovations (computed as $\mathbf{y}_t - \mathbf{H}_t\boldsymbol{\xi}_t^{-} \equiv E(\mathbf{y}_{t,i} - \mathbf{H}_t\boldsymbol{\xi}_{t,i}^{-})$, where $E(\cdot)$ is the ensemble mean operator), we can then update the bias via

$$\mathbf{b}_t^{+} = \mathbf{b}_t^{-} - \gamma \mathbf{K}_{\mathbf{x},t}(\mathbf{y}_t - \mathbf{H}_t\boldsymbol{\xi}_t^{-}). \tag{6}$$

A key assumption of this algorithm is that the bias error covariance $\mathbf{P}_{b,t}$ is a small fraction of the state error

covariance; that is, $\mathbf{P}_{b,t} = \gamma \mathbf{P}_{\mathbf{x},t}$. Based on this assumption, the gain for the bias ($\mathbf{K}_{b,t}$) can be computed as a fraction of the gain for the state, which has already been computed. Here, we use $\gamma = 0.2$ for Catchment and $\gamma = 0.05$ for Noah. The difference in γ is motivated by the different surface layer thicknesses in the two models, but like the perturbations parameters that govern the model error, these bias parameters have not been optimized and are justified primarily by the success of the assimilation.

As formulated above, the bias algorithm estimates a single bias parameter (per state and per location). Here, we implemented a variant that estimates a separate bias parameter for eight different times of day (0000, 0300, . . . , 2100 UTC). Because this requires the estimation of eight bias parameters per state and per location, we refer to this algorithm as b8. Additional variants of the dynamic bias algorithm are discussed in the appendix.

Equation (6) implies that, in practice, the bias estimates can be thought of as an exponential moving (time) average of the LST increments. Unlike the scaling approach, the dynamic bias estimation strategy adapts to slow changes in bias over time. A major disadvantage of this strategy is the implicit assumption that only the model is biased, which contradicts previous findings that retrievals from different sensors may be biased against each other (e.g., Trigo and Viterbo 2003). It is therefore critical that any bias between retrievals from different sensors is small compared to the bias between retrievals and the model estimates.

For this study we implemented the a priori scaling method and the dynamic bias estimation schemes independently. The latter can be invoked with and without a priori scaling. If invoked without a priori scaling, the dynamic bias estimation corrects for static (long term) biases as well as the shorter-term ''bias'' that results from transient differences between the model and observational estimates. If invoked after a priori scaling, the dynamic bias estimation mostly corrects for transient bias. It can also be considered a tool for remembering assimilation increments that would otherwise be forgotten within a single model time step (because of the small heat capacity associated with LST). This is particularly important for CLSM as section 6 will show.

5. Experiment design and skill metrics

The experiment domain consists of the catchments (or grid cells) that contain the CEOP station locations (Fig. 1). The land models were spun up by cycling 10 times through the 4-yr period from 1 January 2001 to 1 January 2005. The models were then integrated in ensemble mode (12 members; using the perturbations settings of Table 2) for the same 4-yr period. These open-loop integrations also recorded the LST innovations (without applying any increments) for the computation of the model and retrieval statistics that are required for the a priori scaling approach. The (ensemble) assimilation integrations covered the same 4-yr period and were validated against the 27 months of CEOP observations from 1 October 2002 through 1 January 2005.

For each of the two land models, we conducted one open-loop (no assimilation) ensemble integration and four different experiments in which ISCCP retrievals were assimilated assuming an observation error standard deviation of 2 K. Two of the four assimilation integrations (per model) were performed with the (unscaled) ISCCP retrievals (s0, for no scaling), the other two utilized ISCCP retrievals that were scaled to each model's LST climatology prior to assimilation (s1; see section 4c). In each set of two assimilation integrations, one was done without bias correction (b0), and the other used the dynamic bias algorithms (b8; see section 4c). For each model, we thus compare four assimilation integrations: s0b0, s0b8, s1b0, and s1b8. (See the boldface entries in Table A1 in the appendix for an overview of the experiments discussed in the main text.)

All integrations were analyzed by computing RMSE values (from raw time series) and anomaly correlation coefficients (R; from anomaly time series) for LST, LH, SH, and GH with respect to the available in situ CEOP observations. These performance metrics were first computed separately for each station from 3-hourly time series and then averaged over the available stations. Anomaly time series were computed by subtracting the monthly mean seasonal and diurnal cycles (climatology) from the raw data. The climatologies and metrics were computed only from data at times and locations where ISCCP retrievals were available so that only clear-sky conditions are compared to the extent possible. Mean values for a given calendar month and time of day are computed only if a minimum of 20 data points are available; mean monthly values are computed only if mean values are available for all eight times of day for the month in question; and performance metrics are based on at least 100 data points. Figure 3 shows the average fraction of the 27-month experiment period that was used in the computation of the metrics for eight groups of stations with similar data availability. The combined mask of the assimilated ISCCP observations and the in situ measurements (black bars) primarily indicates the impacts of the clear-sky masking, with data availability ranging from 20% to 60%. The additional impacts of the minimum data requirements for the climatologies result in the gray bars, which are typically lower by a few percent.

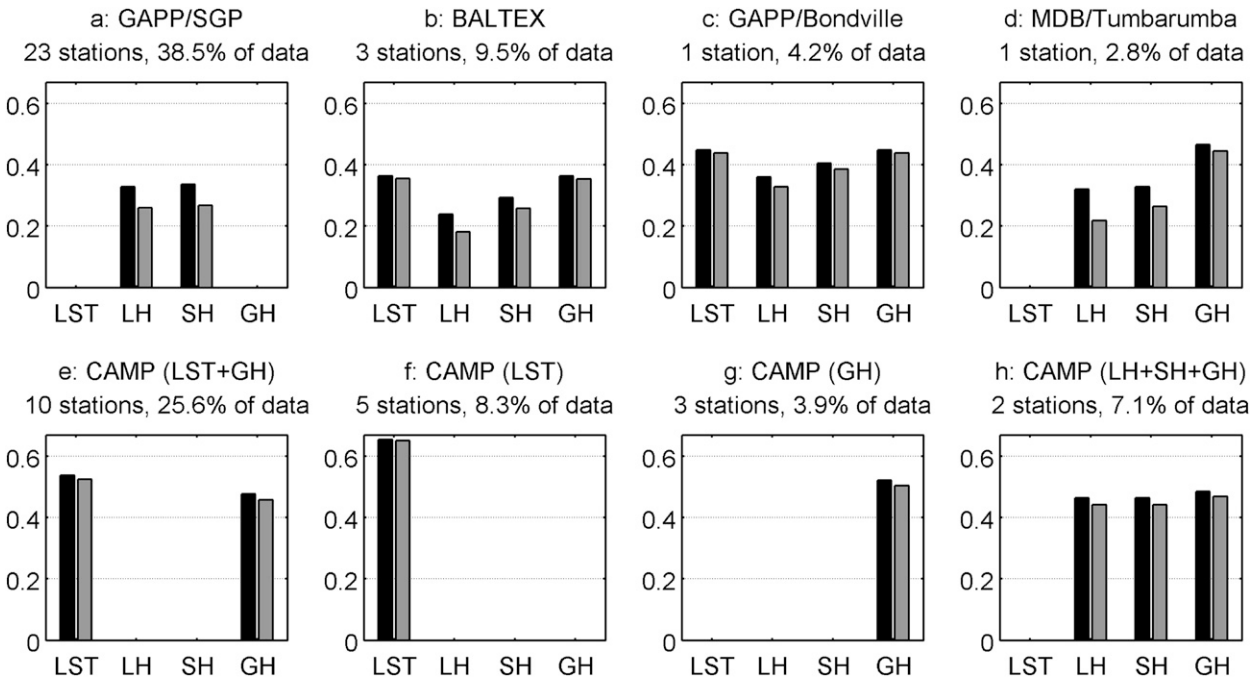

FIG. 3. The fraction of the 27-month experiment period used in the computation of (black bars) RMSE and (gray bars) anomaly R for eight groups of stations [(a)–(g) correspond to data availability categories in Table 1]. Subplot titles indicate the number of stations in each group and the percent of the data used for metrics computation in each group relative to all of the data used.

