22.1. 4DVAR Assimilation of Ground Temperature for the Estimation of Soil Moisture and Temperature

Diandong Ren and Ming Xue*
School of Meteorology and Center for Analysis and Prediction of Storms
University of Oklahoma
Norman OK 73019

1. Introduction

Land surface models (LSMs or called soil-vegetation models) deal with the exchanges of moisture and thermal energy between the land surface and the atmosphere. Many studies have demonstrated the sensitivity of the surface energy budget and atmospheric fields to the formulation of land-surface processes, on virtually all spatial and temporal scales (see, e.g., Henderson-Sellers *et al*. 1996).

Our ability to accurately describe the land surface processes and to forecast the time evolution is severely hindered by process uncertainties and limited availability of appropriate data. There exists a continuing debate about the ways to properly harness the observations to improve LSMs performance (e.g., GCIP3 report, 2002, http://ecpc.ucsd.edu/gcip/2002JGRpapergcipwebs/200210.WEBSJGRsubmit.html). The very limited availability of direct soil temperature and moisture measurements is obviously not adequate for routine model initialization. Recent development in *in situ* measurements of surface fluxes, soil temperature and soil moisture content in the Oklahoma Atmospheric Surface-layer Instrumentation System (OASIS, Brotzge 2000) gives us a unique opportunity to verify LSMs and to test methods for estimating soil state variables as well as uncertain parameters in LSMs.

On the premise that the quality of LSMs output is closely related to the meteorological forcing that drives it, the Land Data Assimilation System (LDAS, Schaake *et al.* 2002; Nijssen *et al.* 2001) performs continuous forced run of a land surface model to assimilate atmospheric information to improve descriptions of land surface conditions. In a similar spirit, the Antecedent Precipitation Index (API) method (Linsley *et al.* 1949) uses weighted summation of past daily precipitation amounts to estimate the content of soil moisture. API is often used in catchment hydrology for studying runoff and infiltration distribution. API has also been used to initialize soil moisture content in LSMs for numerical weather prediction (NWP) and simulation (e.g., Ziegler *et al.* 1997). Both approaches enjoy the advantage of easy implementation and performance. These retrospective approaches are generally case sensitive to changes in location and time period, however. The retrieved "optimal" soil moistures often deviate from the truth (Jones *et al.* 2003), especially with the API method for which much tuning of the weighting parameter is usually needed for specific locations. For the LDAS approach, the moisture analyses depend very much on the skill of the land model used (Henderson-Sellers *et al.* 2002; Qu *et al*. 1998).

During the past two decades, various inverse schemes were also proposed to indirectly infer land surface prognostic variables and model parameters (Mahfouf 1991; Calvet *et al*. 1998; Margulis and Entekhabi 2001; Xu and Zhou 2003). Mahfouf (1991) pioneered the attempt of retrieving top layer moisture from observations of screen-level air temperature and relative humidity. Using the Interactions Soil Biosphere Atmosphere (ISBA, Noilhan and Planton 1989, NP89 henceforth) land surface model, he obtained positive results for a crop area and clear sky conditions. He described two possible approaches: a variational algorithm where a cost function is minimized over an assimilation period, and a sequential assimilation scheme that consists of a set of predictions and static correlation of soil moisture. He validated both methods against *in situ* data collected during a field experiment, using a one-column model to represent the interactions between surface processes and the planetary boundary layer structure. Calvet et al. (1998) tried the inverse estimation of the bulk soil moisture content using surface variables. They argued that knowing the atmospheric forcing (especially precipitation) and four to five surface soil moisture over two weeks are adequate to retrieve the bulk soil moisture by inverting ISBA scheme. They realized a strong relationship between the deeper layer soil moisture and surface soil moisture, especially when the vegetation are in full growth, it is therefore feasible to infer the bulk soil moisture by minimizing error in the prediction of surface soil moisture. Xu and Zhou (2003) discuss a linear regression method for retrieving bulk soil moisture contents from soil temperature profile measurements, based on the

* Corresponding author address: Dr. Ming Xue
School of Meteorology, University of Okalhoma, Norman OK 73019, mxue@ou.edu.

soil heat capacity dependency on soil moisture contents. More recently, with the development of remote sensing techniques, several proposals for assimilating ground surface temperature to infer other related land surface variables were made (e.g., Boni *et al*. 2001; Li and Islam 2002). Boni *et al*. (2001) proposed a scheme for assimilating ground temperature for the estimation of a surface soil moisture index. They found that the optimized surface soil moisture index also leads to satisfactory description of the surface energy balance components. Through a series of sensitivity experiments, Li and Islam (2002), however, found that initial soil moisture profile that optimizes the surface soil moisture description only does not necessarily lead to optimal estimation of surface fluxes, the success of the inversion technique depends, to a large extent, on the successful retrieval of root zone soil moisture content instead.

As part of our effort in developing a general framework for estimating or retrieving the land surface model variables and certain uncertain model parameters, we present in this paper an adjoint-based 4D variational (4DVAR) retrieval system that assimilates ground temperature observations from either remote sensing satellite or surface observation stations. In another word, the ground or skin temperature data are assimilated for the purpose of determining model variables and parameters that are not directly measured. This system, when completed, will also be capable of retrieving these variables using near-surface atmospheric measurements or a combination of the two. Through the analysis, questions such as what soil properties can be effectively retrieved from only the surface temperature information are addressed.

Compared to more traditional methods, the main advantages of the adjoint-based 4DVAR method lie with its ability to optimally use observations distributed over time, and observations that are indirectly related to the variables to be determined or retrieved. For a relatively simply system, the method can also be rather efficient. In addition, the adjoint model provides a powerful tool for studying sensitivities of the model output to input parameters, hence provides physical insight on the behaviors of the land surface system.

Our data assimilation system is based on a revised force-restore-based land surface model, which is briefly presented in section 2a). The variational retrieval scheme is discussed in section 2b). Section 3 describes data from the OASIS data that are used by our forward model validation and data assimilation experiments. In the first part of section 4, systematic numerical experiments using synthetic (model simulated) data are performed to test the effectiveness and robustness of the retrieval scheme. Analyses of the causality mechanism involved are performed. Issues such as the relative sensitivity of the target variable, the variable that will be compared against observations, to the initial soil variables are also addressed. The results provide guidance for our data assimilation experiments with real observations in section 4c) where the basic issues such as temporal resolutions and distributions of observations are further discussed. The most informative periods during the daily cycle are identified and an explanation given. The effects the assimilation window length on the quality of retrieval is evaluated and the results are found to be consistent with our understanding on the information content of the observations. A summary is given in section 5.

2. 4DVAR system for retrieving initial state of land-surface model

Four-dimensional variational (4DVAR) data assimilation and/or retrieval systems seek to minimize the misfit between observations distributed over a period of time (called the assimilation window) and the prediction of a forward forecast model (Le Dimet and Talagrand 1986; Talagrand and Courtier 1987). A cost function, typically of a quadratic form, measures such a misfit. The initial condition of the forecast model is adjusted, starting from an initial guess, so as to minimize the cost function. When the variables defining the initial conditions are not directly measured, such variables are said to be retrieved (from the observed quantities) and the entire procedure is often referred to as retrieval. Since such schemes determine the input of a prediction model by constraining the model output, they are called inverse methods.

Efficient minimization algorithms, such as the conjugate gradient method (Navon and Legler, 1987) used in this study, requires the gradient of the cost function with respective to the variables that are to be adjusted in the initial condition (called control variables), and the gradient can be efficiently obtained by a backward-in-time integration of the adjoint model. Here, the adjoint is mathematically defined as the transpose of the tangent linear approximation to the nonlinear forward prediction model (Le Dimet and Talagrand 1986). In a more general system, the control variables can include other parameters such as those found in the formulation of the forward model.

In the standard 4DVAR procedure, the forward prediction model is used as a strong constraint, i.e., it is strictly satisfied during the assimilation period. For this reason, the accuracy of the forward model does affect the accuracy of the retrieval. An additional requirement for a successful retrieval is the existence of a strong connection between the variables that are measured and those to be retrieved. This is referred to as the sensitiv-

ity of the cost-function to the control variables, and the sensitivity is measured by the gradient of the cost function with respect to the particular control variable.

As pointed out earlier, in this study, we attempt to retrieve the state of the soil and vegetation, by using skin temperature measurements. The model used is an improved version of the two-layer soil-vegetation (or land surface) model and the surface physics model from the ARPS (Advanced Regional Prediction System, Xue *et al.* 2000, 2001). For the current application, the atmospheric component is not needed, except for the package that calculates the surface fluxes. In the first part of this section, we will describe the forward model.

a) Forward model

Land surface schemes, although developed based on different concepts and with different levels of complexity, are schemes that solve energy conservation (for soil temperatures) and mass conservation (for soil water content) equations. The aim of land surface scheme is to provide temperature and specific humidity at the lower boundary of atmospheric models (Mahfouf and Viterbo 2001). These two variables are needed in the estimation of heat, water and momentum exchanges between the land surface and the lower atmosphere. Numerically, the link between soil and atmospheric variables is provided through the parameterizations of the surface fluxes, usually based on similarity theory.

1). Soil temperature equations

In the original derivation of the force-restore scheme (Bhumralkar 1975; Blackadar 1976) for surface soil temperature, Bhumralkar (1975) assumed the daily mean temperature is the same at all soil depths. This assumption may be valid in the spring and fall seasons of a year, but is not true in most parts of the world at summer and winter (de Vries 1963; Ren and Xue 2003). Ren and Xue (2003) discussed this issue and proposed a method for proper incorporation of the mean soil temperature lapse rate in the force-restore model. The most significant refinement by Ren and Xue (2003) to the standard force-restore scheme is to include the effect of seasonal mean vertical temperature gradient, defined in terms of the seasonal mean temperature difference between surface and deep soil. Here the seasonal mean soil temperature is defined as the running mean of soil temperature over one to two weeks, a period long enough to remove diurnal temperature changes while retaining seasonal variations.

