Ensemble Square Root Filter Assimilation of Near-Surface Soil Moisture and Reference-Level Observations into a Coupled Land Surface-Boundary Layer Model

ZHANG Shuwen* (张述文), LIU Yanhua (刘彦华), and ZHANG Weidong (张卫东)

Key Laboratory for Semi-Arid Climate Change of the Ministry of Education, Key Laboratory of Arid Climate Change and Disaster Reduction of Gansu Province, and College of Atmospheric Sciences, Lanzhou University, Lanzhou 730000

(Received August 31, 2012; in final form January 15, 2013)

ABSTRACT

The potential for using the ensemble square root filter data assimilation technique to estimate soil moisture profiles, surface heat fluxes, and the state of the planetary boundary layer (PBL) is explored. An observing system simulation experiment is designed to mimic the assimilation of near-surface soil moisture observations ($\theta^{\rm o}$) and in-situ measurements of 2-m temperature ($T^{\rm o}$), 2-m specific humidity ($Q^{\rm o}$), and 10-m horizontal winds [$\boldsymbol{V}^{\rm o}$=($U^{\rm o}$, $V^{\rm o}$)]. The background forecasts are generated by a one-dimensional coupled land surface-boundary layer model (CLS-BLM) with soil, surface-layer and PBL parameterization schemes similar to those used in the Weather Research and Forecasting (WRF) model. Soil moisture, surface heat fluxes, and the state of the PBL evolve on different characteristic timescales, so the minimum assimilation time intervals required for skillful estimates of each target component are different. Correct estimates of the soil moisture profile are obtained effectively when a 6-h update time interval is used, while skillful estimates of surface fluxes and the PBL state require more frequent updates. The CLS-BLM requires a shorter assimilation time interval to correctly estimate the soil moisture profile than previously indicated by experiments using an off-line land surface model (LSM). Results from assimilating different subsets of observations show that $\theta^{\rm o}$ makes a larger contribution to soil moisture estimates, while $T^{\rm o}$, $\theta^{\rm o}$, and $\boldsymbol{V}^{\rm o}$ are more important for estimates of surface heat fluxes and the PBL state. It is therefore necessary to combine these variables to accurately estimate the states of both the land surface and the PBL. Experimentation with different prescribed observational errors shows that the assimilation system is more sensitive to increases in observational errors than to reductions in observational errors.

Key words: soil moisture, LSM, PBL, ensemble Kalman filter

Citation: Zhang Shuwen, Liu Yanhua, and Zhang Weidong, 2013: Ensemble square root filter assimilation of near-surface soil moisture and reference-level observations into a coupled land surface-boundary layer model. *Acta Meteor. Sinica*, **27**(4), 541–555, doi: 10.1007/s13351-013-0402-6.

1. Introduction

Modeling land-atmosphere interaction has become an important component of research in earth system sciences. The land surface and atmospheric boundary layer compose a tightly coupled system with complicated interactions that modulate weather and climate variability (Boussetta et al., 2008; van den Hurk et al., 2011; Yang and Shen, 2010). Soil moisture plays an important role in these interactions, largely controlling the partitioning of surface-atmosphere energy transfer into fluxes of sensible and latent heat. Realistic simulations of soil moisture are necessary for accurate forecasts using numerical weather prediction (NWP) models. An incorrect initialization of soil moisture may lead to errors in the structure of the planetary boundary layer (PBL) and possibly the development of precipitation, which in turn leads to further errors in soil moisture estimates. Soil moisture initialization not only plays an important role in the prediction of precipitation but also affects the simulation of land surface processes (Svensson et al., 2011;

Supported by the National Key Basic Research and Development (973) Program of China (2013CB430102 and 2012CB956200) and National Natural Science Foundation of China (41075074 and 40775065).
*Corresponding author: zhangsw@lzu.edu.cn.

Koster et al., 2009; Li et al., 2007; Liu et al., 2010).

The representation of the PBL is an important component of any atmospheric model used for studies of weather and climate, regardless of the scale of interest. A more accurate initialization of the PBL structure has been shown to improve short-range forecasts of local thermally-driven flows in a mesoscale model and to positively influence forecasts of cyclogenesis, convective outbreaks, and frontal propagation in a large-scale NWP model (Hacker and Snyder, 2005). The accuracy of the initial PBL structure also influences three-dimensional background fields derived from NWP models for use in process studies (such as studies of air quality and plume dispersion) (Rémy and Bergot, 2009).

The initialization of NWP models is complicated by a relative lack of observations of soil moisture and the state of the PBL. One alternative approach is to use data assimilation techniques within the model framework to retrieve soil moisture profiles or the PBL structure, thus combining direct or indirect observations with model predictions. For example, remote sensing radiometric observations of surface temperature can be used to constrain surface soil moisture and heat fluxes (Reichle et al., 2010) through assimilating microwave brightness temperature or satellite retrieval of near-surface soil moisture into an off-line land surface model (LSM) (Crow and van den Berg, 2010; Walker et al., 2001; Zhang et al., 2005, 2011). In addition to remote sensing measurements, standard ground-based meteorological observations may be assimilated into a single column model (SCM) to estimate the state of the land surface (Hacker and Snyder, 2005; Hacker and Rostkier-Edelstein, 2007; Hess, 2001; Mahfouf, 1991). These ground-based observations are rich, accurate, and often dense. Combinations of different types of observations can also be used to estimate soil moisture or PBL profiles. For example, Seuffert et al. (2003) assimilated both screen-level parameters (e.g., 2-m temperature and specific humidity) and satellite observations of 1.4 GHz brightness temperature to predict sensible heat fluxes, root zones, and near-surface soil moisture. Similarly, Margulis and Entekhabi (2003) assimilated both radiometric surface temperature and screen-level meteorological variables (temperature and humidity) into a model of the atmospheric boundary layer and land surface to estimate surface heat fluxes and the states of the land surface and PBL.

A number of different data assimilation approaches have been adopted for use in constraining the states of the land surface and PBL, from simple direct insertion to complex flow-dependent ensemble-based methods. For example, optimal interpolation (OI) based on statistically derived soil-atmosphere characteristics has been used to correct the soil moisture field (Mahfouf, 1991), while nudging and incremental update methods have been used to improve the initial PBL structure in a mesoscale model (Ruggiero et al., 1996). Other studies have used variational analysis to retrieve soil moisture as the minimum of a cost function that expresses the difference between observations and a specified background state (Seuffert et al., 2003; Zhang et al., 2007). Variational data assimilation techniques have also been used to retrieve soil moisture or PBL profiles (Reichle et al., 2001; Margulis and Entekhabi, 2003). Ensemble Kalman filters (EnKF) and other ensemble-based data assimilation methods have also been widely used to constrain soil moisture and the structure of the PBL (Reichle and McLaughlin, 2002; Hacker and Rostkier-Edelsten, 2007; Tian et al., 2008; Zhang et al., 2010). These ensemble-based methods are able to capture the transient coupling and decoupling of the land surface/PBL with the free atmosphere aloft. Furthermore, they provide flow-dependent estimates of background error covariance, including non-stationary and anisotropic correlations between observations and model background states. This capability has proved very useful for skillful estimation of the PBL state.

Different data assimilation methods have been compared to assess their relative skills for estimating soil moisture. Walker et al. (2001) compared direct insertion with the extended Kalman filter (EKF) method, finding that the latter is superior to the former with a quicker retrieval of the soil moisture profile. Reichle et al. (2002) showed that the EnKF method provided a slightly better estimate of soil moisture than the EKF method. Zhou et al. (2006) compared the performance of EnKF for estimating soil moisture with that of a sequential importance resampling particle filter (which can give exact solutions for large ensemble sizes). They showed that EnKF provides a good approximation for non-normal, nonlinear land surface problems despite its intrinsic assumption of normality.