We analyze two performance metrics, RMSE and anomaly R, to highlight the advantages and disadvantages of the various assimilation approaches. Given the typically strong seasonal and diurnal cycles of the LST and land surface fluxes, metrics based on raw data are dominated by errors in the climatology. Metrics based on anomalies, by contrast, primarily capture day-to-day variations. RMSE values measure how closely the data agree in their original units and are affected by a mean bias or a mean difference in the amplitude of variations. On the other hand, R values are not affected by such biases and only capture the correlation between the estimates and the validating observations. The choice of metric depends on the application at hand (Entekhabi et al. 2010). RMSE values are most relevant if absolute errors matter most. In other cases, anomaly R values may be of most relevance, for example, in model-based applications (such as numerical weather prediction) that could correct for known biases in the mean and variance.

6. Results

a. *Aggregate performance of retrieval and model estimates*

The station-average RMSE and R metrics, evaluated against the CEOP in situ observations as discussed in the previous section, measure the aggregate performance of the satellite, model, and assimilation estimates. In this section, we highlight the performance of select assimilation integrations in terms of select metrics (see the appendix for a complete table of the metrics and algorithms). Before turning to the assimilation integrations, however, we first assess the skill of the satellite retrievals and of the model integrations without assimilation (open loop). The top panel in Fig. 4 illustrates the RMSE values computed from the raw LST estimates, which are 4.9 K for CLSM (yellow bar), 5.5 K for Noah (light blue bar), and 7.6 K for ISCCP (black bar).

The bottom panel in Fig. 4 shows corresponding RMSE values for model estimates of LH, SH, and GH, which range from 51 to 67 W m^{-2}. These values are within a factor of 2 of the typical measurement errors (around 30 W m^{-2}) for surface turbulent fluxes (Finkelstein and Sims 2001; Hollinger and Richardson 2005) and are consistent with literature estimates (Kalma et al. 2008).

The situation is similar for the anomaly R metric, shown in Fig. 5 for LST only. The models show reasonable skill in terms of reproducing standardized anomalies, with anomaly R values of 0.61 for CLSM and 0.63 for Noah. Again, ISCCP retrievals are considerably less skillful than model estimates (anomaly R value of 0.52). The small difference in the anomaly R values for the two model estimates indicates that the difference in

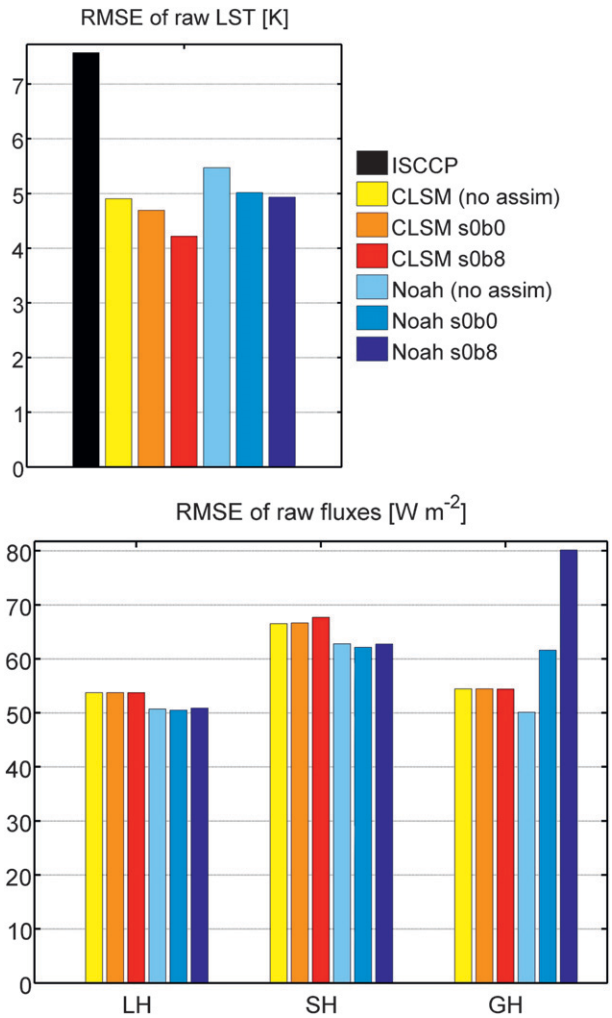

FIG. 4. RMSE vs CEOP in situ observations for (top) LST and (bottom) flux estimates from ISCCP retrievals (LST only), model integrations, and select assimilation integrations without a priori scaling.

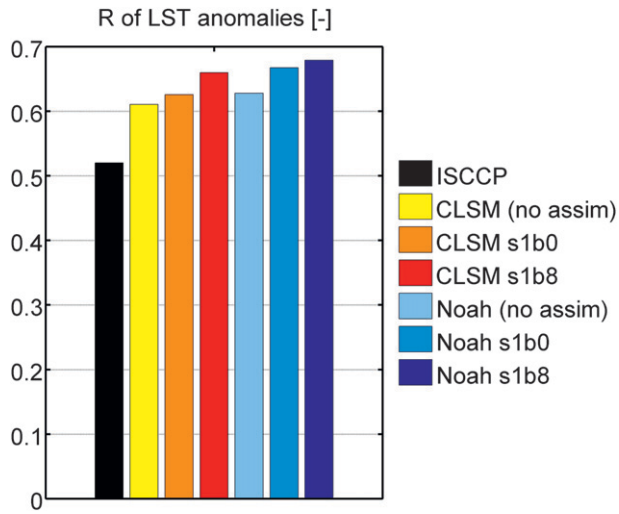

FIG. 5. The R value vs CEOP in situ observations for LST anomalies from ISCCP retrievals, model integrations, and select assimilation integrations with a priori scaling.

the models' RMSE is largely due to bias; that is, CLSM has smaller bias than Noah when compared to the in situ observations. The first important results of Figs. 4 and 5 are therefore that (i) the CLSM and Noah open-loop integrations show similar levels of skill when compared to CEOP in situ observations and that (ii) the model estimates of LST are significantly better than the ISCCP retrievals.