The revised force-restore soil temperature equations that we use in this study are as follows,

$$\begin{cases} \dfrac{\partial T_{sfc}}{\partial t} = C_T (R_n - LE - H) - \omega (T_{sfc} - T_{dp} - \pi d\gamma), \\ \dfrac{\partial T_{dp}}{\partial t} = -\dfrac{1}{\tau}[T_{dp} - T_{sfc} + \gamma \pi d], \end{cases} \tag{1}$$

where T_{sfc} is ground surface or skin temperature, R_n is net radiation flux and is provided by OASIS observations in our experiments. *LE* is latent heat flux, *H* sensible heat flux, ω the frequency of diurnal oscillation ($\omega = 2\pi / \tau$, with τ is equal to 24 hours). C_T is thermal conductivity, T_{dp} is deep soil temperature, γ is the "lapse rate" of the seasonal mean soil temperature, $d = \sqrt{2\lambda_T / C_V \omega}$ is the *e*-folding depth of downward propagation of surface temperature oscillation signals. Here λ_T is soil thermal diffusivity. In the ARPS implementation, a single (mixed) heat budget is considered for the bare ground and the vegetation therefore C_T represents the average conductivity of bare ground and vegetation, and

$$C_T^{-1} = [(1 - veg) C_G^{-1} + veg C_V^{-1}] \ . \tag{2}$$

Here *veg* is the fractional vegetation coverage, C_V is the inverse of vegetation heat capacity, and C_G the inverse of volumetric ground heat capacity which is parameterized as

$$C_G = C_{Gsat} (w_{sat} / w_{dp})^{\frac{b}{\ln 100}} \ . \tag{3}$$

Here C_{Gsat} is the inverse of soil heat capacity at saturation (depends on soil textural properties), w_{sat} the saturation soil water content, w_{dp} the bulk/deep-layer soil moisture content, and *b* the slope of log-retention curve, i.e., Clapp-Hornberger (1978) parameter. Retrieval of soil wetness condition is possible partially because of Eqs. (2) and (3), or the soil moisture dependence of soil heat capacity. For this reason, the adequate modification to C_V by Pleim and Xiu (1995) is considered critical for our retrieval. However, unlike the simplistic scheme of Xu and Zhou (2003), our dynamic system includes more sensitive channels, such as transpiration from deep layer which signifies another mechanism for deep soil moisture to influence on surface energy partition and hence surface temperature evolution.

2). Soil moisture equations

The soil moisture content has a strong impact on the humidity of the low-level atmosphere. Land surface processes are especially relevant for warm season precipitation in temperate region, where, on daily basis, evaporation constantly extracts water from land surface and this effect propagates to deeper soil through evapo-

ration for bare soil or intercepted water reservoirs and through transpiration for vegetated areas. The atmospheric precipitation process, in return, recharges the soil in the hydraulically active layer, where soil moisture conditions control the partitioning of surface energy fluxes, ecosystems, and biogeochemical cycles. The existence of roots facilitates the water flow through the soil-vegetation-atmosphere system through transpiration during daytime and hydraulic lift (M. Caldwell, personal communication) during nighttime, and thus significantly enlarges the depth of hydraulically active soil layer. In this study, we follow the governing equations for three water reservoirs as in NP89,

$$\begin{cases} \dfrac{\partial w_{sfc}}{\partial t} = \dfrac{C_1}{\rho_w d_1}\left(P_g - E_g\right) - \dfrac{C_2}{\tau}\left(w_{sfc} - w_{geq}\right), & (4a) \\ \dfrac{\partial w_{dp}}{\partial t} = \dfrac{1}{\rho_w d_2}\left(P_g - E_g - E_{tr}\right), & (4b) \\ \dfrac{\partial w_v}{\partial t} = vegP - \left(E_v - E_{tr}\right) - R_v. & (4c) \end{cases}$$

Here the three prognostic variables include the soil surface wetness (w_{sfc}), the bulk/deep-layer soil moisture (w_{dp}), and the canopy interception water (w_v). w_{sfc} and w_{dp} are defined in terms of volumetric water content in the soil. These time-dependent parameters are forced by precipitation reaching the ground (P_g), the bare ground evaporation (E_g); evaporation of the wet part of vegetation (E_v); the transpiration of the dry part of canopy (E_{tr}); and vegetation dripping (R_v). P_g equals the total precipitation (per unit surface area, P) reaching the ground ($P(1-veg)$) plus that possibly dripped from the canopy (R_v). The surface moisture is restored to equilibrium by moisture sources from the thick underlying soil layer by the second term on the right hand side of Eq. (4a). The time scale at which this restoring process acts is prescribed *a priori* in the form of time constant τ which is set to one day, and the two soil-layer depths for soil moisture are d_1 and d_2.

Thus, the interaction between the soil and atmosphere varies as a function of the vegetation coverage, vegetation type, soil type, and hydraulic conductivity. These properties are specified by diagnostic variables C_1 (representing the hydraulic property of soil affecting the infiltration at the surface), C_2 (the subsurface conductivity) and C_T. These parameters are formulated in terms of basic soil parameters such as soil moisture at field capacity (w_{fc}, or filled reservoir capacity), wilting point (w_{wilt}, below which the plant transpiration ceases), saturation soil water content (w_{sat}, the maximum possible water content), w_{geq} (the surface volumetric moisture at the balance of gravity and capillary forces), and various other thermal and hydraulic properties of the soil as described in Appendix A.3 of Noilhan and Mahfouf (1996). Since all soil properties are specified according to the 11 soil types of the U. S. Department of Agriculture (USDA) soil textural classification (Clapp and Hornberger 1978), the only soil data required for this model are the soil types by textural classifications.

Although the soil moisture equations are kept formally the same as in NP89, coefficients in the equations have been modified according to their later work (Noilhan and Mahfouf 1996). We also follow several revisions by Xiu and Pleim (2001), which include the modifications to aerodynamic resistance R_a, the functional form of soil moisture availability (to PBL) h_u, and stomatal resistance R_s .

b) The 4DVAR assimilation system

To establish a 4DVAR retrieval system based on the two-layer ISBA like land surface model in the ARPS, we first developed the tangent linear counterpart (TLM) of the land surface model then the adjoint of the TLM.

The development of TLM and adjoint models, especially the letter, by hand is a tedious process that is prone to error. Fortunately standard procedures exist that can be used to validate the codes. In this study, the TLM is first validated by comparing the TLM solution with the difference of two nonlinear model solutions starting from sufficiently close initial conditions. The adjoint code is verified using the inner product of TLM and adjoint solutions, and the inner product should be time invariant (LeDimet and Talagrand, 1986). The relative magnitude of the fluctuation of this time invariant quantity is verified to be less than 10^{-7} for a two-day assimilation cycle using reasonably small initial perturbations and double precision.

Based on our assumed data availability (only ground temperature data are available), the cost-function, J, used in this study is defined as the quadratic difference between modeled and measured ground surface temperature, i.e.,

$$J = \frac{1}{2}\sum_{n=1}^{N}\left(\frac{T_{sfc,n}^{o} - T_{sfc,n}^{f}}{\sigma_{sfc}}\right)^2, \qquad (5)$$

where superscript '*o*' means observation and '*f*' is model forecast. σ_{sfc} is the standard deviation of the surface temperature observation errors. The magnitude

of σ_{sfc} is actually trivial in this study since only one type of data is assimilated and we assumed exact forward model. In our system, the control variables are the surface and deep layer soil temperature, surface and deep layer soil moisture and canopy water content at the beginning of the assimilation window.

Once a correct adjoint system is available, all the sensitivity information can be conveniently obtained via a single backward integration of this adjoint model. Using an optimization scheme, errors of ground surface temperature prediction are then systematically minimized subject to the constraints imposed by the model. An automatic tool TAFLINK from Fastopt (http://fastopt.com), which provides gradient test results, was used to test the optimization system.

In the minimization procedure, it is also necessary to properly scale the control variables so that they are of similar order of magnitude. The scaling improves in conditioning of the optimization procedure (Jones *et al.* 2003). It will be shown that our optimization system can perform retrieval for all control variables simultaneously inasmuch as the small perturbation and over-determination requirements are satisfied. Small perturbation requirement demands that our initial guesses should not be too far from the truth to void the tangent linear assumption. Over-determination requires that the observations (ground temperatures here) outnumber the total control variables.

3. OASIS Data

In this study, we use the OASIS data set at the Norman, Oklahoma site (Elevation: 360 m; Latitude: 35 15' 20"; Longitude: 97 29'; Slope: 0.0) for validating the forward model, forcing the assimilation system and for evaluating the retrieval results. The same data set has been used for model verification by Brotzge and Weber (2002). The meteorological parameters available include surface temperature, water vapor mixing ratio, wind speed, surface pressure, and precipitation rate at 5 min intervals (except that precipitation is accumulated from 00Z). The soil moisture and temperature are available in terms of half-hourly averaged time series at 5, 25, 60 and 75 cm depths. Vegetation parameters (vegetation type and coverage, LAI and NDVI index) are estimated biweekly. The OASIS year round continuous and direct measurements of soil moisture, temperature and all four components of surface fluxes provide the opportunity for rigorously validating and improving the dynamic framework of LSMs, and for objective determination of most uncertain soil and vegetation parameters via retrieval techniques.

Our selected time period between 00UTC, 4 and 28 August 2000 represent a synoptically quiescent period with clear sky conditions and wind speed generally less than 5 m s^{-1}. Forced by periodical net radiation flux, the air pressure, air temperature (maximum temperature of 42° C), water vapor mixing ratio and soil temperatures within 25 cm depth all show apparent daily cycles. This period shows a drying down process of the surface soil moisture. The volumetric soil water content at 5cm did not drop sharply during the first 48 hours starting from 00UTC, August 4, indicating that the process enters stage II of drying after significant rainfall (Idso *et al*. 1974). The soil moisture measurements at the remaining three depths show little changes on daily basis.

