Assimilation of near-surface temperature using extended Kalman filter

Praveen Kumar *, Amy L. Kaleita

Environmental Hydrology and Hydraulic Engineering, Department of Civil and Environmental Engineering,
University of Illinois, Urbana, IL 61801, USA

Abstract

In order to improve the soil temperature profile predictions in land-surface models, an assimilation scheme using the extended Kalman filter is developed. This formulation is based on the discretized diffusion equation of heat transfer through the soil column. The scheme is designed to incorporate the knowledge of the uncertainties in both the model and the measurement. Model uncertainty is estimated by quantifying the model drift from observations when the model is initialized using the observed values. Furthermore, the initial error covariance has a significant influence on the performance of the assimilation scheme. It is shown that an inaccurate initial value for the error covariance can actually diminish the predictive capabilities of the model. When an appropriate initial error covariance is specified, using the top layer soil temperature observations in the assimilation scheme allows for improved predictive capabilities in lower layers. Observations at 30 min intervals have a significant effect on the model predictions in the lower layers. Assimilation of observations at 24 h intervals also has an effect on the lower layer predictive capability of the model, albeit more slowly than the 30 min assimilation scenario.

1. Introduction

Coupled land–atmosphere model simulations have substantially improved our understanding of the feedback between climate and terrestrial systems and enhanced our ability for short term weather and climate forecasts. However, since these models and their associated parameters are only an approximate representation of reality, the solutions tend to diverge from reality over a period of time. This can be due to either inherent biases in the model or the propagation of errors in the initial conditions due to non-linearities. If observations are available either periodically or intermittently, they can be inserted into the model to improve its performance. This process, called data assimilation, attempts to estimate the best state of a system at any time based on a combination of predicted state of the model and observations at that time [4]. This improved estimate of the model states provide better prediction. This technique has been very useful in atmospheric models due to the availability of data from a large network for meteorological observation operating at daily or shorter time scales. Data assimilation is now being exploited for terrestrial hydrology using the rapidly growing body of remote-sensing measurements [12].

Among the several variables that are involved in hydrological and climatological model studies, surface temperature and the accuracy of its prediction plays a crucial role in determining the predictive capability of such models. It influences the partitioning of incoming radiant energy into ground, sensible and latent heat fluxes. Outgoing longwave flux is a function of the surface temperature. In addition, surface temperature impacts the modeled snow cover on the ground during winter. A failure to properly predict the temperature state in the model can result in an inadequate characterization of snow accumulation and melt, thereby impacting the entire water and energy balance. A model with a tendency to over-predict the surface temperature can result in a delay in the formation of snow cover, and in early snow melt. During summer, the over-prediction in the surface temperature can result in excessive

* Corresponding author.
E-mail address: kumar1@uiuc.edu (P. Kumar).

0309-1708/03/$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S0309-1708(02)00098-2
evapotranspiration and convection near the surface which can impact the boundary layer development and subsequently the atmospheric dynamics. An under-prediction has a corresponding adverse effect.

In land–atmosphere models, the prediction of surface temperature is subject to several errors. Thermal diffusivity is a function of the soil texture and soil-moisture content. There is usually a large uncertainty associated with specifying soil texture at the scales at which these models are run. In addition, the soil-moisture dynamics are often inadequately represented, resulting in significant errors in the estimates of the soil-moisture states in the different soil layers, which in turn adversely impacts the thermal conductivity estimates. Given the importance of accurate representation of surface temperature in land–atmosphere interaction studies, it is necessary to utilize methodologies, such as data assimilation, that constrain the error growth resulting from these uncertainties.

Reliable observations of near-surface temperature can often be obtained from various satellite instruments, such as advanced very high resolution radiometer (AVHRR) using the split-window technique [9,13–15]. These observations can be assimilated into land-surface models (LSMs) to improve their predictive capability. Only a small body of research exists in the assimilation of near-surface temperature obtained from passive sensors [10] although schemes for joint assimilation of temperature and soil-moisture using microwave remote sensing have been proposed [5]. However, soil-moisture measurements with acceptable errors from space-based sensors are still not available. Given said lack of adequate soil-moisture observations, it is worthwhile to develop a scheme for assimilating temperature and soil-moisture estimates independently. Lakshmi [11] implemented a simple assimilation scheme for surface temperature to improve soil-moisture model estimates. A two-layer soil column was used, with assimilation being accomplished by a simple average of predicted and observed surface temperatures. While this approach is straightforward and applicable under a wide variety of circumstances, it does not allow for incorporation of the knowledge about the relative reliability of the model predictions and the observations. The Kalman filter algorithm, however, bases the mathematics of the assimilation on the relative error of each of the components.