For a rough analysis of the sampling errors in the station-average metrics, we first computed 95% confidence intervals Δ_i for each station ($i = 1, \ldots, n$). The values of Δ_i are determined by the values of the estimated metric and by the amount of data used in its computation (based on the Fisher transformation for anomaly R and on a chi-square distribution for RMSE). We then computed 95% confidence intervals for the station-average metrics Δ_{avg} by averaging the width of the individual confidence intervals across all stations and dividing by the square root of the number of stations; that is, $\Delta_{\mathrm{avg}} = \mathrm{mean}(\Delta_i)/\sqrt{n}$. This analysis of sampling error reveals that the 95% confidence intervals for all R values discussed here are less than ± 0.01. For RMSE values, 95% confidence intervals are less than ± 0.1 K for LST and less than ± 1 W m^{-2} for surface fluxes. In the following, RMSE and anomaly R values are rounded accordingly (skill difference values rounded after computing the difference). We also investigated the sensitivity of the ISCCP skill to horizontal resolution by processing the ISCCP retrievals onto the 0.5° grid of the Noah integrations. The anomaly R is unchanged (0.52) and the RMSE decreases only marginally (from 7.6 to 7.4 K). The sensitivity may be small because the additional spatial averaging in the 1.0° dataset filters out some of the noise in the ISCCP observations and because most of the RMSE results from larger-scale errors.

b. *Aggregate performance of assimilation estimates*

Obviously, the superior skill of the model LST estimates relative to the skill of the ISCCP retrievals limits the improvements that can be expected from assimilating the ISCCP data. Nevertheless, as will be shown, modest yet statistically significant improvements can be achieved through the assimilation of ISCCP LST retrievals. We expect that the use of a priori scaling produces better anomaly estimates, whereas omitting a priori scaling should yield better improvements in terms of absolute numbers (i.e., raw data) due to likely

biases in model climatology. Thus, for RMSEs computed from raw data (Fig. 4), we focus on assimilation without a priori scaling (s0b0, s0b8). The CLSM assimilation integration without dynamic bias correction (s0b0; orange bars in Fig. 4) is characterized by only minor improvements in LST (0.2 K for RMSE) and virtually no changes in the land surface flux estimates (when compared to the open-loop skill). Recall that CLSM's prognostic "surface temperature" represents the canopy and the top 5 cm of soil and is associated with a very small heat capacity. Consequently, without dynamic bias estimation and correction, increments from the assimilation of ISCCP retrievals dissipate quickly and have little impact on the model state in CLSM. Increments applied to the Noah model's top-layer (10 cm) soil temperature, on the other hand, have a somewhat more noticeable effect on the model state. For assimilation without a priori scaling and without dynamic bias correction (s0b0), the RMSE for Noah LST estimates is improved by 0.5 K (medium blue bar in Fig. 4). Similar to CLSM, the Noah assimilation estimates for the latent and sensible heat fluxes are comparable to the open-loop estimates. However, the RMSE value for the Noah s0b0 estimates of GH increases by 12 W m^{-2} (more on this later).

Next, we analyze the skill of the assimilation integrations with dynamic bias estimation and correction (s0b8), also illustrated in Fig. 4. For CLSM, adding dynamic bias estimation (red bar in Fig. 4) enhances the LST improvements (over the open loop) to 0.7 K in terms of RMSE. Using a bias-corrected model forecast at every model time step [Eq. (5)] enhances the impacts of the LST increments in the CLSM assimilation integrations and thereby yields enhanced improvements from the assimilation of the ISCCP retrievals (relative to not using the dynamic bias algorithm). For Noah, using the dynamic bias algorithm (s0b8; dark blue bar in Fig. 4) yields only slightly better LSTs than does the s0b0 assimilation integration without dynamic bias correction (RMSE now reduced by 0.6 K over the open loop). At the same time, however, the deterioration of the GH estimates is exacerbated in s0b8. The RMSE value for assimilation estimates of GH increases by 30 W m^{-2} when compared with the open-loop RMSE. The LH and SH estimates from s0b8 are again comparable to open-loop estimates.

For the analysis of assimilation integrations *with* a priori scaling (s1b0, s1b8), we focus on the anomaly R metric, shown in Fig. 5. Qualitatively, the results for these integrations are similar to those obtained without a priori scaling. In CLSM, a priori scaling alone (s1b0) yields only small improvements in LST (the anomaly R increases by 0.02). Most of the impact is realized through dynamic bias estimation (the anomaly R for LST increases by 0.05 over the open loop). For Noah, on the other hand, the anomaly R for the LST already increases by 0.04 (over the open loop) when only a priori scaling is applied. The estimates get only slightly better when dynamic bias correction is added (increase of 0.05 in anomaly R for LST). There is also a deterioration in the Noah GH estimates (relative to the open loop) when a priori scaling is used (with or without bias correction), but generally the loss of skill in GH is mitigated through a priori scaling (see discussion of Table A1 in the appendix). Noah assimilation estimates of LH and SH have marginally better anomaly R values than do the open-loop estimates (see the appendix), but these improvements are comparable in magnitude to the size of the 95% confidence intervals and we do not consider them to be relevant.

In summary, using dynamic bias estimation for CLSM (s0b8, s1b8) provides the best assimilation estimates and enables modest LST improvements (over the open loop) of up to 0.7 K in terms of RMSE and up to 0.05 in terms of the anomaly R. Flux estimates from these CLSM assimilation integrations are essentially identical to open-loop estimates: assimilation of LST does not lead to improved flux estimation. For Noah, assimilation without a priori scaling and without dynamic bias estimation (s0b0) already yields most of the benefits of assimilating the satellite retrievals. Using a priori scaling and/or dynamic bias estimation yields only small additional improvements. For Noah, LST improvements (over the open loop) are similar to those for CLSM: up to 0.6 K for RMSE and up to 0.05 in terms of anomaly R. Noah assimilation estimates of SH and LH are similar to the open-loop estimates, but the assimilation estimates of GH are considerably worse than the open-loop estimates (up to a 30 W m^{-2} increase in RMSE), with lesser degradation seen when a priori scaling is used. Without a priori scaling, the worsening of the GH estimates in Noah may well outweigh the benefits of the LST improvements.

As shown earlier (Figs. 1 and 3, Table 1), most CEOP stations either have LST and GH measurements or have LH and SH measurements. Only four stations measure LST and all three fluxes. The results in this section must be interpreted with this caveat in mind. Conceivably, if a large number of stations were able to provide LST, LH, and SH measurements together, and if our analyses were limited to that set of stations, we might indeed be able to show that improved LST estimates from assimilation correspond to improved estimates of the turbulent fluxes. Given the limitations of the available in situ measurements, we cannot know for sure. Such improvements might also require the increased sensitivity

of model flux estimates to LST assimilation, which may be difficult to achieve in an offline system that is dominated by surface forcings. Note, however, that 15 out of the 19 stations that have LST measurements also have GH measurements, suggesting that for Noah the LST improvements from assimilation are connected to the worsening of GH estimates.

As discussed in section 5, the RMSE and anomaly R values shown here are based on 3-h average data, including the nighttime and wintertime when fluxes are small and noise may overwhelm the signal. In a separate analysis (not shown), we also computed the performance metrics from daily average data. While RMSE values are generally lower when based on daily averages (as can be expected from increased averaging) and R values are also somewhat different in overall magnitude, the relative performance is similar regardless of whether metrics are computed from daily or 3-h average data, and the main conclusions of this section thus remain unchanged.

In a second separate analysis (not shown), we assessed the performance based strictly on summertime data and again found that our conclusions remain the same.

c. Seasonal and diurnal cycles

Estimates of the mean seasonal and diurnal cycles provide additional insights into the modeling and assimilation of LST. Figure 6 (left panels) shows the mean seasonal cycle of LST at two locations: Bondville in the midwestern United States and BJ-SAWS3 in Tibet. At both locations, the seasonal cycle estimates of the CLSM and Noah open-loop integrations agree fairly closely with each other (to within 2 K), primarily because both models are driven with the same surface meteorological forcing data. At Bondville, the open-loop estimates of the seasonal cycle also agree closely with the in situ CEOP observations. At the Tibetan station, however, the open-loop estimates are biased low by about 5 K (relative to the in situ observations). In contrast, ISCCP estimates of the seasonal cycle are biased high by about 3 K at BJ-SAWS3 and differ by up to 5 K in the first half of the year at Bondville.