The possible assimilation period, 4-28 August 2000, a total of 24 days, is divided into three consecutive periods of 8-day each. During this period, vegetation is at a slightly stressed stage of growth, although with a rather high vegetation coverage (0.75, estimated based on the study of Brotzge and Weber (2002) for 20 May 2000) and LAI=0.72. The vegetation is slightly stressed because NDVI = 0.5 rather than 0.55 for 20-22 May period when vegetation is rather active. The soil moisture contents at the top measurement depth (5 cm) fall close to the wilting point value (22% for silty clay soil) at the driest hours of the day. This is also shown by the Halstead coefficient $((1-\delta)R_a/(R_a+R_s)+\delta$ being around 0.06 for this period. Here δ is wet fraction of canopy as parameterized by Deardorff (1978). Considering that the dew formation is insignificant, the stomatal resistance is nearly 20 times that of aerodynamic resistance.

The surface ground temperature to be assimilated can be inferred from soil profile measurements by an extrapolation technique (Jackson, 1997; Ren and Xue, 2003). However, for our selected periods, the *a priori* determined K_T tends to give lower surface temperature amplitude due to an overestimation of the scaling depth. We thus use directly measured infrared surface temperature rather than the extrapolated surface temperature.

In the experiments described below, the site specific surface roughness characteristics are estimated using the specifications by Brotzge and Weber (2002) (e.g., surface dynamic roughness z_o = 0.004 m). The seasonal mean surface-deep layer soil temperature difference (Column 3 in Table 1) is estimated using the method described by Ren and Xue (2003). Soil and vegetation properties at Norman site are also listed in Table 1.

4. Numerical experiments

The aim of this study is to test the feasibility and accuracy of retrieving initial values of the prognostic variables of the Interactions Soil Biosphere Atmosphere (ISBA, Noilhan and Planton 1989)-like LSM through merely minimizing the trajectory difference between modeled and observed surface temperatures. Both synthetic and real observations are utilized in the following numerical experiments. Using synthetic data is advantageous for validating an optimization system because we can control and know the quality of observations. Theoretical issues can also be more readily addressed by assimilating synthetic data. The ultimate goal is to assimilate real and often noisy data where the issues of data sparsity also arise. In this section, we first discuss synthetic data assimilation, where we present mainly heuristic examples to illustrate the mechanism. We will then discuss real data assimilation and analyzing related issues. Before proceeding to the variational retrieval experiments, the soundness of the forward system is first examined.

a) Forward model verification

A systematic refinement and verification study had been performed for our land surface scheme using OASIS data by Brotzge and Weber (2002) and further by the current authors (Xue and Ren 2003) from Norman, Oklahoma site, as described in Section 3. Here we highlight some of the findings of the model validation. For these experiments, the surface/skin temperature, superficial and deep soil moisture values are initialized with OASIS observations. The deep soil temperature is initialized using OASIS data, according to a procedure described in Ren and Xue (2003). The time step size used for LSM integration is 30 minutes. The model is run in 1-D stand alone mode, with time-dependent surface atmospheric variables specified using OASIS measurements. Figure 1 shows that the skin surface temperature is predicted accurately, with peak value differences being less than 2 K in general. There is no apparent phase error and a slight warm bias in the prediction is less than 1 K. The time series for deep soil temperature (T_{dp}) indicate that the revisions to the force-restore temperature equations (Eq. (1)) are very necessary. Otherwise, the deep soil temperature will drift upward and tends to assume a same daily average as the surface temperature. The time series of the superficial soil moisture measurements show one large spike, due to presumable measurement error. We must point out that, for the predicted superficial soil moisture appears out of phase from the observation. The missing hydrolic lift processes is believed to be the cause and a remedy has been proposed in another work by the current authors. Because of the shallowness of the surface layer (therefore smallness of the effect on soil heat capacity), the difference does not affect our retrieval in any significant way. The prediction of the deep soil moisture (w_{dp}) is rather satisfactory, with the drying trend corrected predicted. The model-measurement difference is less than 0.005 $m^3\ m^{-3}$ (except at the time of a possibly wrong measurement in the afternoon of August 5), the magnitude of typical errors of the Campbell Scientific 229-L heat dissipation instrument used for soil moisture measurement. For this period of simulation, during daytime, our model suffers systematic overestimation to both latent (*LE*) and sensible (*H*) heat fluxes, with peak differences as large as 50 W m^{-2} in *H*. During nighttime, model overestimates *H* while underestimates *LE*. The magnitudes are however both small (less than 15 W m^{-2}) and well within the instrument error ranges for these variables.

b) Assimilation experiments with synthetic data

As the initial stage of assimilation system testing, using simulated data is beneficial since there are no uncertainties associated with model and measurement errors. We created the synthetic time series of ground temperature and other model state variables and fluxes (which are not assimilated but can be used for model verification purposes) by sampling a forward run of the LSM every half hour. This forward run is the same one presented and verified in the previous subsection. The initial correct land surface conditions are $\left(T_{sfc}^0, T_2^0, w_g^0, w_2^0, w_{canopy}^0\right)$ = (307.16K, 301.16K, 0.2562 $m^3\ m^{-3}$, 0.2802 $m^3\ m^{-3}$, 0.0 $m^3\ m^{-3}$), where superscript '*0*' means initial value**.** Although the exact value of σ_{sfc} in Eq. (5) is not significant, the measurement uncertainties are specified as 2*K* and 0.005 $m^3\ m^{-3}$ for soil temperature and moisture, respectively, for quantifying our initial guess errors of these parameters.

The first retrieval experiment was done by adding a positive error of 0.025 ($m^3\ m^{-3}$) to the initial deep soil moisture w_{dp}^0. We choose an assimilation window of two days, starting at 00 UTC, August 4, 2000. The time series of the adjoint variable are plotted in Fig. 2, for the first iteration in the minimization process. Except for the adjoint variable corresponding to canopy interception (which is constantly zero because there was no precipitation or dew formation during this period), they are plotted from the least sensitive variable (top panel ad_T_{sfc}) to the most sensitive variable (bottom panel ad_w_{dp}).

Except for ad_T_{sfc}, all of the remaining adjoint variables are equal to zero at the end of the assimilation

window of two days. The ad_T_{sfc} at this time holds the difference between modeled and synthetically "observed" surface temperature, due to initial deep soil moisture error. Mathematically, this is a direct result of the terminal condition placed on the adjoint model. Physically, this represents the deterministic causal nature of the system, in that perturbations in the state variables beyond the assimilation window have no influence on the cost function. As the adjoint model proceeds backward in time toward the initial condition, each adjoint variable will accumulate information that influences the cost function in different ways depending on their respective roles in the forward dynamic system. In a certain sense, the backward marching of the adjoint model signifies an information collection process that marks the influence of each state on the trajectory fitness of T_{sfc} over this assimilation period. Thus, the first usage of this diagram is to identify the relative sensitivity among different variables.

The second goal is to identify the positive or negative feedback that is triggered by the introduced initial condition error. For example, in Fig. 2, we see that the most sensitive adjoint variable is the one corresponding to deep soil moisture. The positive feedback is accumulated all the way to about 2000 (at the initial time), and is more than two orders of magnitude larger than the adjoint variables for soil temperature. If the optimization scheme is allowed to vary only one specific variable to minimize the cost function, it will surly pick the initial deep soil moisture. After analyzing the convergence of the corresponding estimated state variable to its actual value (Fig.3), we found that the adjustment to w_{dp} is the most significant one, while the adjustments to T_{sfc} and T_{dp} are insignificant. Similar analysis can be applied to any interim time step using the result of Hall and Cacuci (1983), i.e., the adjoint variable at any time $t = \tau$ is directly related to the sensitivity of cost function to a perturbation in its corresponding state variable at that particular time. Thus, we can cross compare between states to see which is most important at different times during the model integration.

Fig. 2 indicates the feedback (to the innovation in w_{dp}) from T_{sfc} is a negative one, and is especially significant during daytime heating period. Thus, increased deep soil moisture increases *LE* and reduces the energy used for heating the ground. The feedback from deep soil temperature is similar but with phase shift resulted from the larger heat inertia of the deep soil layer. The feedback from w_{sfc} is positive, since it may gain (or loss less) soil moisture from the capillary effects through exchanging with deep soil reservoir. The most sensitive period is also during daytime, indicating that the soil moisture exchange is passively forced by the evapotranspiration process.

Without model error and with complete surface temperature measurements, the minimization is very efficient. After three iterations, the cost function is reduced to only 0.1 percent the initial value. The iterations were carried out five times before terminating, satisfying the smallness of the criteria, which is defined as the cost function difference between two adjacent iterations.

With the assimilated model parameters, we integrated the forward model for one more time to obtain the state estimates. Figure 3 illustrates how much improvement is gained in representing the states by assimilating merely the ground temperature observations. Once the convergence has been reached, the assimilated T_{sfc} trajectory nearly coincidences with the observed one as expected. The two soil moisture values and the deep temperature values, although no corresponding observations for them are assimilated, can also be successfully updated, through the model dynamics as a bond. Their initial guess trajectory are all apparently deviant from the corresponding true trajectories. The surface fluxes are also significantly improved.

We must emphasize that T_{sfc}^0 is the least sensitive variable in this experiment. If the sensitivity channels of w_{sfc} and w_{dp} are closed, however, initial value of ground temperature experiences large and proper adjustment to its true value. Unfortunately, the simulation of *LE* and *H* using the retrieved temperature values (T_{sfc}^0 and T_{dp}^0) are both poor (Fig. 4), because the initial guess errors associated with soil moistures (panels c and d) cannot be effectively removed by adjusting initial soil temperatures.

Because the LSM is run in stand alone mode, the effects on PBL structure cannot be shown. In reality, because of the slow rate at which the deep soil moisture evolves and the limited number of *in situ* observations that can be anticipated, the start-up bias is likely to persist for weeks. As will soon be seen, the relative importance of superficial and deep soil moisture contents depends on site-specific land-cover conditions. For less vegetated area, surface soil moisture may play a more important role than shown here based on Norman site characteristics.

For a deeper understanding of the physics, we examined the contour of the cost function for the assimilation period (figures not shown). We found that the cost function sphere is very irregular. For example, for the cross-section of w_{sfc} and w_{dp}, the contour of the cost function is elongated along w_{sfc}, resulting from the relative strengths of the evaporation from layer one and transpiration from both layers one and two. With the actual values of vegetation cover *veg=0.75* and root depth 1 m and the atmospheric conditions of our se-

lected period, the transpiration from layer 1 is negligible and transpiration from layer 2 is about 10 times of the evaporation from layer 1. Thus the results are much more sensitive to w_{dp} than to w_{sfc} as long as vegetation growth is not extremely stressed. The behavior can be different for a different set of parameters. For instance, if the vegetation coverage is specified close to 0, w_{sfc} could be more accurately determined while the information on w_{dp} would be degraded.