Thus, in this paper we discuss an extended Kalman filter (EKF) algorithm for the assimilation of near-surface temperature measurements and discuss various implementation issues rather than assess the impact of the assimilation on the overall LSM performance. Since satellite observations provide temperature measurements of a thin surface layer, there are essentially two ways in which they can be incorporated into LSMS. The first is by specifying them directly into the surface energy balance equation. The second is by specifying them as the upper boundary condition in the solution of the vertical soil temperature profile. The latter approach is adopted here. This will be discussed further in subsequent sections.

This paper is organized as follows. In Section 2 the EKF scheme is developed. The data used in the simulation is described in Section 3. The results are discussed in Section 4 and summary and conclusions given in Section 5.

2. Algorithm

2.1. Background

Data assimilation [4] is a technique to minimize model forecast error by periodically incorporating observations and updating the model predicted states in such a way that the dynamic consistency in the conservation laws is maintained. In this study we use the Kalman filter approach for the assimilation of near-surface temperature measurements. In comparison to other methods (see [4] for details) the Kalman filter formulation is attractive due to its ability to correct the model states as well as provide an estimate of the associated error.

For this study we used the LSM developed by Bonan [1]. This model was designed to be a comprehensive model for land–atmosphere interaction studies, combining biogeochemical, hydrologic, biophysical, and ecosystem processes in a physically and biologically realistic manner. Running at sub-hourly time scales, it captures the diurnal cycle and models processes such as sensible and latent heat fluxes, soil heat transfer, infiltration and runoff, and soil water transfer. Required driving variables and model parameters as used in the current simulations are given in Table 1.

The various temperature layers in the model are shown in Fig. 1. The ground temperature is obtained by iteratively solving the surface energy balance equation. For non-vegetated surface this equation is given as

\[ -S_g + L_g(T_g) + H_g(T_g) + \lambda E_g(T_g) + M_g(T_g) + S_{in} = 0 \]  

(1)

where \( S_g \) is the net solar radiation absorbed by the ground, \( L_g \) is the longwave radiation, and \( H_g, \lambda E_g, M_g \)
and $S_n$ are the sensible, latent, ground and snow melt heat fluxes, respectively. All the heat fluxes are functions of the ground temperature $T_g$. For a vegetated surface, both ground and vegetation temperature ($T_v$) are obtained by simultaneously solving the energy balance for the vegetation and the ground given as:

$$-S_v + L_v(T_g, T_v) + H_v(T_g, T_v) + \lambda E_v(T_g, T_v) = 0 \quad (2)$$
\[-S_g + L_g(T_g, T_v) + H_g(T_g, T_v) + \dot{E}_g(T_g, T_v) + M_g(T_g, T_v) + S_h = 0 \tag{3}\]

In the above equations, $S_g$ is the solar radiation absorbed by the vegetation, and the subscript ‘v’ represent the variables related to the vegetation. The total flux is obtained by adding the individual components from vegetation and ground (for additional details see Bonan [1]). In both the above cases the ground heat flux is defined as

\[M_g = \frac{2k_i}{\Delta z_i} (T_g - T_i) \tag{4}\]

where $T_i$ is the temperature of the first layer soil temperature from the previous time step and $\Delta z_i$ and $k_i$ are the thickness of the layer and the thermal conductivity. Once the ground temperature $T_g$ is obtained, the ground heat flux provides the upper boundary condition for obtaining the temperatures in the various soil layers through the solution of the heat flux diffusion equation (see below).

One important issue associated with the assimilation of the satellite temperature estimates is determining the equivalent model temperature corresponding to the observations. As illustrated in Fig. 1, there are potentially two competing representations, namely, the ground temperature $T_g$ and the first soil layer temperature $T_i$. Since $T_g$ is obtained by solving the energy balance equations, its estimate is subject to errors associated with all the empirical parameters involved in the computation of latent, sensible and snow melt heat fluxes, and net solar radiation. These include soil and vegetation characteristics, and the resistances for momentum, heat and water vapor fluxes through the canopy and over bare ground. Consequently, it is very difficult to obtain an estimate of the model error associated with the ground temperature prediction. In the absence of this, an optimally estimated weighted average of the predicted and observed ground temperature cannot be specified. This problem is further compounded because of the iteration process for the solution of the energy balance equations (1), (2) and (3).