By construction, the seasonal and diurnal cycle estimates of the assimilation integrations with a priori scaling closely match those of the open-loop integration (not shown). Figure 6 also shows the seasonal climatology of the assimilation integrations without a priori scaling and with dynamic bias estimation and correction (s0b8). As expected, the s0b8 assimilation integrations draw more closely to the ISCCP retrievals (when compared to the open loop). At BJ-SAWS3, this fortuitously brings the assimilation estimates of the seasonal cycle into better agreement with the in situ observations than either the ISCCP or the open-loop estimates.

Figure 6 (right panels) also illustrates the mean August diurnal cycle estimates at Bondville and BJ-SAWS3. At Bondville, the Noah open-loop estimates have a slightly higher diurnal amplitude than the CLSM estimates (because LST for Noah exceeds that of CLSM by up to 2 K during the day and by less than 1 K during the night). At BJ-SAWS3, the open-loop integrations agree closely with each other. Similar to the seasonal cycle estimates, the open-loop estimates at Bondville are in reasonable agreement with the in situ observations but are biased low (by about 5 K) at BJ-SAWS3. ISCCP retrievals, on the other hand, exhibit a weaker diurnal amplitude at Bondville than do the model or in situ observations. At BJ-SAWS3, ISCCP retrievals are biased high (compared to the in situ observations) in the morning and midday but biased low in the evening and at nighttime.

Again, the LST diurnal cycle estimates of the CLSM s0b8 assimilation integrations draw toward the ISCCP retrievals by construction (Fig. 6, right panels). This brings them closer to the in situ observations at BJ-SAWS3 but makes CLSM's nighttime estimates at Bondville worse when compared to the open loop. By contrast, the LST from the Noah s0b8 assimilation integration is similar to the open-loop integration during daytime. Only during the evening hours (at Bondville and BJ-SAWS3) are the Noah s0b8 LST estimates noticeably closer to the ISCCP retrievals. The delayed impacts of the LST assimilation are probably a result of the phase lag between the diagnostic LST (observations space) and the top-layer soil temperature (state space) in Noah.

d. Filter diagnostics

Internal filter diagnostics offer further clues about the performance of the assimilation algorithms. For a filter that operates according to its underlying assumptions (that various linearizations hold, that model and observation errors are unbiased, uncorrelated, and normally distributed), the time average of the (ensemble mean) innovations sequence $(\mathbf{y}_t - \mathbf{H}_t\mathbf{x}_t^-)$ equals zero. Moreover, the standard deviation of the "normalized" innovations $(\mathbf{y}_t - \mathbf{H}_t\mathbf{x}_t^-)(\mathbf{H}_t\mathbf{P}_{\mathbf{x},t}\mathbf{H}_t^{\mathrm{T}} + \mathbf{R}_t)^{-0.5}$ equals one (Reichle et al. 2002a). The latter diagnostic compares the actual spread in the innovations to what the filter expects. A simple interpretation is that the assumed error bars of a model forecast and its corresponding observation must have an appropriate overlap.

Figure 7 displays the distribution of these two internal filter diagnostics across the CEOP stations listed in Table 1. The top panel in Fig. 7 indicates that without a priori scaling of the ISCCP observations and without dynamic bias estimation (s0b0), biases of several Kelvin

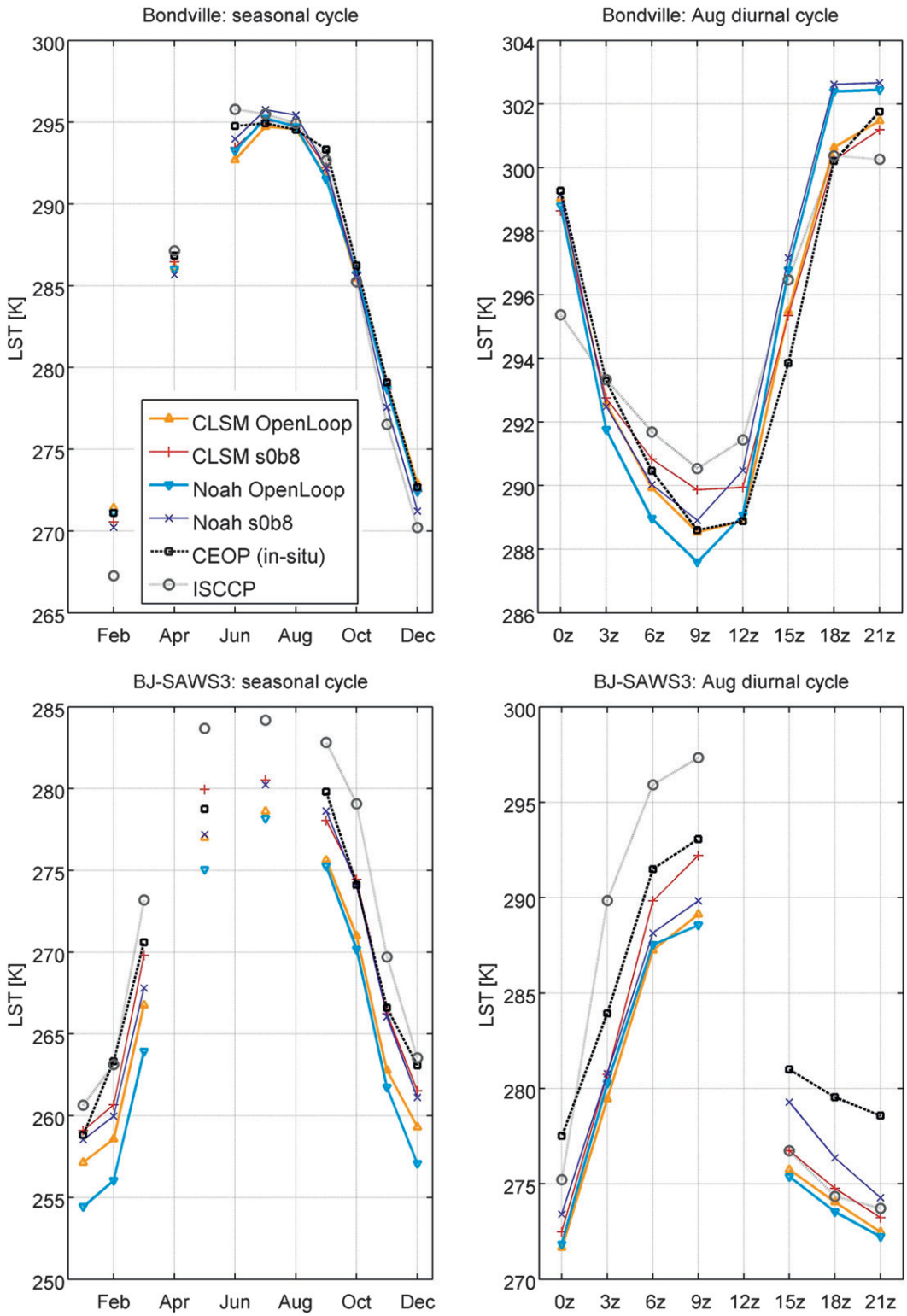

FIG. 6. LST (left) annual seasonal and (right) August diurnal cycle (UTC) at (top) Bondville and (bottom) BJ-SAWS3 for CEOP, ISCCP, model, and assimilation data.