The studies mentioned above have focused on estimating soil moisture profiles or PBL states separately. In this study, we estimate the soil moisture profile and the PBL state simultaneously by assimilating either near surface soil moisture observations alone or a suite of observations that includes near surface soil moisture, 2-m temperature and humidity, and 10-m horizontal wind. We also evaluate the contributions of each observation to improving the estimated soil moisture profile, surface heat fluxes, and PBL state. The data assimilation is performed using an unperturbed-observation ensemble square root filter (EnSRF) (Whitaker and Hamill, 2002), and the background forecasts are generated using a one-dimensional (1D) coupled land surface-boundary layer model (CLS-BLM). The CLS-BLM contains the same parameterizations as the Weather Research and Forecasting (WRF) model for sub-grid scale processes associated with the PBL, surface layer and land surface (Hacker and Rostkier-Edelstein, 2007; Pagowski et al., 2005). Section 2 briefly introduces the CLS-BLM and the data assimilation algorithm. Section 3 describes the experimental setup. Section 4 documents experimental results, including estimated soil moisture profiles, surface heat fluxes, PBL profiles, and sensitivity to observational errors. Section 5 provides a summary and discussion of the results.

2. Model description and assimilation algorithm

2.1 One-dimensional column model

This study uses the 1D column model CLS-BLM. CLS-BLM has the same physical parameterizations as the WRF model, and is forced externally by tendencies derived from WRF forecasts (e.g., horizontal advection, downward short- and long-wave radiative fluxes and geostrophic wind). The PBL is simulated using the Mellor-Yamada-Janjić scheme (Janjić et al., 2001) and land surface processes are modeled using the Noah LSM (Chen and Dudhia, 2001).

The Noah LSM, which was initially developed in 1993, is directly related to many hydrologic models. The model framework is based on a 1D approach to soil-vegetation-atmosphere interactions that solves the coupled energy and water budgets at the land surface and within the unsaturated zone. Noah LSM is a stand-alone model; it can either be run off-line (driven by atmospheric forcings) or coupled with an atmospheric model. The model state variables include soil moisture and soil temperature in 4 layers, skin temperature (bare soil or vegetation), canopy water storage, and a variety of storage variables related to snow processes (Ek et al., 2003). The Noah LSM has been extensively tested over a wide range of climate regimes. Recent improvements to the model have been documented by Niu et al. (2011) and Yang Zong-Liang et al. (2011).

The column model is adopted because it is easy to use in ensemble-based data assimilation experiments while maintaining the complexity of the original WRF model with respect to the PBL structure and land-atmosphere interactions. The latter feature is essential for this study. Detailed descriptions of the CLS-BLM have been provided by Pagowski et al. (2005) and Hacker and Rostkier-Edelstein (2007).

2.2 Data assimilation algorithm

The ensemble-based data assimilation method generally takes one of two forms: the stochastic filter or the deterministic filter. The latter approach is designed to avoid sampling errors associated with the use of "perturbed observations" in the stochastic filter. A variety of deterministic filters have been proposed, including the ensemble transform Kalman filter, the ensemble adjustment Kalman filter, and EnSRF (Whitaker and Hamill, 2002). Here, we use EnSRF for convenience. Let $\boldsymbol{x}^{\mathrm{b}}$ be an m-dimensional background vector, with $\boldsymbol{x}$ the sum of the ensemble mean $\bar{\boldsymbol{x}}$ and $\boldsymbol{x}'$ the deviation from the ensemble mean (i.e., $\boldsymbol{x} = \bar{\boldsymbol{x}} + \boldsymbol{x}'$). Let $\boldsymbol{y}$ be a p-dimensional vector of observations. In this study, y is composed of observations of soil moisture ($\theta^{\rm o}$), 2-m temperature ($T^{\rm o}$), 2-m humidity ($Q^{\rm o}$), and 10-m horizontal wind ($\boldsymbol{V}^{\rm o}$). The observation operator $\boldsymbol{H}$ transforms model states into the observational space. $\boldsymbol{P}^{\rm b}$ is the $m \times m$-dimensional background error covariance matrix, and R is the $p \times p$-dimensional observational error covariance matrix. The minimum variance estimate of the error in the analysis $\boldsymbol{x}^{\rm a}$ is formulated for EnSRF as

$$\bar{\boldsymbol{x}}^{\rm a} = \bar{\boldsymbol{x}}^{\rm b} + \boldsymbol{K}(\bar{\boldsymbol{y}} - \boldsymbol{H}\bar{\boldsymbol{x}}^{\rm b}), \tag{1}$$

$$\boldsymbol{x}'^{\rm a} = \boldsymbol{x}'^{\rm b} - \tilde{\boldsymbol{K}}\boldsymbol{H}\boldsymbol{x}'^{\rm b}, \tag{2}$$

where $\boldsymbol{K}$ is the Kalman gain given by

$$\boldsymbol{K} = \boldsymbol{P}^{\rm b}\boldsymbol{H}^{\rm T}(\boldsymbol{H}\boldsymbol{P}^{\rm b}\boldsymbol{H}^{\rm T} + \boldsymbol{R})^{-1}, \tag{3}$$

and $\tilde{\boldsymbol{K}}$ is expressed as

$$\tilde{\boldsymbol{K}} = \boldsymbol{P}^{\rm b}\boldsymbol{H}^{\rm T}\left[\left(\sqrt{\boldsymbol{H}\boldsymbol{P}^{\rm b}\boldsymbol{H}^{\rm T} + \boldsymbol{R}}\right)^{-1}\right]^{\rm T} \cdot\left[\sqrt{\boldsymbol{H}\boldsymbol{P}^{\rm b}\boldsymbol{H}^{\rm T} + \boldsymbol{R}} + \sqrt{\boldsymbol{R}}\right]^{-1}. \tag{4}$$

The superscript T stands for the matrix transpose. $\boldsymbol{P}^{\rm b}$ is calculated according to

$$\boldsymbol{P}^{\rm b} = \overline{\boldsymbol{x}'^{\rm b}\boldsymbol{x}'^{\rm bT}} = \frac{1}{n-1}\sum_{i=1}^{i=n}\boldsymbol{x}'^{\rm b}\boldsymbol{x}'^{\rm bT}, \tag{5}$$

where n is the ensemble size.

$\boldsymbol{H}\boldsymbol{P}^{\rm b}\boldsymbol{H}^{\rm T}$ and $\boldsymbol{R}$ reduce to scalars for an individual observation. Assuming $\tilde{\boldsymbol{K}} = \alpha\boldsymbol{K}$ (with a constant α), then a scalar quadratic equation for α can be obtained with the solution

$$\alpha = \left(1 + \sqrt{\frac{\boldsymbol{R}}{\boldsymbol{H}\boldsymbol{P}^{\rm b}\boldsymbol{H}^{\rm T} + \boldsymbol{R}}}\right)^{-1}.$$

Detailed descriptions of EnSRF have been set forth by Whitake and Hamill (2002) and Yang Yi et al. (2011).

3. Experimental setup

An observing system simulation experiment (OSSE) is designed to reduce the interference of model errors. The true model states are assumed to be exactly known and the observations are generated by adding known random errors to the true model states. An OSSE is carried out between 13 and 29 August 2003 at a single point with sparse, flat land cover located at 37.6°N, 96.7°W. The weather during the testing period was fine, with no large advection processes or precipitation.