In contrast, the diffusion equation for the heat flux in the soil column offers a viable alternative because of its relatively simple formulation. In a recent study [9] found that AVHRR surface temperature estimates were strongly correlated with in situ ground temperature measurements at a depth of 5 cm. They found that the absolute mean bias and rms error was 1.25 and 2.7 K, respectively. This suggests that the first layer soil temperature in LSMs can be used for the assimilation, provided the thickness of the layer is sufficiently small. The description below provides the model formulation to accomplish this objective.

### 2.2. Model description

In LSM the dynamics of soil temperature are modeled using the diffusion equation for heat in a one-dimensional vertical column. The governing equation is:

\[c \frac{\partial T}{\partial t} = - \nabla \cdot (k \nabla T) \tag{5}\]

where $c$ is the volumetric soil heat capacity ($J \cdot m^{-3} \cdot K^{-1}$), $T$ is the soil temperature ($K$), $t$ is time, $k$ is the thermal conductivity of the soil ($W \cdot m^{-1} \cdot K^{-1}$), and $z$ is the soil depth. In order to develop the Kalman filter equations we discretize the vertical column into layers of unequal depth. The soil column is modeled as six layers with thicknesses $\Delta z_i$ of 10, 20, 40, 80, 160, and 320 cm sequentially from the surface. Soil thermal and hydraulic properties are defined at the center of each layer and are functions of the soil type and soil-moisture state. Following the formulation of the LSM model, we use a Crank–Nicholson scheme to arrive at a system of equations of the form

\[\frac{c \Delta z_i}{\Delta t} (T_i^{n+1} - T_i^n) = \frac{1}{2} \left( -3X_i - X_i + X_i^{n+1} + X_i^{n+1} \right) \tag{6}\]

where the superscripts $n$ and $n + 1$ indicate values at consecutive time steps, $\Delta t$ is the time step ($s$), and $X_i$ is the heat flux from layer $i$ to $i + 1$. Assuming that the heat flux across each layer is constant, $X_i$ can be derived as

\[X_i = -\left( \frac{T_i - T_{i+1}}{(\Delta z_i/2k_i) + (\Delta z_{i+1}/2k_{i+1})} \right) \tag{7}\]

The thermal conductivity is parameterized as a function of soil texture and soil-moisture content. Using the formulation of Farouki [6], the thermal conductivity, $k_i$, in the layer $i$ is specified as

\[k_i = \left[ k_s (1 - n_i) k_w - 0.15 \right] \frac{\theta_i}{\theta_s} + 0.15 \tag{8}\]

where

\[k_s = 8.8(\%sand) + 2.92(\%clay) \tag{9}\]

\[\theta_i = 0.489 - 0.00126(\%sand) \tag{10}\]

$\theta_i$ is the soil-moisture in the $i$th soil layer, $\theta_s$ is the saturated soil-moisture, $k_s$ is the saturated hydraulic conductivity of the soil, and $k_w$ is the thermal conductivity of water.

Rearranging Eq. (6), we arrive at the following equation:

\[\epsilon_1 T_{i-1}^{n+1} + \epsilon_2 T_i^{n+1} + \epsilon_3 T_{i+1}^{n+1} = -\epsilon_1 T_{i-1}^n + \epsilon_4 T_i^n - \epsilon_3 T_{i+1}^n \tag{11}\]
where
\begin{align}
\epsilon_1 &= -\frac{1}{m_1} \\
\epsilon_2 &= \frac{1}{m_1} + \frac{1}{m_2} + \frac{1}{m_3} \\
\epsilon_3 &= -\frac{1}{m_2} \\
\epsilon_4 &= -\frac{1}{m_1} - \frac{1}{m_2} + \frac{1}{m_3}
\end{align}
and
\begin{align}
m_1 &= \frac{\Delta z_{i-1}}{k_i} + \frac{\Delta z_i}{k_i} \\
m_2 &= \frac{\Delta z_i}{k_i} + \frac{\Delta z_{i+1}}{k_{i+1}} \\
m_3 &= \frac{\Delta t}{\Delta z_i}\epsilon_i
\end{align}