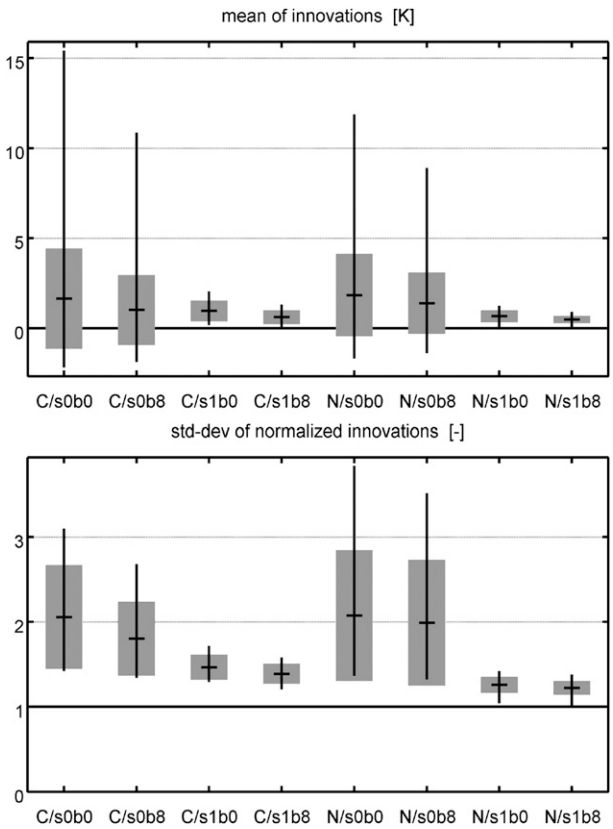

FIG. 7. (top) Mean of innovations (K) and (bottom) standard deviation of normalized innovations (dimensionless) for Catchment (C) and Noah (N) assimilation integrations. The box plots indicate the average, standard deviation, and minimum and maximum of the respective innovations diagnostic across the stations listed in Table 1.

typically persist in the model forecast and are reflected in the mean of the innovations. (The innovations statistics of the open-loop integrations are essentially the same as those of the s0b0 integrations and are not shown in Fig. 7.) A very modest reduction of the bias can be achieved with the b8 dynamic bias estimation algorithm. If in addition to dynamic bias estimation the observations are also scaled prior to data assimilation (s1b8), the innovations mean at all stations is, by construction, much closer to zero. These results hold for assimilation into both land models: CLSM and Noah.

The standard deviation of the normalized innovations, shown in the bottom panel of Fig. 7, exceeds the target value of one at almost all stations and for almost all algorithms. This indicates that the model and/or the observation error standard deviations were underestimated. As is the case for the innovations mean, a priori scaling brings the standard deviation of the normalized innovations much closer to its expected value of one. This implies that a large part of the mismatch between the actual spread in the innovations and the expected spread is simply due to bias. Finally, the fact that the innovation diagnostics are comparable for the CLSM and Noah assimilation integrations indicates that the assimilation performance (relative to its unknown optimum) is comparable for the two models, which lends further support to the broad conclusions reached in this paper. In summary, a priori scaling in combination with dynamic bias estimation exhibits the best performance in terms of the innovations diagnostics, independent of the land model used.

e. Impacts of bias in data assimilation

Estimates from a properly designed assimilation system should be no worse than open-loop estimates. The example in Fig. 8 further illustrates the potentially serious detrimental impacts of not addressing bias properly in data assimilation. Figure 8 shows LST and land surface flux time series from select Noah integrations for a few days in August of 2003 at the MGS station in Mongolia. At this location, LST and GH estimates from the open-loop integrations agree fairly well with CEOP in situ observations. The daytime peak LST estimates from ISCCP, however, are warmer by as much as 30 K. This extreme bias may be due to one or more of the reasons discussed in section 2.

At any rate, when the unscaled ISCCP retrievals are assimilated without a priori scaling (s0b0, s0b8), LST assimilation estimates are drawn toward the extreme temperatures in the ISCCP retrievals (Fig. 8). However, since Noah is not designed to accommodate such extreme temperatures, and because the surface meteorological forcing remains unchanged in the system, the Noah assimilation integrations without a priori scaling respond with unrealistic and excessive estimates of the sensible and ground heat flux, most notably on 13 August 2003. Because the impacts on LST of assimilating unscaled ISCCP retrievals without bias correction (s0b0) are more limited, the corresponding flux estimates are less pathological than in the s0b8 case with dynamic bias correction. For reference, Fig. 8 also shows an assimilation integration with a priori scaling and dynamic bias correction (s1b8), which does not produce such unrealistic flux estimates.

The situation is similar for LST from the CLSM assimilation integrations for the same location and time period (not shown), but for CLSM we obtain unrealistic estimates of the latent heat flux for select assimilation integrations without a priori scaling. To summarize, Fig. 8 illustrates the pitfalls of assimilating LST retrievals that are severely biased against the model LST. While the assimilation can be designed to produce LST estimates that are closer to the satellite retrievals, there may be

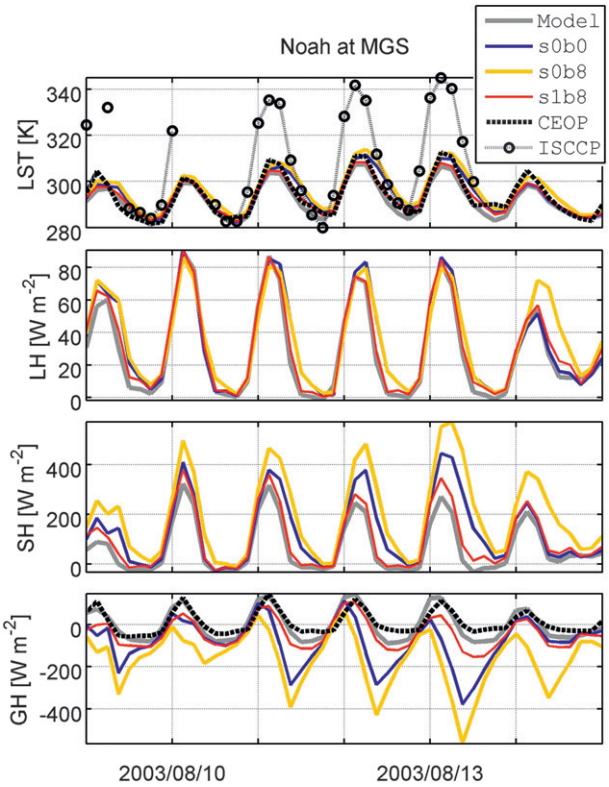

FIG. 8. (top) LST, (second from top) LH, (second from bottom) SH, and (bottom) GH for Noah integrations, ISCCP retrievals, and CEOP observations at the MGS station.

unintended and undesirable side effects in terms of the land surface fluxes and the surface energy balance.

7. Summary and conclusions

An ensemble-based data assimilation method, the EnKF, was used with and without a priori scaling of observations and/or dynamic bias estimation methods to assimilate satellite retrievals of LST into two different land surface models at 48 CEOP sites scattered around the globe (Fig. 1). The two land models, CLSM and Noah, follow distinct modeling approaches for land surface temperature (Fig. 2). CLSM has model prognostic "surface temperature" variables, whereas Noah diagnoses the surface temperature from the surface energy balance. The LST, sensible, latent, and ground heat flux estimates from the data assimilation integrations were validated against 27 months of CEOP in situ observations (Fig. 3).

The main conclusions from the experiments are as follow.