The initial PBL conditions and surface radiation are drawn from the WRF climatology, with the geostrophic wind forcing set to 3 and –9 m s^{-1}. The model is configured to run on 60 non-uniform vertical levels with the lowest model level at 45 m above ground level. $T^{\rm o}$, $Q^{\rm o}$, and $\boldsymbol{V}^{\rm o}$ are accordingly not predicted by the model; rather, they are calculated according to Monin-Obukhov similarity. The observation operator $\boldsymbol{H}$ is not explicitly needed (Reichle et al., 2002a). Forecasts were initiated at 0300 UTC with a horizontal spacing of 4 km and a time step of 60 s.

Rather than using the 4-layer discretization of the soil column in the original Noah LSM, we divide the 1-m soil column into 50 layers. The thickness of the topmost layer is 1 cm, and the thickness of all remaining layers is 2 cm. This approach increases the accuracy of the estimates and provides a fair comparison with previous data assimilation results using an offline LSM. Additional details of the data assimilation environment are listed in Table 1.

We adopt a relatively large ensemble of 80 members when comparing the OSSE with model states. The initial ensemble of PBL state profiles is constructed by first randomly selecting two forecasts from WRF real-time forecasts at a 12-h lead time, then av-

Table 1. Soil parameters and assimilation conditions used in the OSSE

Soil parameters		Assimilation conditions	
Soil type	Clay loam	Depth	100 cm
Soil moisture at saturation	0.476 m^3 m^{-3}	Number of soil layers	50
Exponent b	5.33	Initial soil moisture	0.25 m^3 m^{-3}
Matric potential at saturation	–63.0 cm	Initial guess	0.20 m^3 m^{-3}
Hydraulic conductivity at saturation	25 cm day^{-1}	Observations	$\theta^{\rm o}$, $T^{\rm o}$, $Q^{\rm o}$, and $\boldsymbol{V}^{\rm o}$

eraging them with weighting coefficients μ and $1-\mu$ (with μ sampled from a uniform distribution U[0,1]). The initial soil moisture is intentionally poor to account for the poor knowledge of actual soil moisture, with an initial value of 0.2 m^3 m^{-3} (approximately half the saturated soil moisture). The initial ensemble of soil moisture is generated by adding a normally-distributed random field with a mean of zero and a standard deviation (STD) of 0.05 m^3 m^{-3} to the initial guess. Observational errors are represented by temporally uncorrelated additive Gaussian noise with STDs of 1 K for $T^{\rm o}$, 0.8 g kg^{-1} for $Q^{\rm o}$, 1.4 m s^{-1} for $|\boldsymbol{V}^{\rm o}|$, and 0.03 m^3 m^{-3} for $\theta^{\rm o}$. The assimilation time interval ranges from 24 to 0.5 h. Further details are provided in Section 4.

We use the mean absolute error (MAE) to quantitatively compare estimates by different approaches. MAE is defined for an estimate f and its true value $f^{\rm t}$ as $\mathrm{MAE} = T_{\rm n}^{-1}\sum_{i=1}^{T_{\rm n}}|f - f^t|$, where $T_{\rm n}$ is the number of the verification times. The benefit of the estimate relative to an open loop simulation (i.e., an 80-member ensemble forecast without data assimilation with the same ensemble members and initial conditions as the assimilation system) is estimated using the relative MAE reduction (RMR)

$$\mathrm{RMR} = (\mathrm{MAE}^{\rm OP} - \mathrm{MAE})/\mathrm{MAE}^{\rm OP} \times 100\%,$$

where $\mathrm{MAE}^{\rm OP}$ is the mean absolute error from the open loop simulation.

4. Results and analysis

4.1 Results from assimilating only the near-surface soil moisture observations

Previous assimilation studies have used off-line meteorological forcing data to drive the LSM without any coupling between the land surface and the atmosphere aloft. Here, averaged soil moisture estimates are calculated in four representative layers with the midpoint depths of 2.5, 10, 30, and 72 cm from the surface to facilitate comparisons and reduce random errors. Averaged soil moisture conditions obtained by assimilating the near-surface soil moisture observations at time intervals of 6 and 24 h are then compared with the open loop simulations (Fig. 1). The 6-h time interval is chosen to match the 6-h sampling interval typical of synoptic station measurements. Estimates of soil moisture based on 6-h assimilation interval are more accurate than those based on 24-h assimilation interval in all layers. The accuracy of soil moisture estimates decreases gradually with increases in soil depth. This is consistent with previous data assimilation results using off-line LSM simulations (Zhang et al., 2005; Walker et al., 2001). Assimilation of $\theta^{\rm o}$ on 6-h interval is performed consistently throughout the assimilation period and provides a means of quickly retrieving the true soil moisture in each layer. By contrast, assimilation of $\theta^{\rm o}$ on 24-h interval performs well during the early stages but its impact gradually diminishes with time (particularly in the deeper layers). In contrast with previous results using off-line LSM simulations (Zhang et al., 2005; Walker et al., 2001), our simulations of soil moisture take longer to converge to a skillful estimate. Moreover, the accuracy of the soil moisture estimates drops relative to that of the off-line simulations when similar experimental settings are used.

Unlike previous data assimilation experiments using off-line LSM simulations, CLS-BLM can be used to estimate surface heat fluxes and the PBL states in addition to the soil moisture profile. Sensitivity tests indicate that accurate estimation of surface heat fluxes will require a much shorter assimilation time interval because surface heat fluxes change very quickly after sunrise. This rapid adjustment of surface heat fluxes to the diurnal cycle is completely different from soil water transfer processes, which occur on a much slower timescale. Moreover, unlike the background states, heat fluxes are diagnosed rather than updated with each assimilation cycle. Observations only affect heat fluxes indirectly, with gradual adjustments to improved model states. When the update frequency is once per day, the estimated fluxes are almost identical to those calculated by the open loop simulation. The estimated fluxes are still not significantly better than the open loop estimates when the assimilation time interval is reduced to 6 h. The estimated fluxes are only substantially improved when the assimilation time interval is reduced to 0.5 h (Fig. 2). This improvement is especially apparent in the sensible heat flux. Data assimilation improves the estimates more slowly at later integration times due to the convergence of