Grouping the above equation for all layers and adding a model error term \(w_i\) to the \(i\)th layer to account for uncertainty in the model and the parameters, we can combine Eq. (11) for all the layers in a vector form as
\[
F(T^{n+1}) = G(T^n) + w^n
\]
where
\[
T^n = \{T_i^n\}_{i=1,N}
\]
and
\[
F(T^{n+1}) = \{(-\epsilon_1 T_{i-1}^n + \epsilon_2 T_i^n + \epsilon_3 T_{i+1}^n)\}_{i=1,N}
\]
\[
G(T^n) = \{(-\epsilon_1 T_{i-1}^n + \epsilon_4 T_i^n - \epsilon_3 T_{i+1}^n)\}_{i=1,N}
\]
\[
w^n = \{w_i^n\}_{i=1,N}
\]
\[
E\{w^n\} = 0
\]
In the above equations, the \(T^n\) superscript denotes matrix transpose. The model error has the covariance
\[
E\{w^n w^T\} = Q^n
\]
which is assumed to be diagonal but not a multiple of the identity.

Once the value of \(T^{n+1}_i\) which balances equation Eqs. (1), or (2) and (3) has been determined, Eq. (4) is used to calculate \(M^n_{\Delta t}\) at the current time step. This ground heat flux is then used in the determination of the soil temperature profile. For the first soil layer \((i = 1)\), we assume a known heat flux into the soil, \(S^n_0 = M^n_{\Delta t}\). For the bottom layer \((i = N)\), we assume that the layer is sufficiently deep and therefore there is zero heat flux out of the layer, \(S^n_N = 0\).

### 2.3. State-space formulation

In this section we derive the dynamical evolution equation for the state vector \(T^n = (T_1^n, T_2^n, \ldots, T_N^n)^T\) and the error covariance matrix. Let \(\hat{T}^n\) denote the conditional mean of \(T^n\) based on all observations prior to time \(t = t_0 + n\Delta t\). This is the same as the predicted value at \(t = t_0 + n\Delta t\) based on all observations up to the prior time step. Let \(\tilde{T}^n\) denote the conditional mean of \(T^n\) based on all observations up to and including time \(t = t_0 + n\Delta t\). Note that this is the same as the optimally estimated value at \(t = t_0 + n\Delta t\) based on all observations up to the time step \(t = t_0 + n\Delta t\). Expressed mathematically,
\[
\hat{T}^n = E\{T^n|\text{all observations prior to } t = t_0 + n\Delta t\}
\]
\[
\tilde{T}^n = E\{T^n|\text{all observations up to and including } t = t_0 + n\Delta t\}
\]
In general, \(\hat{T}\) refers to the conditional mean of \(T^n\), a superscript \(n\) denotes the time step, and subscripts + and − refer to after and before assimilation, respectively.

The Taylor series expansion of \(F(T^{n+1})\) about the prior estimate \(\hat{T}^n\) is given as
\[
F(T^{n+1}) = F(\hat{T}^n) + DF(\hat{T}^n)[T^{n+1} - \hat{T}^n] + \text{higher order terms}
\]
where
\[
DF(\hat{T}^n) = \frac{\partial F}{\partial T^n}\bigg|_{T^n = \hat{T}^n}
\]
is the \(N \times N\) Jacobian matrix evaluated at \(\hat{T}^n\). Similarly \(G(T^n)\) can be expanded about the optimally estimated value \(\tilde{T}^n\) as
\[
G(T^n) = G(\tilde{T}^n) + DG(\tilde{T}^n)[T^n - \tilde{T}^n] + \text{higher order terms}
\]
where
\[
DG(\tilde{T}^n) = \frac{\partial G}{\partial T^n}\bigg|_{T^n = \tilde{T}^n}
\]
is the \(N \times N\) Jacobian matrix evaluated at \(\tilde{T}^n\). A careful inspection of Eqs. (28), (30) and (11) reveals that both \(DF(\hat{T}^n)\) and \(DG(\tilde{T}^n)\) are tri-diagonal matrices.