1) There are strong biases between LST estimates from in situ observations, land modeling, and satellite retrievals that vary with season and time of day. Biases of a few Kelvin are typical, with larger values exceeding 10 K (Figs. 6–8).
2) The skill of the LST estimates from the CLSM and Noah land model integrations is superior to that of the ISCCP satellite retrievals (Figs. 4 and 5).
3) Assimilation of ISCCP LST retrievals into the land surface models can improve LST estimates by up to 0.7 K for RMSE and by up to 0.05 for the anomaly R, while not making surface turbulent fluxes worse (Figs. 4 and 5).
4) Gross errors in surface flux estimates can result if biases are not taken into account properly, with a combination of a priori scaling and dynamic bias estimation methods yielding the best overall results (Figs. 4, 5, and 8).
5) Assimilation diagnostics for integrations without a priori scaling strongly reflect the underlying biases, indicating that without a priori scaling the assimilation system is far from operating in accordance with its underlying assumptions (Fig. 7).
6) Provided the assimilation system is properly configured for each land model, the benefits from the assimilation of LST retrievals are comparable for both land models (Figs. 4 and 5).

There are many reasons why the improvements from the assimilation of satellite LST, while statistically significant, turn out to be modest. First and foremost, the skill of the satellite data is modest and much lower than that of the model to begin with. The information gained by assimilating the satellite data into the model is therefore naturally limited. More accurate LST retrievals from newer datasets that have been developed with a focus on LST should alleviate this constraint in future studies. In the present study, the parameters of the assimilation system, including the perturbation (or model error) parameters and the parameters of the bias algorithm (α, γ), were not optimized. Additional calibration may further improve the results and may also reveal differences in what can be achieved with a given land model structure. Finally, even if the assimilation estimates were perfect, the performance metrics would not show it because of errors in the in situ data and because of the mismatch of the spatial and temporal characteristics of the satellite, model, and in situ datasets.

By design, the present study was limited to LST (state) estimation in an uncoupled land modeling system. Ideally, the land model parameters would be calibrated to minimize LST biases prior to data assimilation, or perhaps even dynamically within the data assimilation system. Such a calibration could rely on sophisticated parameter estimation methods (Vrugt et al. 2003). Perhaps more importantly, though, is that the surface

meteorological forcings in the present experiments were fixed. In other words, LST increments did not feed back onto the atmospheric state. Future experiments will explore LST assimilation into a coupled land–atmosphere model with the methods proposed here. The relative performance of the algorithms may very well change in the coupled environment.

Another path for future research involves investigating further the role of specific aspects of LST modeling. In the present paper, Noah was integrated in its default configuration, including a 10-cm-thick surface layer, which implied a small phase shift between the Noah diagnostic LST (used in the computations of the innovations) and the Noah top soil temperature (to which the increments were applied). Use of a thinner soil layer may alleviate the problems related to the phase shift in Noah between LST (observation space) and the top-layer soil temperature (state space).

Our results suggest that the potential for improving the surface flux estimates through the assimilation of LST retrievals is limited. However, one of the most common uses of LST observations at present is arguably the estimation of surface fluxes over irrigated land or wetland areas (Kalma et al. 2008) because flux improvements may be particularly important in areas where there are additional inputs of river water or groundwater. To the best of our knowledge, the sites considered here do not receive water inputs other than rainfall and thus do not address potential improvements in flux measurements over irrigated or wetland areas.

Even if the assimilation of LST retrievals manages to improve our LST estimates only modestly and fluxes not at all, the impacts may be significant. In coupled (land–atmosphere) data assimilation systems, improved LST may enable the *atmospheric* assimilation of additional radiance observations from surface-sensitive channels that have so far been disregarded, thereby possibly providing substantial indirect benefits. Obviously, the present study only scratches the surface of a very complex problem that has been a challenge for many years. Nevertheless, given the relative abundance of LST observations from satellites and the importance of accurate LST estimates, in particular within the context of atmospheric data assimilation, the results of the present study offer an encouraging step forward in land data assimilation.

Acknowledgments. Thanks to Michael Bosilovich, John Eylander, Peter Norris, and Christa Peters-Lidard for helpful comments. Partial support for this study was provided through NASA (NNX08AH36G) and the Air Force Weather Agency (F2BBBJ7080G001). We are grateful for access to the many datasets that supported this work, in particular the CEOP and ISCCP projects. Surface meteorological data were provided by H. Kato, J. Meng, and M. Rodell of the GLDAS project, with data contributions from the NOAA Earth System Research Laboratory. Computing was supported by the NASA High End Computing Program.

APPENDIX

Additional Bias Estimation Algorithms

In addition to the dynamic bias algorithm b8 discussed in section 4c, we also tested additional variants of the algorithm used by Bosilovich et al. (2007). Specifically, we tested the simplest possible variant that estimates a single, ''mean'' bias estimation parameter per state and per location, here referred to as b1 (also discussed in section 4c). Bosilovich et al. (2007) also introduced a bias parameterization with sine and cosine functions that accounts for variations in the diurnal cycle of the bias. Here, we implemented two variants: a ''diurnal'' bias parameterization (constant term plus sine and cosine waves with a period of 1 day) and a ''semidiurnal'' bias parameterization (diurnal terms plus sine and cosine waves with a period of ½ day). The diurnal and semidiurnal algorithms estimate three and five bias parameters, respectively, per state and location, and are referred to as b3 and b5.

For the b8 integrations discussed in the main article, we always applied both the state increments [Eq. (2)] and the bias increments [Eq. (6)]. More generally, though, any assimilation integration that uses dynamic bias estimation can also be done without applying the state increments [as, in fact, was implemented by Bosilovich et al. (2007)]. We also tested these variants. Per model, we therefore tested a total of 18 different assimilation integrations, listed in Table A1.

A close examination of Table A1 reveals that, generally, the assimilation integrations show more skill when more bias parameters are used. Reductions in RMSE values for LST are greater by up to 0.4 K for CLSM and by up to 0.2 K for Noah when comparing the b8 and b1 integrations. Corresponding differences in anomaly R values for LST are up to 0.03, respectively, for both models. As can be expected, these differences across bias estimation algorithms of varying complexity tend to be greater without a priori scaling. Also in line with our expectations, taking into account the diurnal cycle of the bias (in b3, b5, and b8, as opposed to b1) has the biggest impact.

Finally, applying the state increments in addition to the bias increments contributes only a small amount of skill, typically less than 0.1 K in terms of RMSE

TABLE A1. Skill of model and assimilation integrations vs CEOP in situ observations. Integrations shown in boldface are those discussed in the main text. Anomaly R for the LST from ISCCP is 0.52; the RMSE is 7.6 K.