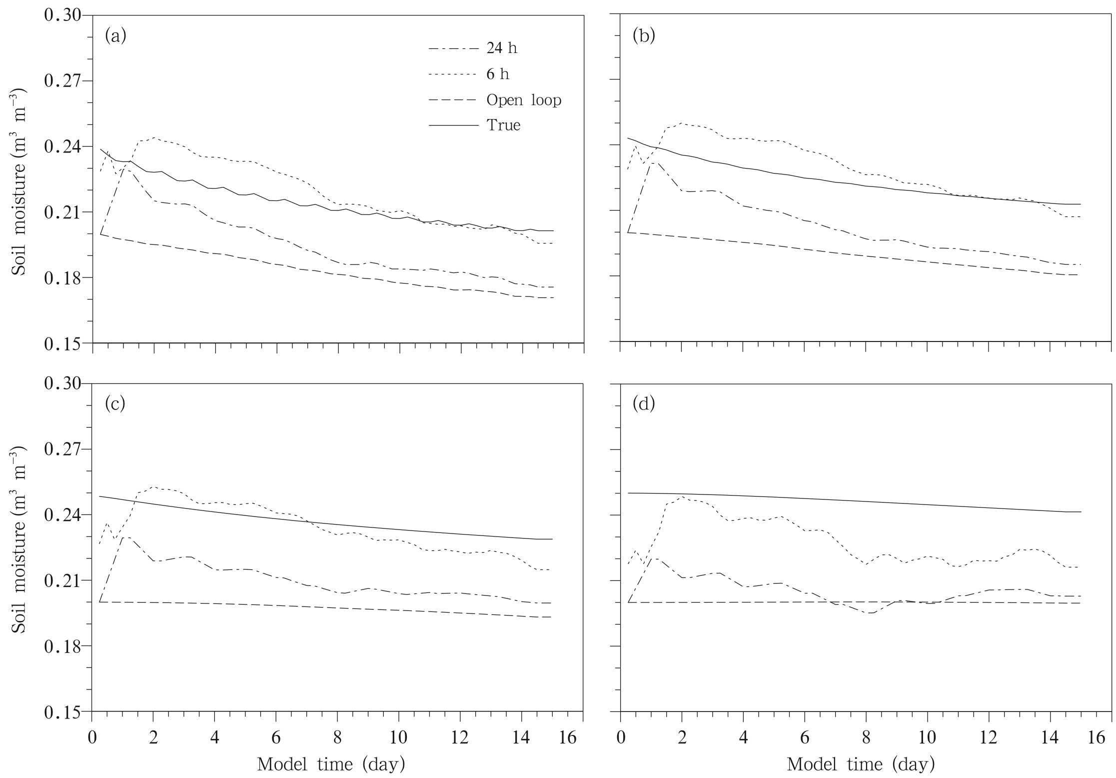

Fig. 1. Averaged soil moisture estimates in four representative layers with the midpoint depths of (a) 2.5, (b) 10, (c) 30, and (d) 72 cm from the surface using assimilation time intervals of 6 h (dotted line) and 24 h (dash-dotted line). The true soil moisture (solid line) and open loop estimate (dashed line) are also shown.

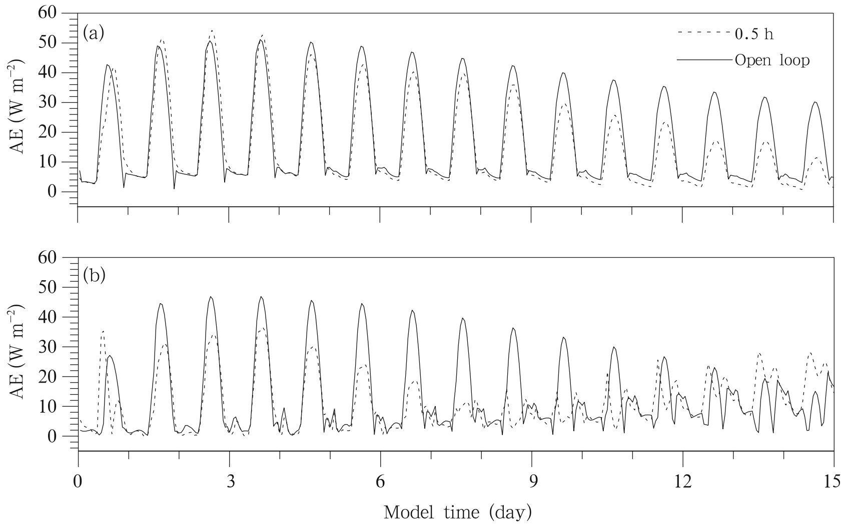

Fig. 2. Absolute errors (AE) in (a) latent heat flux and (b) sensible heat flux relative to the true fluxes when only θ^{o} is assimilated at an assimilation time interval of 0.5 h (dashed line). The absolute errors for the open loop fluxes are also shown (solid line).

the ensemble members, with data assimilation estimates of the sensible heat flux at the longest integration times even slightly worse than the open loop results. The maximum absolute error of the estimated fluxes during the 15-day data assimilation period always appears at approximately midday, between 1000 and 1600 LT (local time). On average, the largest reduction in MAE lies in the estimates of surface sensible heat flux. The RMR for sensible heat flux relative to the open loop flux is 44.5% at midday and 25.4% for the whole day (Table 2). The estimated latent heat flux is progressively improved throughout the 15-day data assimilation period (Fig. 2a). This improvement may be attributable to simultaneous improvements in soil moisture estimates, which control the surface evaporation when the soil is very dry (as it is in this case). The next subsection will show that estimates of the surface heat fluxes (especially latent heat flux) can be improved still further when additional surface observations are assimilated.

This data assimilation experiment offers no improvement over the open loop ensemble for estimated profiles of temperature, humidity, and wind in the PBL, even when the assimilation time interval is reduced to 0.5 h (figure omitted). This result suggests that assimilation of θ^{o} alone offers little potential for improving estimates of the PBL state.

4.2 Results from assimilating all observations

This subsection describes an extension of the experiments discussed in Subsection 4.1, in which all observations (i.e., θ^{o}, T^{o}, Q^{o}, and $\boldsymbol{V}^{\mathrm{o}}$) are assimilated into CLS-BLM. Consistency with the observational frequency (i.e., once every 6 h) at the conventional meteorological observation stations is assured by initially adopting a 6-h assimilation time interval. A 0.5-h time interval is also used for comparison with the results reported in Subsection 4.1. Figure 3 shows average soil moisture estimates in the four specified layers with an assimilation time interval of 6 h. The results obtained from assimilating only θ^{o} and the open loop simulations of soil moisture are included for comparison. The estimates of soil moisture when all observations are assimilated are almost the same as those obtained when only θ^{o} is assimilated, although the former are slightly more consistent with the true soil moisture. This result implies that assimilating T^{o}, Q^{o}, and $\boldsymbol{V}^{\mathrm{o}}$ does not substantially improve estimates of soil moisture, confirming that the largest contribution to improvements in simulated soil moisture profiles comes from assimilating θ^{o}. This inference is supported by an additional experiment in which only T^{o}, Q^{o}, and $\boldsymbol{V}^{\mathrm{o}}$ are assimilated. In this case, the estimates are only slightly better than the open loop soil moisture (figure omitted).

Similar to the experiment in which only θ^{o} is assimilated, the estimated surface heat fluxes are not improved relative to the open loop fluxes if the assimilation time interval is set to 6 h (figure omitted). By contrast, reducing the assimilation time interval to 0.5 h results in substantial improvements in the estimated surface heat fluxes relative to both the θ^{o}-only and open loop experiments, particularly around midday (Fig. 4). Assimilating surface observations leads to particularly large improvements (i.e., reductions in MAE) in the latent heat flux, with an RMR of 57.8% at midday and an RMR of 46.2% for the whole day (Table 2). Estimates of sensible heat flux are also improved relative to the experiment in which only θ^{o} is assimilated (c.f., Fig. 2).

The improvement in PBL state profiles after assimilating all observations is investigated using the MAE over the full 15-day period because the PBL state changes often. Four representative times (0800,

Table 2. MAE and RMR of estimated surface fluxes over the 15-day assimilation period

	Only assimilating θ^{o}		Assimilating θ^{o}, T^{o}, Q^{o}, and $\boldsymbol{V}^{\mathrm{o}}$		Open loop
	MAE	RMR	MAE	RMR	MAE
LHF (midday)	35.3	11.3%	16.8	57.8%	39.8
LHF (whole day)	24.3	8.6%	14.3	46.2%	26.6
SHF (midday)	17.2	44.5%	14.4	53.5%	31.0
SHF (whole day)	17.3	25.4%	16.2	30.2%	23.2

1400, 2000, and 0200 LT) are chosen to reflect the diurnal evolution of the PBL structure.