Substituting the above two equations into Eq. (19) and neglecting higher order terms we get
\[
F(\tilde{T}^n) + DF(\hat{T}^n)[T^{n+1} - \hat{T}^n]
\]
\[
= G(\tilde{T}^n) + DG(\tilde{T}^n)[T^n - \tilde{T}^n] + w^n
\]
On rearranging we get
\[
DF(\hat{T}^n)T^{n+1} = DG(\hat{T}^n)T^n + B(\hat{T}^n) + w^n
\]
where
\[ B(\tilde{T}^n) = G(\tilde{T}^n) - DG(\tilde{T}^n)\tilde{T}^n - F(\tilde{T}^n) + DF(\tilde{T}^n)\tilde{T}^n \]  
(34)

However, since in this case \( G \) and \( F \) are linear functions of \( T \), Eq. (34) can be simplified to
\[ B(\tilde{T}^n) = G(\tilde{T}^n) - DG(\tilde{T}^n)\tilde{T}^n = M^n \]  
(35)

where \( M^n \) accounts for the ground flux and can be expressed as \( M^n = (M^n, 0, \ldots, 0)^T \). Eq. (33) describes the dynamical evolution of the state in an implicit form. In the explicit form it can be written as
\[ T^{n+1} = [DF(\tilde{T}^n)]^{-1}DG(\tilde{T}^n)T^n + [DF(\tilde{T}^n)]^{-1}B(\tilde{T}^n) \]  
(36)

To derive the error covariance propagation equation, we first take the expected value of Eq. (28) and subtract it from Eq. (28), and noting that the a priori estimate \( E\{\tilde{T}^{n+1}\} = T^{n+1} \), we get
\[ F(T^{n+1}) - E\{F(T^{n+1})\} = DF(\tilde{T}^n)[T^{n+1} - \tilde{T}^{n+1}] \]  
(37)

Taking expected value of Eq. (19) and substituting the above, noting that the noise is uncorrelated with other terms, we get
\[ F(T^{n+1}) - E\{G(T^n)\} = DF(\tilde{T}^n)\{T^{n+1} - \tilde{T}^{n+1}\} \]  
(38)

When evaluating expected values at the \( n+1 \) time step, \( E\{G(T^n)\} = \tilde{T}^n \) and therefore we obtain
\[ F(T^{n+1}) - G(\tilde{T}^n) = DF(\tilde{T}^n)[T^{n+1} - \tilde{T}^{n+1}] \]  
(39)

Substituting Eq. (19) in the above gives
\[ G(T^n) - G(\tilde{T}^n) + w^n = DF(\tilde{T}^n)[T^{n+1} - \tilde{T}^{n+1}] \]  
(40)

Substituting from Eq. (30) and neglecting higher order terms we have
\[ DG(\tilde{T}^n)T^n - \tilde{T}^n] + w^n = DF(\tilde{T}^n)[T^{n+1} - \tilde{T}^{n+1}] \]  
(41)

Multiplying both sides of the above equation by their transpose and taking expected values we get
\[ DF(\tilde{T}^n)P_{n+1}^{-1}DF(\tilde{T}^n)^T = DG(\tilde{T}^n)P_{n+1}^{-1}[DG(\tilde{T}^n)]^T + Q^n \]  
(42)

or equivalently
\[ P_{n+1}^{-1} = [DF(\tilde{T}^n)]^{-1}[DG(\tilde{T}^n)P_{n+1}^{-1}[DG(\tilde{T}^n)]^T + Q^n] \]  
(43)

where the a priori and a posteriori error covariances are defined, respectively, as
\[ P_{n+1}^{-1} = E\{[T^{n+1} - \tilde{T}^{n+1}][T^{n+1} - \tilde{T}^{n+1}]^T\} \]  
(44)
\[ P_n = E\{[T^n - \tilde{T}^n][T^n - \tilde{T}^n]^T\} \]  
(45)

2.4. Extended Kalman filter algorithm

Let us assume that the model describing the measurement is given as
\[ Z^n = H^nT^n + v^n \]  
(46)

where \( v^n \) is the measurement error with covariance \( V^n = E\{v^n v^n^T\} \) and \( E\{v^n v^n^T\} = 0 \), i.e., there is no correlation between the system and the observation noise. In our case, when measurement in only the top soil layer is available, \( Z^n = [z^n_1, 0, \ldots, 0]^T \) is an \( N \times 1 \) matrix and \( H \) is an \( N \times N \) matrix having a value of 1 in diagonal elements for layers with available observation (here the first layer) and 0 elsewhere. The Kalman filtering algorithm proceeds as follows:

1. Initial conditions: Since no assimilation is performed at the initial time step, we start with the initial conditions
\[ \tilde{T}_0 = (T_0^0, \ldots, T_K^0)^T \]  
(47)
\[ P_0 = P_0 \] specified a priori covariance  
(48)

2. State estimate extrapolation: Solve for \( \tilde{T}^{n+1} \) from the equation
\[ \tilde{T}^{n+1} = [DF(\tilde{T}^n)]^{-1}DG(\tilde{T}^n)T^n + [DF(\tilde{T}^n)]^{-1}B(\tilde{T}^n) \]  
(49)

where \( DF(\tilde{T}^n), DG(\tilde{T}^n), \) and \( B(\tilde{T}^n) \) are defined in Eqs. (29), (31) and (34).