				CLSM								Noah							
				Anomaly R				RMSE				Anomaly R				RMSE			
Assimilation	Apply state increments?	A priori scaling?	Bias algorithm	LST (−)	LH (−)	SH (−)	GH (−)	LST (K)	LH (W m^{-2})	SH (W m^{-2})	GH (W m^{-2})	LST (−)	LH (−)	SH (−)	GH (−)	LST (K)	LH (W m^{-2})	SH (W m^{-2})	GH (W m^{-2})
No	**N/A**	**N/A**	**N/A**	**0.61**	**0.31**	**0.21**	**0.25**	**4.9**	**54**	**67**	**54**	**0.63**	**0.26**	**0.13**	**0.24**	**5.5**	**51**	**63**	**50**
Yes	**Yes**	**s0**	**b0**	**0.63**	**0.31**	**0.21**	**0.25**	**4.7**	**54**	**67**	**54**	**0.67**	**0.27**	**0.14**	**0.14**	**5.0**	**50**	**62**	**62**
Yes	Yes	s0	b1	0.64	0.31	0.21	0.26	4.6	54	67	53	0.65	0.28	0.14	0.09	5.1	51	63	82
Yes	Yes	s0	b3	0.66	0.31	0.21	0.26	4.4	54	68	54	0.66	0.28	0.14	0.08	5.1	51	63	86
Yes	Yes	s0	b5	0.66	0.31	0.21	0.25	4.4	54	68	54	0.66	0.28	0.14	0.08	5.1	51	63	85
Yes	**Yes**	**s0**	**b8**	**0.67**	**0.31**	**0.21**	**0.25**	**4.2**	**54**	**68**	**54**	**0.68**	**0.28**	**0.14**	**0.11**	**4.9**	**51**	**63**	**80**
Yes	No	s0	b1	0.63	0.31	0.21	0.26	4.7	54	67	53	0.64	0.28	0.13	0.10	5.3	51	64	78
Yes	No	s0	b3	0.65	0.31	0.21	0.26	4.4	54	68	54	0.65	0.28	0.13	0.09	5.2	51	64	80
Yes	No	s0	b5	0.66	0.31	0.21	0.26	4.4	54	68	54	0.66	0.28	0.13	0.09	5.1	51	63	80
Yes	No	s0	b8	0.66	0.31	0.21	0.25	4.3	54	68	54	0.67	0.28	0.14	0.15	5.0	51	63	75
Yes	**Yes**	**s1**	**b0**	**0.63**	**0.31**	**0.21**	**0.25**	**4.8**	**54**	**67**	**54**	**0.67**	**0.27**	**0.14**	**0.20**	**5.2**	**50**	**62**	**53**
Yes	Yes	s1	b1	0.65	0.31	0.21	0.25	4.6	54	67	54	0.67	0.27	0.14	0.15	5.1	50	63	58
Yes	Yes	s1	b3	0.67	0.31	0.21	0.25	4.5	54	67	54	0.67	0.27	0.14	0.15	5.1	50	63	59
Yes	Yes	s1	b5	0.67	0.31	0.21	0.25	4.5	54	67	54	0.67	0.27	0.14	0.15	5.1	50	63	59
Yes	**Yes**	**s1**	**b8**	**0.66**	**0.31**	**0.21**	**0.25**	**4.6**	**54**	**67**	**54**	**0.68**	**0.27**	**0.14**	**0.17**	**5.1**	**50**	**62**	**56**
Yes	No	s1	b1	0.64	0.31	0.21	0.25	4.6	54	67	54	0.66	0.27	0.13	0.15	5.1	51	63	57
Yes	No	s1	b3	0.66	0.31	0.21	0.25	4.6	54	67	54	0.66	0.27	0.13	0.15	5.1	50	63	57
Yes	No	s1	b5	0.66	0.31	0.21	0.25	4.5	54	67	54	0.67	0.27	0.13	0.15	5.1	50	63	57
Yes	No	s1	b8	0.65	0.31	0.21	0.25	4.6	54	67	54	0.67	0.27	0.13	0.18	5.1	50	63	55

reduction (or 0.01 in terms of R increase) for LST estimates. In the case of Noah, applying the state increments contributes commensurately to the worsening of the GH estimates.

To summarize, as long as the dynamic bias algorithm takes the diurnal cycle into account, the differences that result from the exact number of bias parameters used or that result from not applying the state increments are much smaller than the assimilation improvements over the open loop. In other words, the lessons learned in the main article about the assimilation of LST retrievals in general and about using a priori scaling and/or dynamic bias correction are insensitive to the details of the dynamic bias estimation algorithm, provided the algorithm considers the diurnal cycle of the bias.

REFERENCES

Aires, F., C. Prigent, and W. B. Rossow, 2004: Temporal interpolation of global surface skin temperature diurnal cycle over land under clear and cloudy conditions. *J. Geophys. Res.,* **109,** D06214, doi:10.1029/2003JD003527.

Andreadis, K., and D. Lettenmaier, 2006: Assimilating remotely sensed snow observations into a macroscale hydrology model. *Adv. Water Resour.,* **29,** 872–886.

Balsamo, G., J.-F. Mahfouf, and S. Belair, 2007: A land data assimilation system for soil moisture and temperature: An information content study. *J. Hydrometeor.,* **8,** 1225–1242.

Bloom, S., and Coauthors, 2005: Documentation and validation of the Goddard Earth Observing System (GEOS) Data Assimilation System—Version 4. NASA Tech. Doc. Vol. 10460626, NASA Tech. Rep. Series on Global Modeling and Data Assimilation, Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, 187 pp. [Available online at http://gmao.gsfc.nasa.gov/pubs/docs/Bloom168.pdf.]

Boni, G., D. Entekhabi, and F. Castelli, 2001: Land data assimilation with satellite measurements for the estimation of surface energy balance components and surface control on evaporation. *Water Resour. Res.,* **37,** 1713–1722.

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. *J. Meteor. Soc. Japan,* **85A,** 205–228.

Caparrini, F., F. Castelli, and D. Entekhabi, 2004: Variational estimation of soil and vegetation turbulent transfer and heat flux parameters from sequences of multisensor imagery. *Water Resour. Res.,* **40,** W12515, doi:10.1029/2004WR003358.

Castelli, F., D. Entekhabi, and E. Caporali, 1999: Estimation of surface heat flux and an index of soil moisture using adjoint-state surface energy balance. *Water Resour. Res.,* **35,** 3115–3125.

Crow, W. T., and E. F. Wood, 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.

——, and R. H. Reichle, 2008: Comparison of adaptive filtering techniques for land surface data assimilation. *Water Resour. Res.,* **44,** W08423, doi:10.1029/2008WR006883.

Dee, D. P., 2005: Bias and data assimilation. *Quart. J. Roy. Meteor. Soc.,* **131,** 3323–3343, doi:10.1256/qj.05.137.

De Lannoy, G. J. M., R. H. Reichle, P. R. Houser, V. R. N. Pauwels, and N. E. C. Verhoest, 2007: Correcting for

forecast bias in soil moisture assimilation with the ensemble Kalman filter. *Water Resour. Res.,* **43,** W09410, doi:10.1029/2006WR005449.

Drusch, M., 2007: Initializing numerical weather prediction models with satellite derived surface soil moisture: Data assimilation experiments with ECMWF's Integrated Forecast System and the TMI soil moisture dataset. *J. Geophys. Res.,* **112,** D03102, doi:10.1029/2006JD007478.

Dunne, S., and D. Entekhabi, 2006: Land surface state and flux estimation using the ensemble Kalman smoother during the Southern Great Plains 1997 field experiment. *Water Resour. Res.,* **42,** W01407, doi:10.1029/2005WR004334.

Ek, M., K. Mitchell, L. Yin, P. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah land-surface model advances in the NCEP operational mesoscale Eta model. *J. Geophys. Res.,* **108,** 8851, doi:10.1029/2002JD003296.

Entekhabi, D., R. H. Reichle, R. D. Koster, and W. T. Crow, 2010: Performance metrics for soil moisture retrievals and application requirements. *J. Hydrometeor.,* **11,** 832–840.

Finkelstein, P. L., and P. F. Sims, 2001: Sampling error in eddy correlation flux measurements. *J. Geophys. Res.,* **106** (D4), 3503–3509.

Garand, L., 2003: Toward an integrated land-ocean surface skin temperature analysis from the variational assimilation of infrared radiances. *J. Appl. Meteor.,* **42,** 570–583.