Unlike when only θ^o is assimilated, estimates of potential temperature profiles are improved relative to the open loop simulations when all observations are assimilated on 6-h interval (Fig. 5). These estimates can be improved still further (especially at low levels) by reducing the assimilation time interval to 0.5 h. Assimilating all observations even yields improvements throughout the profile at night, when the coupling be-

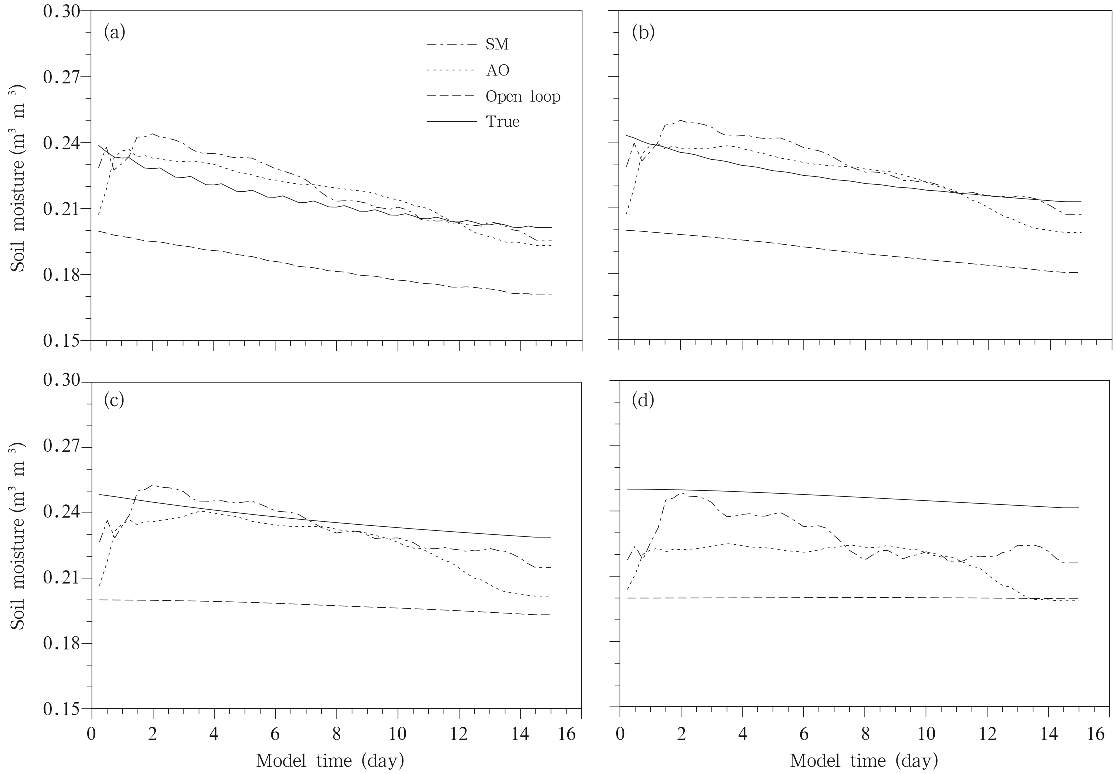

Fig. 3. As in Fig. 1, but for experiments assimilating different sets of observations. SM indicates that only θ^o is assimilated while AO indicates that all observations are assimilated.

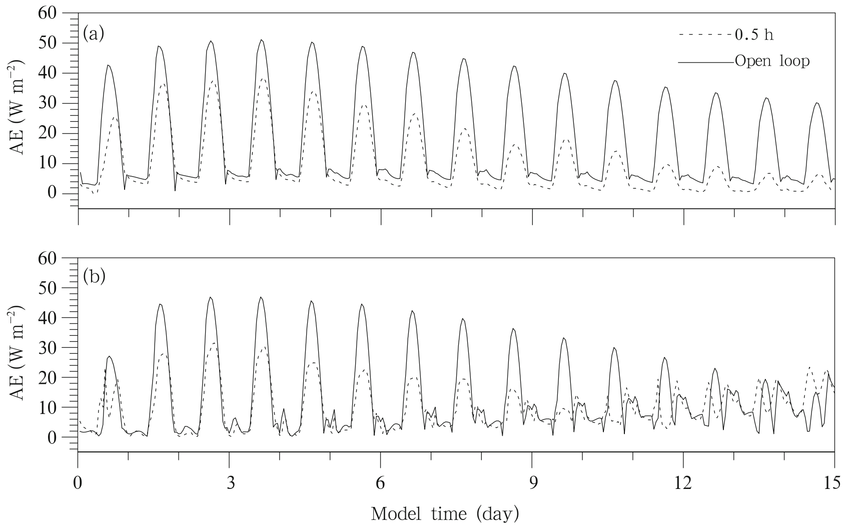

Fig. 4. As in Fig. 2, but for the experiment in which all observations are assimilated.

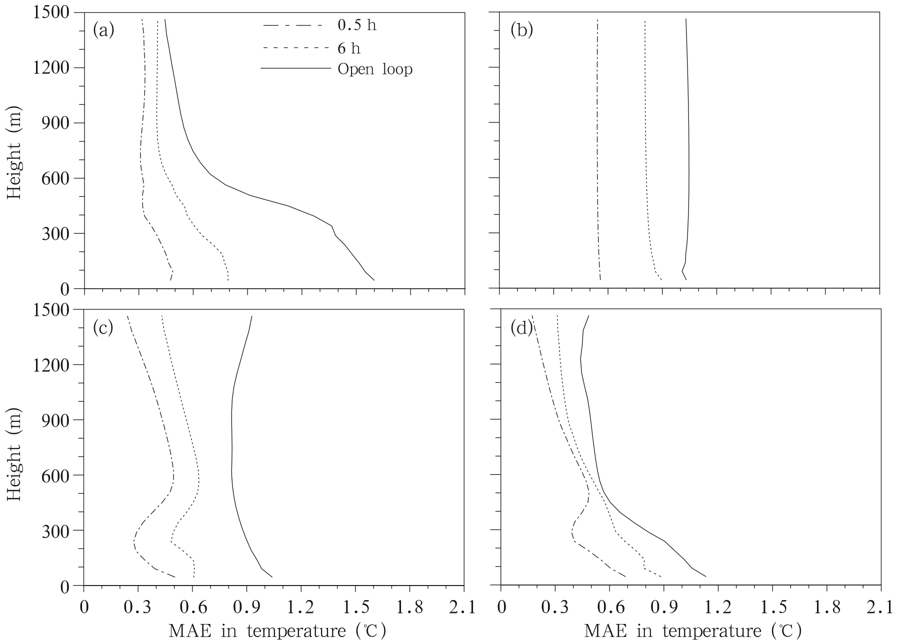

Fig. 5. Vertical profiles of MAE in potential temperature at (a) 0800, (b) 1400, (c) 2000, and (d) 0200 LT averaged over the 15-day analysis period for experiments in which all observations are assimilated at intervals of 0.5 h (dash-dotted line) and 6 h (dotted line). Results from the open loop simulation (solid line) are also shown.

tween the surface layer and the upper residual layer becomes small (Figs. 5c and 5d). Turbulent fluxes weaken at night, while the model lacks the advective tendency that forces ensemble members to respond to external changes. Accordingly, the ensemble mean changes little with time. On the other hand, useful information about the daytime state propagates into the residual layer at nighttime. These two factors may both contribute to the skill in estimates of nighttime PBL temperature profiles.

The improvements in specific humidity profiles (Fig. 6) are similar to the improvements in temperature profiles. The MAE throughout the humidity profile is reduced relative to the open loop simulation at all four representative times. Furthermore, the MAE is smaller when the assimilation time interval is shorter. Unlike the temperature profiles, the MAE in specific humidity changes smoothly in the vertical direction with similar improvements at all levels. The reason for this is that the ensemble means and increments of humidity all change smoothly with height.

Improvements in the profile of horizontal wind (Fig. 7) are also similar to improvements in both temperature and humidity profiles. One difference is that shorter assimilation time interval provides greater improvements in estimates of horizontal wind than in estimates of temperature or specific humidity in the afternoon and at night. This difference arises because the state of PBL winds often changes quickly in the afternoon while the state at night relaxes gradually toward the geostrophic wind, resulting in a rapid growth of internal errors. Increasing the frequency of updates helps to prevent error growth and improve the estimates.