Similar experiments are repeated in which errors are added to T_{sfc}, T_{dp}, and w_{sfc}. The results are summarized in Table 2. Because the uncertainty in initial surface temperature has least influence on the ensuing surface temperature evolution, it is most difficult for the minimization scheme to pinpoint the exact initial value of surface temperature. On the contrary, the deep soil moisture has the strongest influence to the surface temperature evolution, thus is the easiest one to retrieve. This is also consistent with our nonlinear sensitivity experiments evaluating the influence of each surface parameter on surface fluxes and ground surface temperature. We found *LE*, *H* and T_{sfc} are very sensitive to the soil water contents, and the influence lasts longer than ten days for a reasonable perturbation on deep soil moisture and barely one day for superficial soil moisture. For the evolution of T_{sfc}, the influence of initial value uncertainty of deep temperature is much low and the one of T_{sfc} itself is quite insignificant. Viewing form another angle, the initial guess error associated with surface temperature is least important and the initial errors in soil moisture are far more important in affecting the surface temperature. Assimilation scheme for initial land surface conditions should concentrate on the better estimation of soil moisture conditions.

Now, the control variables $\left(T_{sfc}^0,\ T_{dp}^0,\ w_{sfc}^0,\ w_{dp}^0,\ w_{canp}^0\right)^T$ are perturbed simultaneously by $\left(-1.0K, +0.5K, -0.02\ m^3\ m^{-3}, +0.01 m^3\ m^{-3}, 0.0\right)^T$, or starting from an initial guess vector of $\left(306.16\text{K}, 301.66\text{K}, 0.2362\ \text{m}^3\text{m}^{-3}, 0.2902\ \text{m}^3\text{m}^{-3}, 0.0\right)^T$. It takes the optimization scheme 8 steps to converge to $\left(306.16K,\ 301.66K,\ 0.2613\ m^3 m^{-3},\ 0.2803\ m^3 m^{-3},\ 0.0\right)^T$ as shown in Fig. 5. The root mean square error (*rms*) in surface temperature prediction was reduced to 7.5% of its initial amount (Fig. 5a).

In addition to the case shown in Fig. 5, other scenarios were tested with different perturbation magnitudes. In general, the estimated values were robust for reasonable perturbations in the prior values. If the perturbations are too large, however, there can be other estimates that reproduce the measured ground temperature but are different from the actual state of the system (i.e., solution bifurcation). This ambiguity can be reduced in actual application for coupled runs by introducing other observations to the cost function, i.e., PBL atmospheric measurements.

Robustness is also an issue, especially for using real data, where model error is usually involved. We tested the robustness of the assimilation scheme by adding zero mean Gaussian noise of different standard deviations (*std*) to the synthetic observations of T_{sfc}. The final degree of closeness as measured by the cost-function degrades with increasing noise level. Figure 6 is the case for retrieving initial conditions starting from a perturbation to the control variable vector by $\delta U = \left(-1K, 0.5K, -0.02\ m^3 m^{-3}, 0.025\ m^3 m^{-3}, 0\right)^T$. The initial value retrieval does not work as the noise level surpasses 2.0K *std*. This also attests to the importance of our modification to the original force-restore scheme shown in section 2a). Although the effects on surface temperature prediction is small (generally less than 1 K), as will soon be shown, our modification is vital for a successful retrieval by assimilating real observations.

c) Assimilation experiments with real measurements

In contrast to using synthetic data, two more difficulties must be dealt with for real data assimilation, namely, data sparseness and model error. We use OASIS measures T_{sfc} time series of 8 days long (00Z, 04-12 August 2000) in the following experiments except for those explicitly stated.

Using complete observations, after 5 iterations, the initial guess errors (Fig. 7) can be effectively removed, especially those associated with soil moistures. We also performed two other experiments assuming data availability at a 6-hour wide window centered on local noon (~18Z at Norman site) and a 3-hour wide window centered at local midnight (~06Z at Norman site). The first is called daytime assimilation and the second nighttime assimilation. We compared the assimilation results of both soil temperatures and surface fluxes (*LE* and *H*). The nighttime assimilation yields worse estimations for all the quantities (Fig. 7b and c). It misrepresented the peak values of surface temperature and accordingly severely underestimated the sensible heat flux, although the nighttime surface temperature itself is simulated rather well and the initial guess error decreased by 80%, the best among the three sampling schemes. The errors in *LE* and *H* seem systematic (for clarity, the 8-day period is bisected into two panels).

This has important implications for understanding that some phases of daily cycle are more informative

for retrieval scheme, with the caveat that measuring scheme performance from only the convergence rate of the cost function is very uncertain. For example, the readers are cautioned not to compare the convergence rate as shown in Fig. 7d, because, except for the complete assimilation case, only those periods with observations (see Fig. 7a for sampling schemes) are used in composing the cost function, not the whole period. In reality, except for intense field measurements, we usually do not know the complete time series of T_{sfc}, since only a couple passes are made by one satellite observation platform per day.

We did not find the daytime periods especially informative for experiments with synthetic data. The reason why the selected daytime period are more informative for real observations assimilation may be because the signal to noise ratio is high for that period, when the forcings are steadily strong. In other words, the weak signal during nighttime tends to be inundated by the instrument errors. We tested a series of assimilation windows which are of equal length (3 hours) and adjacent to each other (3 hours apart) and together cover the whole daily cycle. To our surprise, the quickly heating up period (early morning hours) is the worst assimilation period. We repeated this experiment by varying the window length from 1 hour (2 measurements) to 6 hours (12 measurements). Although there are very noisy periods when the assimilation window is narrow, the general pattern is nearly the same for all cases.

Figure 8 illustrates the effects of assimilation period length. Using OASIS measurements, we performed three retrieval experiments with respective assimilation period of 2, 4, and 8 days, all starts from 00Z, 4 August 2000. Using retrieved initial conditions, an extra forecasting period of 8 days was made for each experiment. The shared forecasting period, i.e., 00Z, 12 August-00Z, 14 August, is shown here. The forecasts for the prognostic model variables (w_{sfc} and w_{dp} are not shown for clarity) are best from using the initial conditions retrieved from T_{sfc} contained in 8 days. This agrees with our assertion about the information redundancy. To overcome the noisy signal contained in the observations, the quantity of observation matters. To have a deeper understanding, we also calculated *rms* errors for this forecasting period. The *rms* errors for using assimilation periods of 2, 4 and 8 days are respectively 1.96, 1.92 and 1.36K for T_{sfc}; 0.98, 0.95 and 0.55 for T_{dp}; 0.016, 0.015 and 0.0037 m^3 m^{-3} for w_{sfc}; 0.02, 0.019 and 0.003 m^3 m^{-3} for w_{dp}; 34, 32, and 20 W m^{-2} for H; and 33, 32 and 19 W m^{-2} for LE. With the increased assimilation period, the estimations for the initial conditions of the two soil moistures are most significantly improved, followed by deep soil temperature and by surface temperature. As a result, the estimations for both latent and sensible heat fluxes are significantly improved.

5. Conclusions

In this study, based on a two-layer soil model and its adjoint, a rather general 4DVARdata assimilation framework for retrieving land surface variables is implemented. The adjoint model is used in this variational data assimilation framework to yield dynamically consistent estimates of land surface states and fluxes that optimally merge observations with the physical model. We performed retrieval experiments for soil state variables and verified the results against OASIS data. Good results were obtained by retrieving the initial soil temperature and moisture using the reduced 4DVAR system when the soil skin temperature in the assimilation window is known. Tests were done with both model-simulated and OASIS-measured data.

In the experiments with synthetic data, the primary goal was to demonstrate the capability of the adjoint model as a general tool for sensitivity analysis and data assimilation. We have used simple illustrative examples of each type of application to give a glimpse of the kind of problems that can be addressed using the model and its adjoint. We first showed that the assimilation works for the initial value retrieval for perturbations larger than the corresponding instrumentation errors. Because the number of control variables is small, large information redundancy makes the retrieval very efficient. Based on our LSM, we also tested the robustness of the estimation scheme by adding random white noise to the synthetic surface temperature time series.

For real OASIS measurements assimilation, initial soil moisture conditions can be successfully retrieved, even though half-hourly surface temperature observations are only available for a narrow 6-hr window centered at local noon. However, assimilating only nighttime surface temperatures causes poor retrieval results. Further investigation shows that the stronger daytime signal to noise ratio explains why daytime periods are more informative.

We examined several outstanding issues in this land surface data assimilation study. If all sensitivity channels are open, it is shown that the data assimilation system for the surface ground temperature is capable of accurately estimating the components of the surface energy balance. This holds for assimilating both synthetic data and real observations. As long as all the sensitivity channels are open, the scheme is able to differentiate the sources of sensitivity. These successful results are a consequence of the explicit inclusion of the vegetation transpiration process (whereas the simple schemes such as Xu and Zhou (2003) lack this) and

the proper initialization of deep soil temperature in our general land surface data assimilation scheme.

We analyzed the assertions of Li and Islam (2002) and Boni *et al*. (2001) and found that both may be correct since the field site used by Boni *et al.* (2001) is much less vegetated and consequently the ground evaporation overwhelms the transpiration process. For a bare ground surface, the soil moisture availability may be the single most important parameter that affects the simulation of the ground latent heat flux. For a highly vegetated area, however, the surface moisture flux is generally dominated by evapotranspiration. The key parameters for realistic simulation of evapotranspiration are then root zone soil moisture content and the canopy resistance. As a conclusion, we believe that for a less vegetated area, because deep soil moisture (i.e., deeper than 50 cm, Deardorff, 1978) seldom varies on a daily basis, numerical weather forecasting, especially for the warm season precipitation, requires an accurate description of the surface soil moisture content and hence an improved knowledge of how the net radiation is partitioned among latent, sensible and ground heat fluxes. On the other hand, because of the slow rate at which the deep soil moisture evolves and the limited number of *in situ* observations that can be anticipated, the start-up bias is likely to persist for weeks. With surface temperature observations, variational data assimilation is good at determining the value of the deep soil moisture.