Hollinger, D. Y., and A. D. Richardson, 2005: Uncertainty in eddy covariance measurements and its application to physiological models. *Tree Physiol.,* **25,** 873–885.

Jin, M., 2004: Analysis of land skin temperature using AVHRR observations. *Bull. Amer. Meteor. Soc.,* **85,** 587–600.

——, R. E. Dickinson, and A. M. Vogelmann, 1997: A comparison of CCM2–BATS skin temperature and surface-air temperature with satellite and surface observations. *J. Climate,* **10,** 1505–1524.

Kalma, J. D., T. R. McVicar, and M. F. McCabe, 2008: Estimating land surface evaporation: A review of methods using remotely sensed surface temperature data. *Surv. Geophys.,* **29,** 421–469, doi:10.1007/s10712-008-9037-z.

Keppenne, C. L., 2000: Data assimilation into a primitive-equation model with a parallel ensemble Kalman filter. *Mon. Wea. Rev.,* **128,** 1971–1981.

Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar, 2000: A catchment-based approach to modeling land surface processes in a general circulation model, 1: Model structure. *J. Geophys. Res.,* **105,** 24 809–24 822.

Kumar, P., and A. L. Kaleita, 2003: Assimilation of near-surface temperature using extended Kalman filter. *Adv. Water Resour.,* **26,** 79–93.

Kumar, S. V., R. H. Reichle, C. D. Peters-Lidard, R. D. Koster, X. Zhan, W. T. Crow, J. B. Eylander, and P. R. Houser, 2008: A land surface data assimilation framework using the land information system: Description and applications. *Adv. Water Resour.,* **31,** 1419–1432, doi:10.1016/j.advwatres.2008.01.013.

Lakshmi, V., 2000: A simple surface temperature assimilation scheme for use in land surface models. *Water Resour. Res.,* **36,** 3687–3700.

Mahfouf, J.-F., K. Bergaoui, C. Draper, F. Bouyssel, F. Taillefer, and L. Taseva, 2009: A comparison of two off-line soil analysis schemes for assimilation of screen level observations. *J. Geophys. Res.,* **114,** D08105, doi:10.1029/2008JD011077.

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

McNider, R. T., A. J. Song, D. M. Casey, P. J. Wetzel, W. L. Crosson, and R. M. Rabin, 1994: Toward a dynamic–thermodynamic assimilation of satellite surface temperature in numerical atmospheric models. *Mon. Wea. Rev.,* **122,** 2784–2787.

Meng, C. L., Z.-L. Li, X. Zhan, J. C. Shi, and C. Y. Liu, 2009: Land surface temperature data assimilation and its impact on evapotranspiration estimates from the Common Land Model. *Water Resour. Res.,* **45,** W02421, doi:10.1029/2008WR006971.

Minnis, P., and M. M. Khaiyer, 2000: Anisotropy of land surface skin temperature derived from satellite data. *J. Appl. Meteor.,* **39,** 1117–1129.

Pan, M., and E. F. Wood, 2006: Data assimilation for estimating the terrestrial water budget using a constrained ensemble Kalman filter. *J. Hydrometeor.,* **7,** 534–547.

Pinheiro, A. C. T., J. L. Privette, R. Mahoney, and C. J. Tucker, 2004: Directional effects in a daily AVHRR land surface temperature dataset over Africa. *IEEE Trans. Geosci. Remote Sens.,* **42,** 1941–1954, doi:10.1109/TGRS.2004.831886.

Reichle, R. H., and R. D. Koster, 2003: Assessing the impact of horizontal error correlations in background fields on soil moisture estimation. *J. Hydrometeor.,* **4,** 1229–1242.

——, and ——, 2004: Bias reduction in short records of satellite soil moisture. *Geophys. Res. Lett.,* **31,** L19501, doi:10.1029/2004GL020938.

——, D. McLaughlin, and D. Entekhabi, 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.

——, R. D. Koster, P. Liu, S. P. P. Mahanama, E. G. 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). *J. Geophys. Res.,* **112,** D09108, doi:10.1029/2006JD008033.

——, W. T. Crow, and C. L. Keppenne, 2008: An adaptive ensemble Kalman filter for soil moisture data assimilation. *Water Resour. Res.,* **44,** W03423, doi:10.1029/2007WR006357.

——, M. G. Bosilovich, W. T. Crow, R. D. Koster, S. V. Kumar, S. P. P. Mahanama, and B. F. Zaitchik, 2009: Recent advances in land data assimilation at the NASA Global Modeling and Assimilation Office. *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications,* S.-K. Park and L. Xu, Eds., Springer, 407–428, doi:10.1007/978-3-540-71056-1.

Rodell, M., and Coauthors, 2004: The Global Land Data Assimilation System. *Bull. Amer. Meteor. Soc.,* **85,** 381–394.

Rossow, W. B., and R. A. Schiffer, 1991: ISCCP cloud data products. *Bull. Amer. Meteor. Soc.,* **72,** 2–20.

——, and ——, 1999: Advances in understanding clouds from ISCCP. *Bull. Amer. Meteor. Soc.,* **80,** 2261–2287.

Seuffert, G., H. Wilker, P. Viterbo, J.-F. Mahfouf, M. Drusch, and J.-C. Calvet, 2003: Soil moisture analysis combining screen-level parameters and microwave brightness temperature: A test with field data. *Geophys. Res. Lett.,* **30,** 1498, doi:10.1029/2063GL017128.

Sini, F., G. Boni, F. Caparrini, and D. Entekhabi, 2008: Estimation of large-scale evaporation fields based on assimilation of remotely

sensed land temperature. *Water Resour. Res.,* **44,** W06410, doi:10.1029/2006WR005574.

Slater, A., and M. Clark, 2006: Snow data assimilation via an ensemble Kalman filter. *J. Hydrometeor.,* **7,** 478–493.

Trigo, I. F., and P. Viterbo, 2003: Clear-sky window channel radiances: A comparison between observations and the ECMWF model. *J. Appl. Meteor.,* **42,** 1463–1479.

Van den Hurk, B. J. J. M., L. Jia, C. Jacobs, M. Menenti, and Z. L. Li, 2002: Assimilation of land surface temperature data from ATSR in an NWP environment—A case study. *Int. J. Remote Sens.,* **23,** 5193–5209.

Vrugt, J. A., H. V. Gupta, L. A. Bastidas, W. Bouten, and S. Sorooshian, 2003: Effective and efficient algorithm for multiobjective optimization of hydrologic models. *Water Resour. Res.,* **39,** W1214, doi:10.1029/2002WR001746.

Walker, J. P., G. R. Willgoose, and J. D. Kalma, 2001: One-dimensional soil moisture profile retrieval by assimilation of near-surface observations: A comparison of retrieval algorithms. *Adv. Water Resour.,* **24,** 631–650.

Wan, Z., and Z.-L. Li, 1997: A physics-based algorithm for retrieving land-surface emissivity and temperature from EOS/MODIS data. *IEEE Trans. Geosci. Remote Sens.,* **35,** 980–996.

Zaitchik, B. F., and M. Rodell, 2009: Forward-looking assimilation of MODIS-derived snow-covered area into a land surface model. *J. Hydrometeor.,* **10,** 130–148, doi:10.1175/2008JHM1042.1.

Zhou, Y., D. McLaughlin, and D. Entekhabi, 2006: Assessing the performance of the ensemble Kalman filter for land surface data assimilation. *Mon. Wea. Rev.,* **134,** 2128–2142.