4.3 Sensitivity to specified observational errors

The results of the previous two subsections are all based on the assumption of specified observational errors (STDs of 1 K for T^{o}, 0.8 g kg^{-1} for Q^{o}, 1.4 m s^{-1} for $|\boldsymbol{V}^{o}|$, and 0.03 m^{3} m^{-3} for θ^{o}). These experiments are henceforth referred to as the base-case assimilation experiments. Observational errors are likely to change in real applications. To investigate the potential impacts of changes in observational errors on the skill of soil moisture, heat flux and PBL state profile esti-

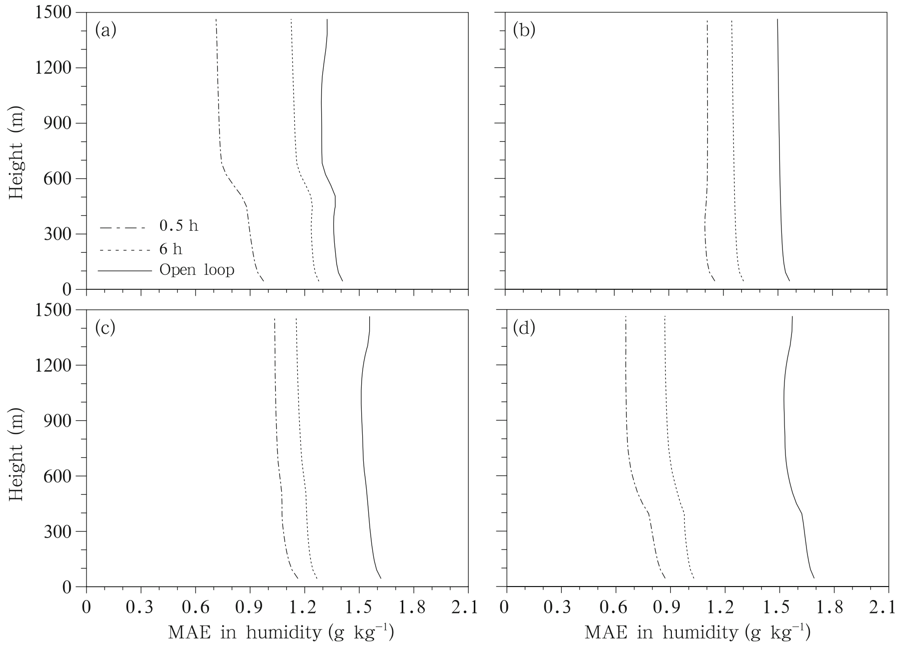

Fig. 6. As in Fig. 5, but for specific humidity.

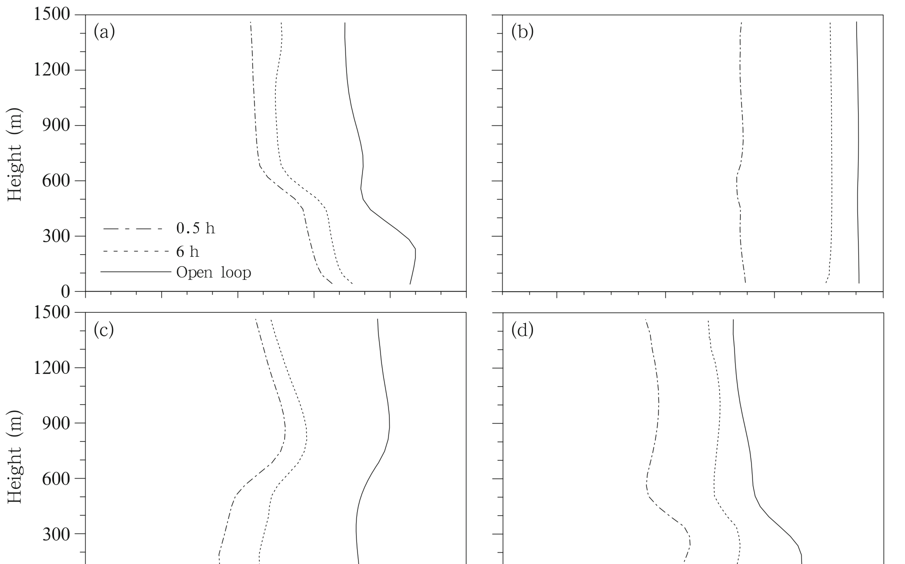

Fig. 7. As in Fig. 5, but for horizontal wind speed.

mates, two additional data assimilation experiments are conducted. Observational errors are reduced by a factor of 5 in the first experiment (errors in soil moisture are reduced by a factor of 3 to 0.01 m^3 m^{-3}) and increased by a factor of 5 in the second (errors in soil moisture are increased to 0.05 m^3 m^{-3}).

Estimates of soil moisture profile are severely affected when observational errors are increased, even becoming worse than open loop soil moisture at the two deepest layers after a few days (figure omitted). The effects on estimates of the PBL potential temperature profile are dependent on the time of day (Fig. 8). Improvements relative to the open loop simulation are only evident at 1400 LT. These improvements are related to increased ensemble spread at the observing location due to strong turbulent mixing. The observed temperature in the surface layer is highly correlated with temperature in the well-mixed PBL at this time, resulting in large analysis increments. At all other times (i.e., 0800, 2000, and 0200 LT), the state estimates are slightly worse than the open loop simulation at some levels because of the smallness of the analysis increments. The specific humidity profile (Fig. 9) is only worse than the open loop simulation at a few levels at 0800 LT. At all other times, the data assimilation still adds skill due to slow reductions in the ensemble spread. Unlike temperature and humidity, the profile of horizontal wind is still improved relative to the open loop at all levels and all times, although the degree of improvement is different at different times (Fig. 10). The winds are always forced by the geostrophic wind distribution, so even at night the ensemble spread does not often become small.

When the observational errors are specified to be smaller than in the base-case experiment (i.e., a closer fit to the true state), estimates of soil moisture near the observing location are generally improved relative to the base-case experiment. The one exception is in the two deepest layers, where small ensemble spread leads to slightly worse estimates at later times (figure omitted). All three PBL state profiles improve consistently at all levels relative to the base-case experiment at 0800, 1400, and 2000 LT. At upper levels at 0200 LT, temperature estimates do not improve with reductions in observational errors, while wind estimates even become slightly worse than the base-case (although the degree of deterioration is relatively weak). These results indicate that the EnSRF assimilation system is very sensitive to increases in observational errors and moderately sensitive to decreases in observational errors.

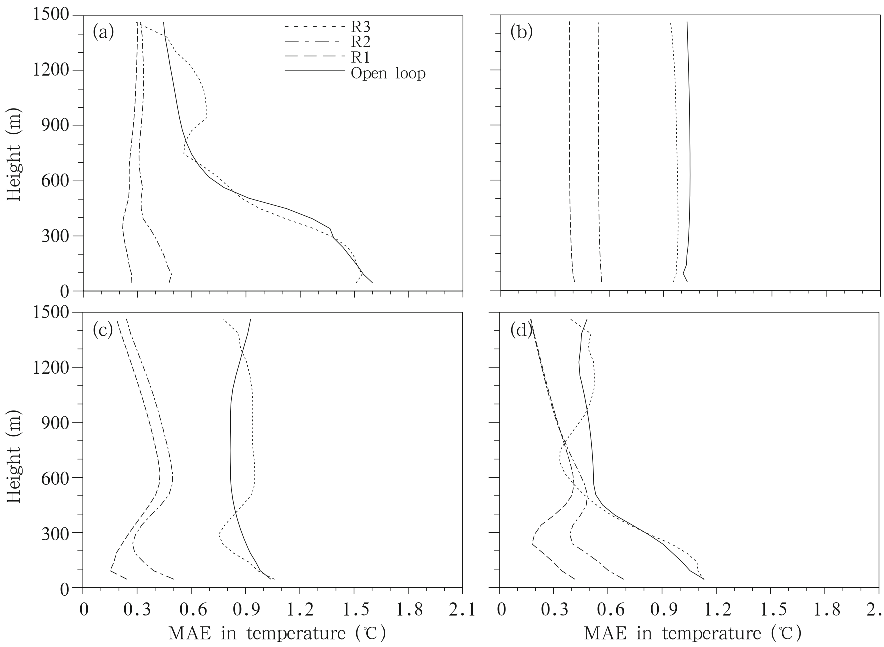

Fig. 8. Profiles of MAE in temperature estimates at (a) 0800, (b) 1400, (c) 2000, and (d) 0200 LT using different ranges of observational errors. R1 (dashed line) indicates the experiment with smaller observational errors, R2 (dash-dotted line) indicates the base-case experiment, and R3 (dotted line) indicates the experiment with larger observational errors.