Acknowledgement

This work was supported by a DOT-FAA grant, and by NSF grants ATM9909007 and ATM0129892 to the second author. The authors thank Dr. Jidong Gao for discussions on adjoint code development and Dr. Jerry Brotzge for making available the OASIS data set and for very helpful discussions on the related issues.

References

Bhumralkar, C. M., 1975: Numerical experiments on the computation of ground surface temperature in an atmospheric general circulation model, J. Appl. Meteo., 14, 1246-1258.

Blackadar, A. K., 1976. Modeling the nocturnal boundary layer. *Prec. Third Symp. Atmos. Turb., Diffusion and Air Quality*, Boston, Amer. Metero. Soc., 46-49.

Boni, G., Castelli, F., Entekhabi, D., 2001. Sampling strategies and assimilation of ground temperature for the estimation of surface energy balance components, IEEE Transactions on Geosciences and Remote Sensing, 39(1), 165-172.

Brotzge, J. A., Weber, D., 2002. Land-surface scheme validation using the Oklahoma atmospheric surface layer instrumentation system (OASIS) and Oklahoma Mesonet data: preliminary results, Meteo. Atmos. Phys., **80,** 189-206.

Brotzge, J. A., 2000. Closure of the surface energy budget. PhD dissertation, School of Meteorology, University of Oklahoma, Norman. 208 pp

Calvet, J.-C., Noilhan, J., Bessemoulin, P., 1998. Retrieving the root-zone soil moisture from surface soil moisture or temperature estimates: A feasibility study based on field measurements. J. Appl. Metero., 37**,** 371-386.

Clapp, R. B., Hornberger, G. M., 1978. Empirical equations for some soil hydraulic properties. Water Resour. Res., 14, 601-604.

Deardorff, J. W., 1978. Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation, J. Geophys. Res., 83(C4), 1889-1903.

de Vries, D. A., 1963. Thermal properties of soils. Physics of Plant Environment, W. R. V. Wijk, Ed., John Wiley & sons, Inc, 210-235.

Hall, M., Cacuci, D. G., 1983. Physical interpretation of the adjoint functions for sensitivity analysis of atmospheric models. *J. Atmos. Sci*., 40, 2537-2546.

Henderson-Sellers, A., McGuffie, K., Pitman, A. J., 1996. The Project for Intercomparison of Land-Surface Parameterization Schemes (PILPS): 1992 to 1995, Climate Dynamics, 12, 849-859.

Henderson-Sellers, A., A. J. Pitman, P. Irannejad, and K. McGuffie, 2002: Land-surface simulations improve atmospheric modeling. *EOS, Trans, AGU*, **83**, 145-152.

Idso, S.B., Reginato, R. J., Jackson, R. D., Kimball, B. A., Nakayama, F.S., 1974: The three stages of drying of a field soil. Soil Sci. Amer. Proc., 38, 831-837".

Jackson, T. J., Southern Great Plains 1997 (SGP97) Hydrology Experiment Plan, http://hydrolab.arsusda.gov/sgp97/.

Jones, A., T. Vukicevic, and T. H. Vonder Haar, 2003: A microwave satellite observational operator for variational data assimilation of soimoisture. *J. Hydrol. Meteorology* (Submitted).

LeDimet, F. X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations-Theoretical aspects, Tellus, 38A, 97-110.

Li, J., and S. Islam, 2002: Estimation of root zone soil moisture and surface fluxes partitioning using near surface soil moisture measurements. J. Hydrology, 259, 1-14.

Linsley, R. K., M. A. Kohler, and J. L. H. Paulhus, 1949: *Applied Hydrology*. McGraw-Hill., 689pp.

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

Mahfouf, J.-F., and P. Viterbo, 2001: Land surface assimilation. ECMWF Meteorological training course lecture series (Printed in March, 2001), pp. 1-23.

Margulis, S. A., Entekhabi, D., 2001. A coupled land surface-boundary layer model and its adjoint, Journal of Hydrometeorology, 2(3), 274-296.

Nijssen, B., R. Schnur, and D. P. Lettenmaier, 2001: Global retrospective estimation of soil moisture using the variable infiltration capacity land surface model, 1980-1993. *J. Climate*, **14**, 1790-1808.

Noilhan, J., Planton, S., 1989. A simple parameterization of land surface processes for meteorological models. Mon. Wea. Rev., 117, 536-549.

Noilhan, J., and J.-F. Mahfouf, 1996: The ISBA land surface parameterization scheme. *Global and Planetary Change* **13,** 145-159.

Navon, I. and D. Legler, 1987: Conjugate-gradient methods for large-scale minimization in meteorology. *Mon. Wea. Rev.*, **115,** 1479-1502.

Pleim, J. E., and A. Xiu, 1995: Development and testing of a surface flux and planetary boundary layer model for application in mesoscale models. *J. Appl. Meteor*., 34, 16-32.

Qu, W., and coauthors, 1998: Sensitivity of Latent Heat Flux from PILPS Land-Surface Schemes to Perturbations of Surface Air Temperature. J. Atmos. Sci., **55**, 1909–1927.

Ren, D. and M. Xue, 2003: An improved force-restore model for land-surface modeling. *J. App. Meteor.***,** Under review.

Schaake, J., and coauthors, 2002: An intercomparison of North American LDAS soil moisture fields. Abstract Volume, Mississippi River Climate and Hydrology Conf., New Orleans, Louisiana.

Talagrand, O. and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. Part I: Theory. *Quart. J. Roy. Meteor. Soc.*, **113,** 1311-1328.

Xiu, A., Pleim, J., 2001. Development of a land surface model. Part I: Application in a Mesoscale Meteorological Model, J. Appl. Meteror., 40**,** 192-208.

Xu, Q., and B. Zhou, 2003: Retrieving soil moisture from soil temperature measurements by using linear regression. *Advances in Atmospheric Sciences*, Vol. 20, No.6 (in press).

Xue, M., Droegemeier, K. K., Wong, V., 2000. The Advanced Regional Prediction System (ARPS) - A multi-scale nonhydrostatic atmospheric simulation and prediction model. Part I: Model dynamics and verification. Meteor. Atmos. Phys., 75**,** 161.

Xue, M., K. K. Droegemeier, V. Wong, A. Shapiro, K. Brewster, F. Carr, D. Weber, Y. Liu, and D.-H. Wang, 2001: The Advanced Regional Prediction System (ARPS) - A multiscale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteor. Atmos. Physics*, **76,** 143-165.

Xue, M., and D. Ren, 2003: Testing of several recent modifications to ARPS land surface model, JP4.16, 18th Conference on Hydrology, 84TH AMS annual meeting, Seattle, WA 11-15 January 2004.

Ziegler, C. L., T. J. Lee, and R. A. Pielke, Sr., 1997: Convective Initiation at the Dryline: A Modeling Study. *Mon. Wea. Rev.*, **125,** 1001-1026.

Table 1. Norman OK site parameters during the period of 4-27 August 2000*

	$\overline{T_{sfc}}$ (K)	$\overline{T_{sfc}} - \overline{T_{dp}}$ (K)	A_0 (K)	A_{dp} (K)
(08/04-08/11)	306.04	5.49	9.7	0.51
(08/12-08/19)	306.70	4.70	13.5	0.96
(08/20-08/27)	307.60	5.46	13.8	1.20

veg	Rs_{min}(s m^{-1})	R_G(W m^{-2})	w_{fc} (m^3 m^{-3})	w_{sat} (m^3 m^{-3})	w_{wilt} (m^3 m^{-3})
0.75	100	30	0.326	0.45	0.22
b	C_{Gsat}(K m^2 J^{-1})	z_0(m)	d_1(m)	d_2 (m)	C_{GV}(K m^2 J^{-1})
4.9	2.56×10^{-6}	0.004	0.1	1.0	1.5×10^{-5}

*Infrared surface temperature data are used in calculating the soil temperature parameters (Notations follow Ren and Xue 2003). The notation for static soil and vegetation parameters follow NP89.

Table 2. List of initial state retrieval experiments using synthetic data

Initial guess for control variable	Initial *rms* error	J_{final}/J_0	# iterations needed	Retrieved initial condition			
				T_{sfc}	T_{dp}	w_{sfc}	w_{dp}
T_{sfc}+3.0K	0.310	0.153	17	306.76	301.23	0.25	0.280
-3.0K	0.315	0.65	40	305.80	301.14	0.25	0.280
T_{dp}+2.0K	0.154	0.25	12	307.16	303.16	0.29	0.280
-2.0K	0.157	0.3	11	307.16	299.16	0.23	0.280
w_{sfc}+0.01	0.042	9.7E-5	7	307.16	301.16	0.256	0.280
-0.008	0.035	3.4E-7	8	307.16	301.16	0.256	0.280
W_{dp}+0.01	1.2	8.7E-11	4	307.16	301.16	0.256	0.280
-0.01	0.94	2.4E-10	4	307.16	301.16	0.256	0.280