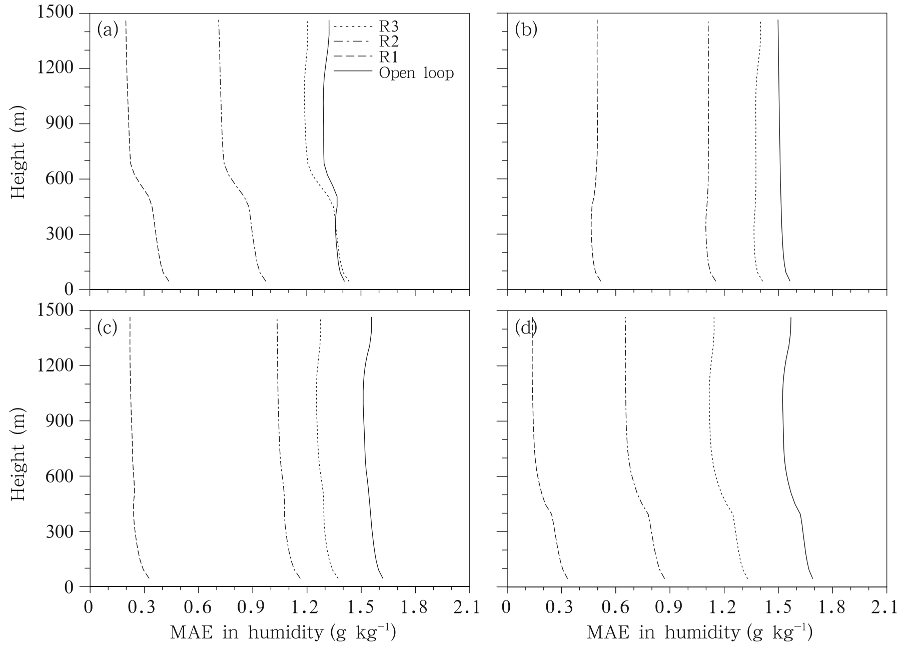

Fig. 9. As in Fig. 8, but for specific humidity.

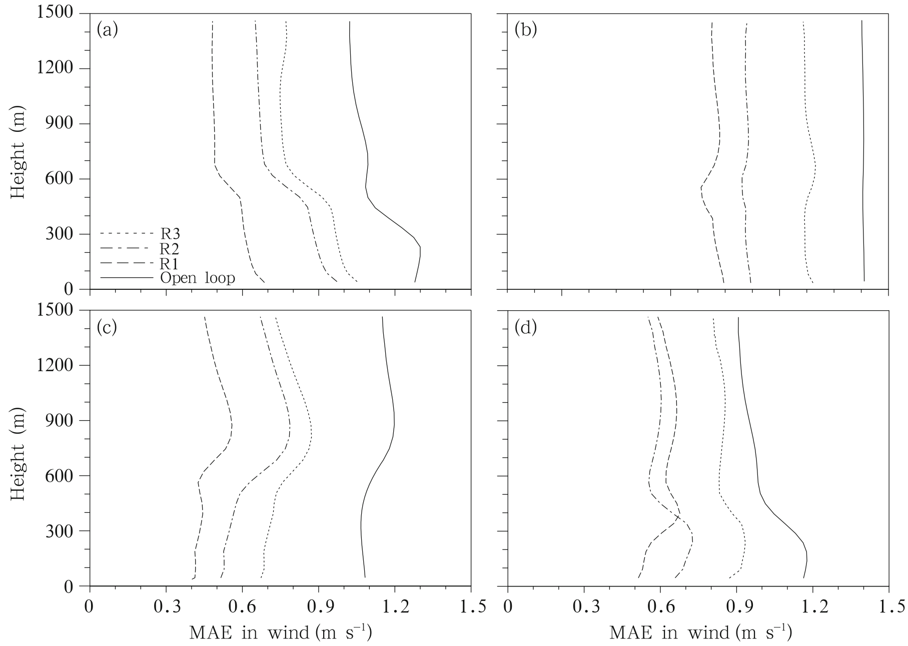

Fig. 10. As in Fig. 8, but for horizontal wind speed.

5. Summary and discussion

An observing system simulation experiment (OSSE) has been designed and carried out to show the ability of data assimilation to improve estimates of the soil moisture profile, surface heat fluxes, and PBL states when different sets of observations are assimilated. The coupled land surface-boundary layer model is a 1D column model with the same PBL parameterization as the WRF model, forced by WRF forecasts. The EnSRF data assimilation algorithm is used. EnSRF does not require perturbed observations and can provide estimates of anisotropic and inhomogeneous background error covariance. The skill of the data assimilation experiments is quantified on a one point scale relative to the true state and open loop simulations during 13–29 August 2003.

The soil moisture profile can be retrieved effectively when a 6-h assimilation time interval is used, regardless of whether the assimilation includes only soil moisture (θ^{o}) or all observations (T^{o}, Q^{o}, $\boldsymbol{V}^{\mathrm{o}}$, and θ^{o}). The results are slightly better when all observations are assimilated. Soil moisture in the deep layer cannot be correctly estimated when the assimilation time interval is 24 h. This result contrasts with the results of previous studies that used an off-line land surface model driven by external atmospheric forcings.

Surface heat flux estimates are only substantially improved when the assimilation time interval is further reduced from 6 to 0.5 h, with large reductions in MAE between 1000 and 1600 LT. Surface heat fluxes are better when all observations are assimilated than when only θ^{o} is assimilated. These improvements are especially significant for estimates of the latent heat flux at midday.

Estimates of the PBL state cannot be substantially improved relative to the open loop simulation when only θ^{o} is assimilated, even when the assimilation time interval is reduced to 0.5 h. However, assimilating all observations yields improvements in estimates of the PBL state when the assimilation time interval is 6 h, even at night. Reducing the assimilation time interval to 0.5 h further enhances the accuracy of estimates of the PBL state. Differences in the timescales of related processes lead to differences in the minimum assimilation time interval required to improve estimates of soil moisture profiles, surface heat fluxes, and PBL state profiles.

Sensitivity tests show that increasing the specified observational errors worsens estimates of deep-layer soil moisture and the vertical structure of the PBL temperature profile. These estimates are sensitive to increases in observational errors. Reducing the observational errors generally improves the skill of state estimates near the observing locations, although temperature and wind estimates at upper levels (which are farther from the observing locations) may worsen slightly at night. This adverse effect is not serious.

Although these results are promising, further research is needed to successfully apply these assimilation techniques to real observations. The experimental case presented here is limited to a short period in summer, when horizontal advection is weak and the coupling between the surface and the atmosphere aloft is strong. A column model with no advection may not provide a correct background forecast for use in EnSRF during winter, when strong advection exists. The most difficult issue when assimilating real observational data is how to address model errors from different sources (e.g., parameterizations of vertical diffusion and turbulent mixing in the PBL, uncertainties in modeling the strength of land-atmospheric coupling and feedbacks on local and regional scales, etc.). Model errors may be further complicated when the land surface is heterogeneous. Although EnSRF potentially has the ability to simultaneously estimate both model states and model uncertainties using the state augmentation technique, initial experimentation along these lines has not been encouraging. More research is still needed to evaluate this approach due to the highly nonlinear relationships between model states and parameters.

Acknowledgments. Three anonymous reviewers are thanked for insightful comments and suggestions. The language editor for this manuscript is Dr. Jonathon S. Wright.

REFERENCES

Boussetta, S., T. Koike., T. Graf, et al., 2008: Development of a coupled land-atmosphere satellite data assimilation system for improved local atmospheric simulations. *Remote Sens. Environ.*, **112**(3), 720–734.

Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface-hydrology model with the Penn state-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.*, **129**, 569–585.

Crow, W. T., and M. J. van den Berg, 2010: An improved approach for estimating observation and model error parameters for soil moisture data assimilation. *Water Resour. Res.*, **46**, W12519, doi: 10.1029/2010WR009402.

Ek, M., K. Mitchell, L. Yin, at al., 2003: Implementation of Noah land-surface model advances in the NCEP operational mesoscale Eta model. *J. Geophys. Res.*, **108**, D22, doi: 10.1029/2002JD003296.

Hacker, J. P., and C. Snyder, 2005: Ensemble Kalman filter assimilation of fixed screen-height observations in a parameterized PBL. *Mon. Wea. Rev.*, **133**, 3260–3275.

——, and D. Rostkier-Edelstein, 2007: PBL state estimation with surface observations, a column model, and an ensemble filter. *Mon. Wea. Rev.*, **135**, 2958–2972.

Hess, R., 2001: Assimilation of screen-level observations by variational soil moisture. *Meteor. Atmos. Phys.*, **77**, 145–154.

Janjić, Z. I., J. P. Gerrity, and S. Nickovic, 2001: An alternative approach to nonhydrostatic modeling. *Mon. Wea. Rev.*, **129**, 1164–1178.

Koster, R. D., Guo Zhichang, Yang Rongqian, et al., 2009: On the nature of soil moisture in land surface models. *J. Climate*, **22**(16), 4322–4335.

Li Xin, Huang Chunlin, Che Tao, et al., 2007: Development of a Chinese land data assimilation system: Its progress and prospects. *Progress in Natural Science*, **17**(8), 881–892.

Liu Jingmiao, Ding Yuguo, Zhou Xiuji, et al., 2010: A parameterization scheme for regional average runoff over heterogeneous land surface under climatic rainfall forcing. *Acta Meteor. Sinica*, **24**(1), 116–122.

Mahfouf, J. F., 1991: Analysis of soil moisture from near-surface parameters: A feasibility study. *J. Appl. Meteor.*, **30**, 1534–1547.

Margulis, S. A., and D. Entekhabi, 2003: Variational assimilation of radiometric surface temperature and screen-level micrometeorology into a model of the atmospheric boundary layer and land surface. *Mon. Wea. Rev.*, **131**, 1272–1288.

Niu, G. -Y., Z. -L. Yang, K. E. Mitchell, et al., 2011: The community Noah land surface model with multi parameterization options (Noah-MP). Part I: Model description and evaluation with local-scale measurements. *J. Geophys. Res.*, **116**, D12109, doi: 10.1029/2010JD015139.

Pagowski, M., J. Hacker, and J. -W. Bao, 2005: Behavior of WRF PBL schemes and land-surface models in 1D simulations during BAMEX. Joint WRF/MM5 Users' Workshop, Boulder CO, July 2005, available at http://www.mmm.ucar.edu/wrf/users/workshops/WS2005/WorkshopPapers.htm.

Reichle, R., D. B. McLaughlin, and D. Entekhabi, 2001: Variational data assimilation of microwave radiobrightness observations for land surface hydrologic applications. *IEEE Trans. Geosci. Remote Sens.*, **39**, 1708–1718.

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

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

——, S. V. Kumar, S. P. P. Mahanama, et al., 2010: Assimilation of satellite-derived skin temperature observations into land surface models. *J. Hydrometeor.*, **3**, 1103–1122.

Rémy, S., and T. Bergot, 2009: Ensemble Kalman filter data assimilation in a 1D numerical model used for fog forecasting. *Mon. Wea. Rev.*, **138**(5), 1792–1810.

Ruggiero, F. H., K. D. Sashegyi, R. V. Madala, et al., 1996: The use of surface observations in four-dimensional data assimilation using a mesoscale model. *Mon. Wea. Rev.*, **124**, 1018–1033.

Seuffert, G., H. Wilker, P. Viterbo, et al., 2003: Soil moisture analysis combining screen-level parameters and microwave brightness temperature: A test with field data. *Geophys. Res. Lett.*, **30**(10), 1498, doi: 10.1029/2003GL017128.

Svensson, G., A. A. M. Holtslag, V. Kumar, et al., 2011: Evaluation of the diurnal cycle in the atmospheric boundary layer over land as represented by a variety of single-column models: The second GABLS Experiment. *Bound.-Layer Meteor.*, **140**(2), 177–206.

Tian Xiangjun, Xie Zhenghui, and A. Dai, 2008: A land surface soil moisture data assimilation system based on the dual-UKF method and the Community Land Model. *J. Geophys. Res.*, **113**, D14127, doi: 10.1029/2007JD009650.

van den Hurk, B., M. Best, P. Dirmeyer, et al., 2011: Acceleration of land surface model development over a decade of GLASS. *Bull. Amer. Meteor. Soc.*, **92**(12), 1593–1600.

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.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924.

Yang Junli and Shen Xueshun, 2010: The construction of SCM in GRAPES and its applications in two field experiment simulations. *Adv. Atmos. Sci.*, **28**(3), 534–550.

Yang Yi, Gong Zhongqiang, Wang Jinyan, et al., 2011: Time-expanded sampling approach for ensemble Kalman filter: Experiment assimilation of simulated soundings. *Acta Meteor. Sinica*, **25**(5), 558–567.

Yang Zong-Liang, Niu Guo-Yue, K. E. Mitchell, et al., 2011: The community Noah land surface model with multi parameterization options (Noah-MP). Part II: Evaluation over global river basins. *J. Geophys. Res.*, **116**, D12110, doi: 10.1029/2010JD015140.

Zhang Shuwen, Li Haorui, Zhang Weidong, et al., 2005: Estimating the soil moisture profile by assimilating near-surface observations with the Ensemble Kalman Filter (EnKF). *Adv. Atmos. Sci.*, **22**(6), 936–945.

——, Zhang Weidong, and Qiu Chongjian, 2007: A variational method for estimating near-surface soil moisture and surface heat fluxes. *Acta Meteor. Sinica*, **21**(4), 476–486.

——, X. B. Zeng, Zhang Weidong, et al., 2010: Revising the ensemble-based Kalman filter covariance for the retrieval of deep-layer soil moisture. *J. Hydrometeor.*, **11**(1), 219–227.

——, Li Deqin, and Qiu Chongjian, 2011: A multimodel ensemble-based Kalman filter for the retrieval of soil moisture profiles. *Adv. Atmos. Sci.*, **28**(1), 195–206.

Zhou Yuhua, D. McLaughlin, and D. Entekhabi, 2006: Assessing the performance of the ensemble Kalman filter for land surface data assimilation. *Mon. Wea. Rev.*, **134**(8), 2128–2142.