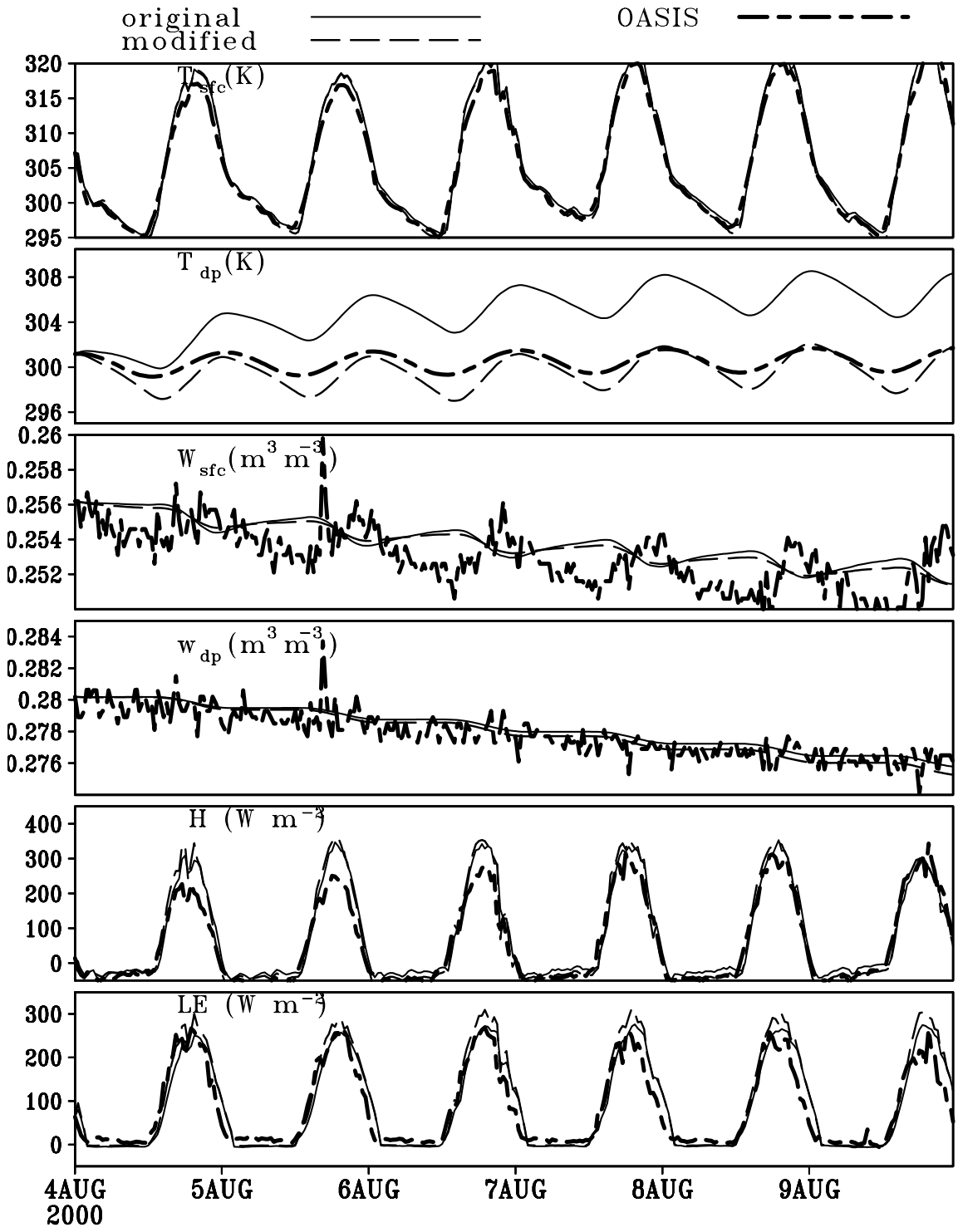

Fig. 1. Model predicted vs. observed (OASIS) soil temperature (T_{sfc} and T_{dp}), moisture (w_{sfc} and w_{dp}), and surface energy fluxes (H and LE). The curves labeled 'original' are for model predictions with the γ terms in Eq.(1), terms added by Ren and Xue (2003).

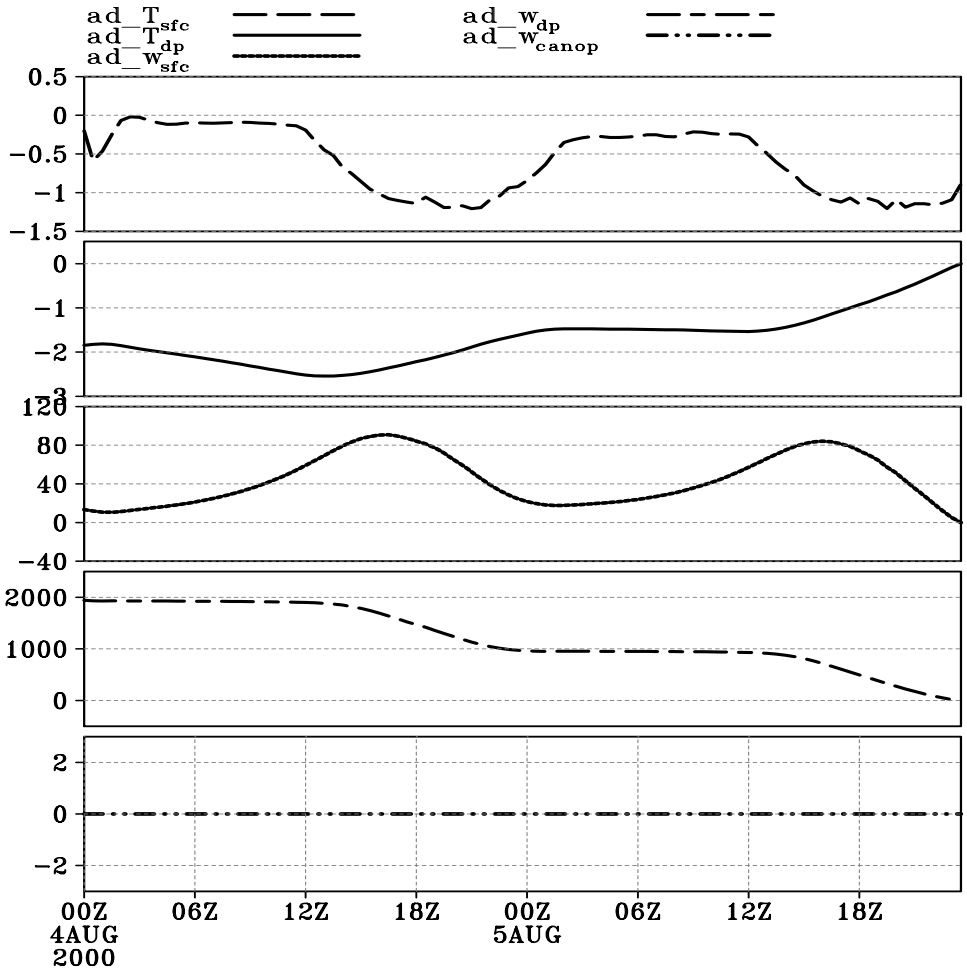

Fig. 2. The time series of adjoint variables from w_{dp}+0.025 at the first iteration of the minimization. Each adjoint variable is marked by its corresponding state variable with an *ad_* prefix.

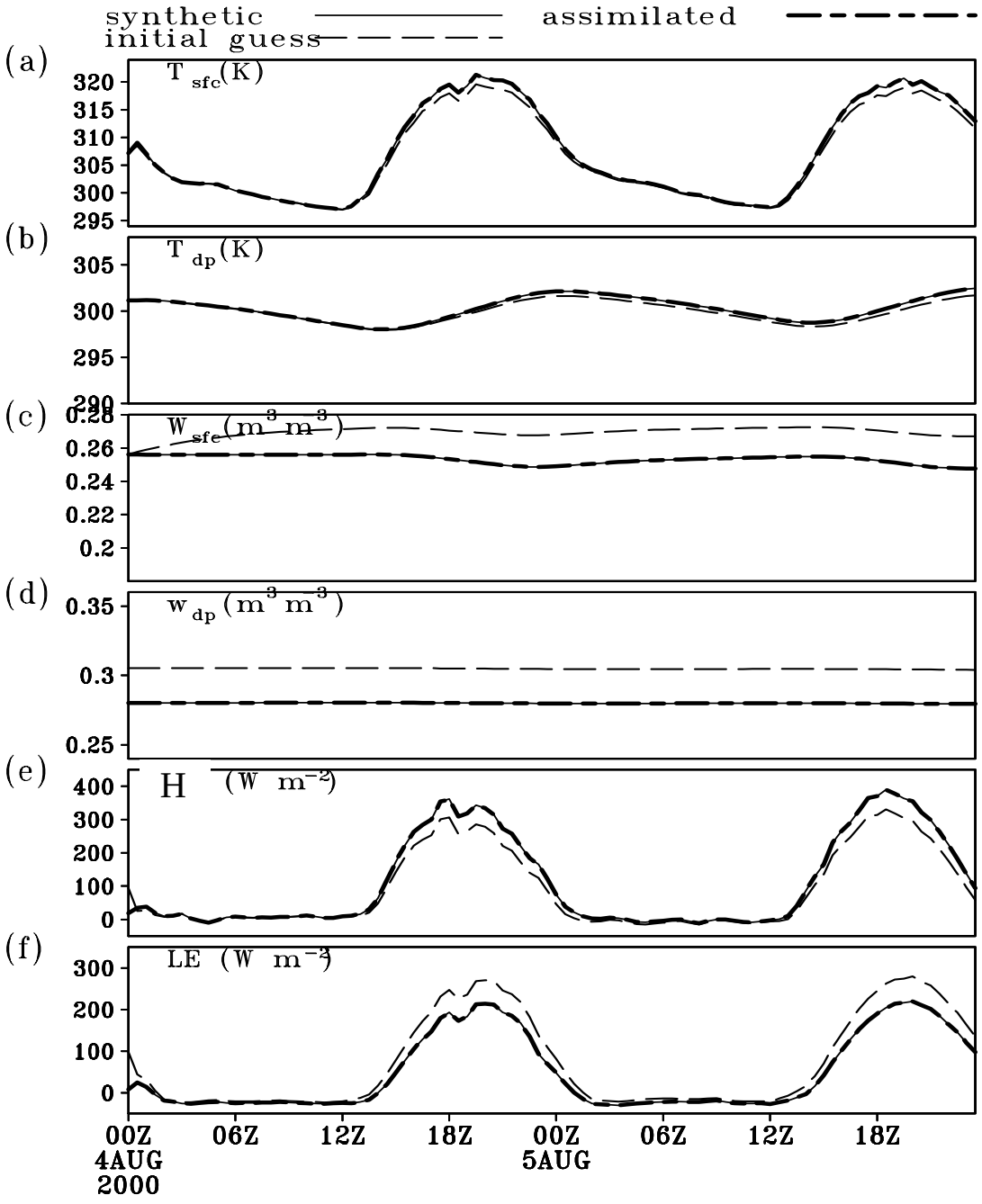

Fig. 3. Comparison of actual state variables and latent and sensible heat fluxes (synthetic) with those using prior (initial guess) and estimated using retrieved values (assimilated). Only ground surface temperatures are assimilated.

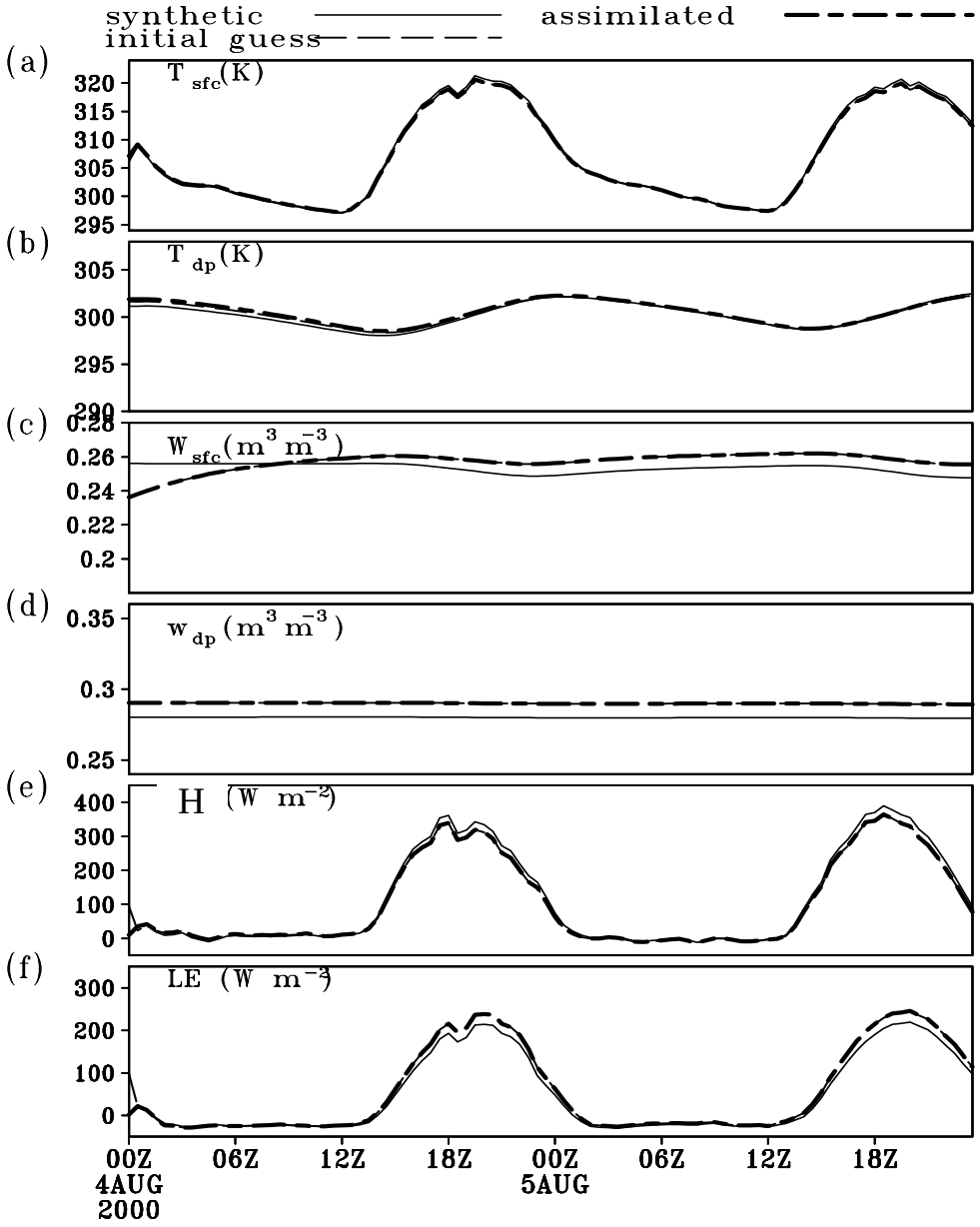

Fig. 4. Only temperature sensitivity channels are open for the assimilation. Although surface temperature time series are accurately retrieved, the surface fluxes are at systematic errors.

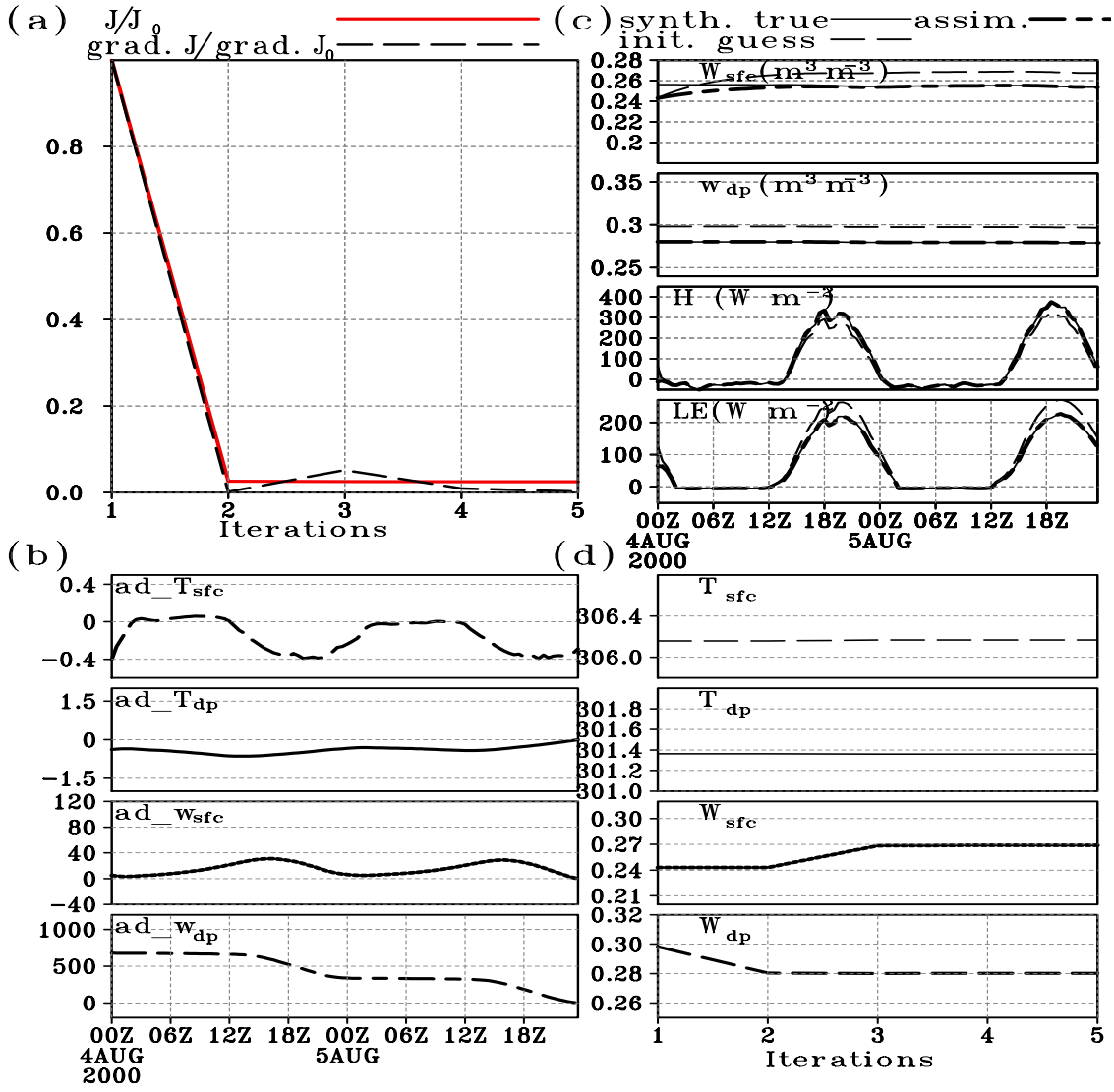

Fig. 5. The case of simultaneously perturbing the control variables. (a) reduction in cost function and the norm of gradient vector; (b) the time series of adjoint variables at the first iteration; (c) comparison between trajectories resulted from prior guess initial conditions and retrieved values; and (d) the convergence of the control variables to the "true" values. This experiment is done using an initial guess state vector: $U = \left(306.16\text{K}, 301.66\text{K}, 0.2362\ \text{m}^3\text{m}^{-3}, 0.2902\ \text{m}^3\text{m}^{-3}, 0.0\right)^T$.

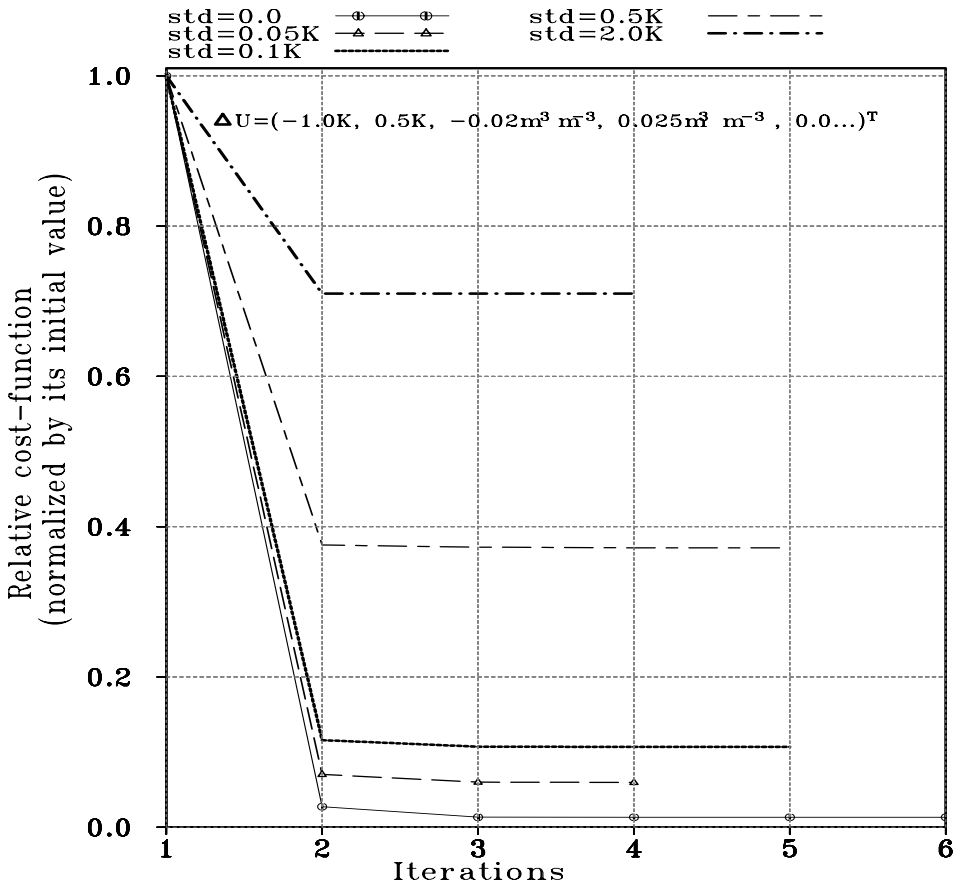

Fig. 6. Robustness of the assimilation scheme. Panel a is produced using δU =(-1.0K, 0.5K, -0.02 m^3m^{-3}, 0.025 m^3 m^{-3}, 0…)^T. The white noise is of 0.05, 0.1, 0.5, and 2.0 K, respectively.

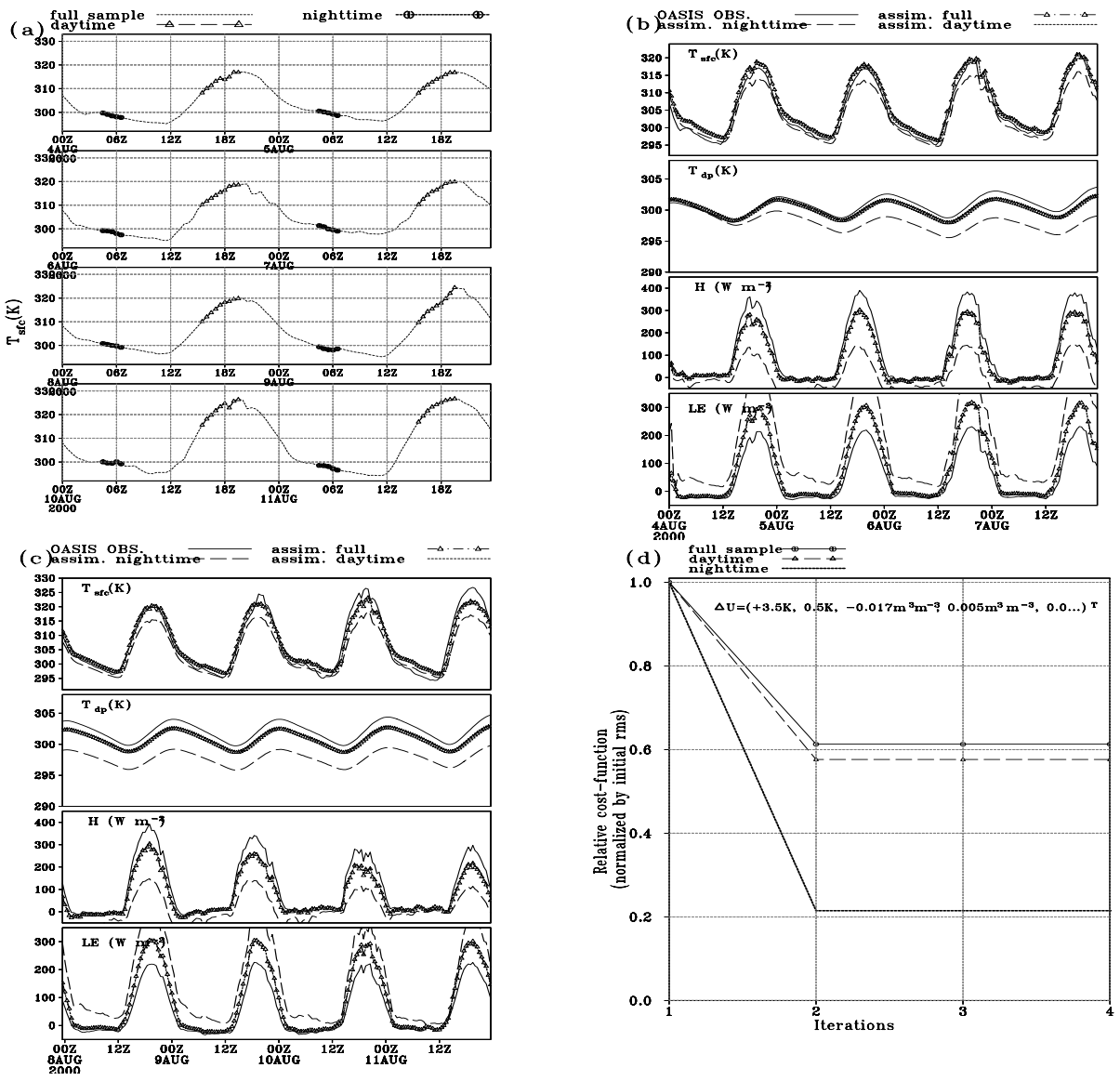

Fig. 7. Sampling strategy experiments. Panel (a) shows the three sampling scheme: The continuous sampling (dotted line), daytime (triangular sign), and the nighttime sampling (circle with vertical bar). Panels (b) and (c) show the respective accuracy in estimation soil temperatures and surface fluxes. Panel (d) indicates the decreasing of the cost-function for these three sampling schemes.

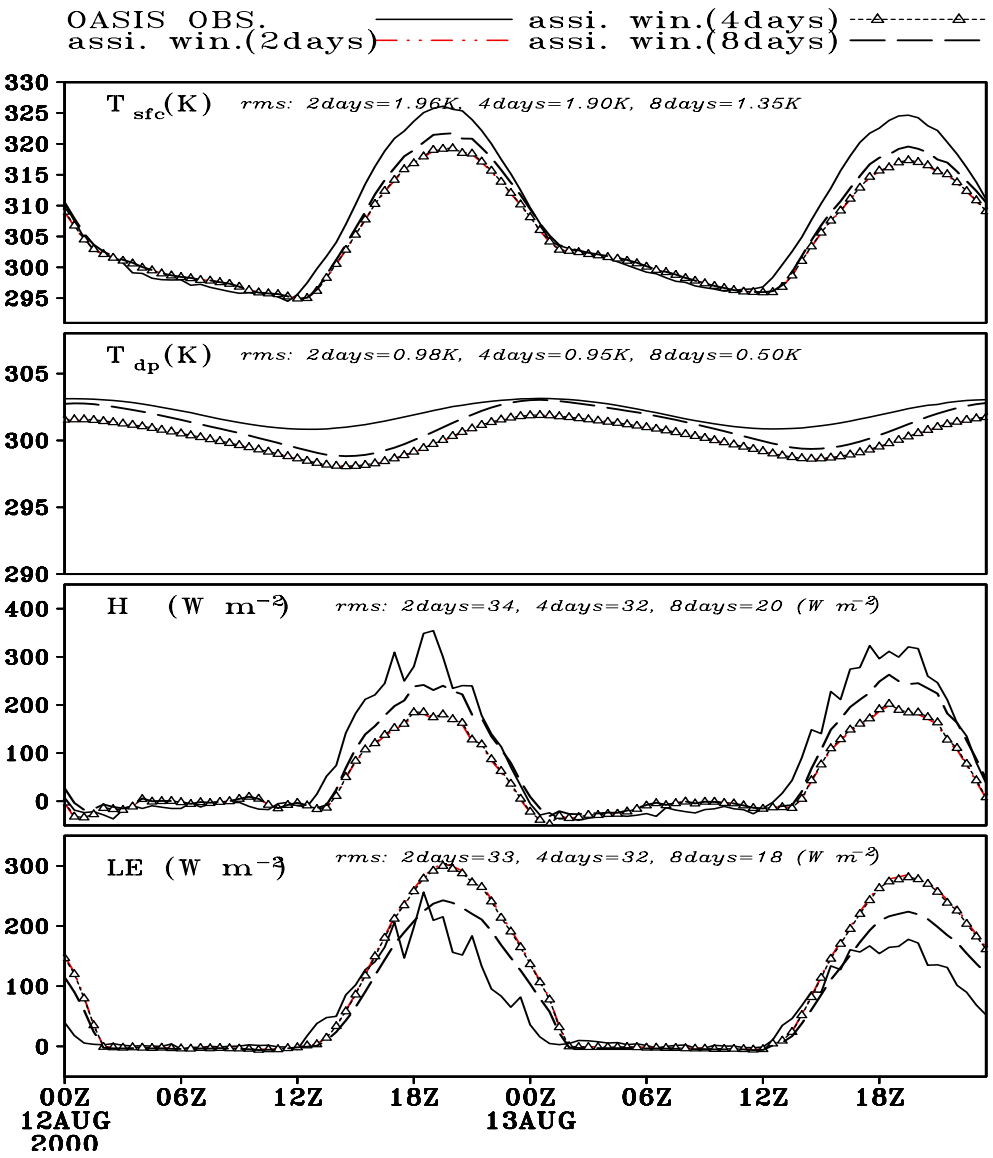

Fig. 8. The effects of assimilation window length. Three experiments with respective assimilation period of 2 (ass. win. 2 days), 4 (ass. win. 4 days), and 8 days (ass. win. 8 days), all starts from 00Z, 4 August 2000, are performed. For each experiment, using retrieved initial conditions, a forward model run is made to produce an extra forecasting 8 days long (forecasting periods follows immediately the assimilation periods). The shared forecast period, i.e., 00Z, August 12-14, is shown here. Real OASIS observations are used in these experiments for initial conditions retrieval. The initial guess errors: $\delta U = \left(1.0K, -0.5K, -0.03m^3m^{-3}, 0.01m^3m^{-3}, 0.0\right)^T$.