3. Error covariance extrapolation: Solve for \( P^{n+1} \) in the equation (see Eq. (43) above)
\[ P^{n+1} = [DF(\tilde{T}^n)]^{-1}[DG(\tilde{T}^n)P_n^{-1}[DG(\tilde{T}^n)]^T + Q^n] \times [DF(\tilde{T}^n)]^{-T} \]  
(50)

4. State estimate observational update:
\[ \tilde{T}_n = \tilde{T}^{n+1} + K^{n+1}[Z^{n+1} - H^{n+1}\tilde{T}^{n+1}] \]  
(51)

where the Kalman gain matrix is
\[ K^{n+1} = P^{n+1}H^nT^{n+1}[H^nT^{n+1}P^{n+1}H^nT^{n+1} + V^{n+1}]^{-1} \]  
(52)

5. Error covariance update:
\[ P^{n+1} = [I - K^{n+1}H^nT^{n+1}]P^{n+1} \]  
(53)

For an explanation of Eqs. (52) and (53) see [2]. As demonstrated through the simulations in the following sections, the above scheme is numerically stable.

3. Data description

We use the data collected during the SGP97 experiment [8]. The SGP97 was a field experiment carried out in the sub-humid environment of Oklahoma during the summer months of 1997, primarily to investigate retrieval algorithms for near-surface soil-moisture developed earlier [7] using the L-band electronically scanned thinned array radiometer instrument aboard an aircraft can be extended to coarser resolutions expected from
satellite platforms. Additional objectives of the experiment were to retrieve vertical profiles of soil-moisture and temperature using in situ and remote sensing data, along with modeling techniques, and to study their influence on the development of the boundary layer.

Meteorological and soil temperature measurements were available from 42 agricultural research service micronet stations in the Little Washita River Watershed. Site 151 (see Fig. 2) at the western edge of the Little Washita basin was used in this analysis. This site was chosen because it is the only site where the vertical profile of temperature is available for a fairly long duration along with the relevant meteorological observations. Input data for the LSM was taken from the Micronet data as well as meteorological and hydrologic data from SGP97 field measurements (see Table 1 for details).

Soil temperature measurements at depths of 3, 10, 20, 40, and 60 cm were available for this site at every half hour interval starting on June 15, 1997 (Julian day 166) and ending on July 9, 1997 (Julian day 190). These observed values, in the absence of another independent data set, are used for validation of the model performance in the lower layers. The top layer measurements serve as observational data for use in the EKF scheme. All simulations are performed using a single column of unit surface area centered at the coordinate of site 151.

4. Results and discussion

In order to investigate several aspects of the assimilation algorithm performance, simulation studies are performed for the following cases:

**Case 1:** LSM simulation, with no assimilation, using a reasonable guess for the initial condition. This case is used to quantify the model error which is subsequently used in the EKF assimilation algorithm.

**Case 2:** Assimilation using synthetic observations for the top soil layer. The synthetic observations are obtained by corrupting the top layer predictions from Case 1 with white noise of specified variance. The predictions of other layers are compared with that obtained in Case 1 to verify the validity of the algorithm.

Two sub-cases are used. In the first, an arbitrary initial error covariance (see Eq. (48)) is specified and in the second an “equilibrium” error covariance is specified as the initial guess. The equilibrium covariance is obtained by starting with the arbitrary guess and running the assimilation for a long period of time. These two sub-cases are used to assess the role of initial error covariance in the performance of the temperature assimilation.

**Case 3:** Assimilation using observations at every half hour and 24 h time steps. Observations at 5 cm depth are specified by interpolating the measurements obtained at depths of 3 and 10 cm (recall that first layer thickness is 10 cm). Measurement error of 0.5% ($\sigma=1.5$ K) was assumed. Observations at other depths are used for the validation of the assimilation scheme. This case is used to assess the impact of assimilation interval on the model performance.

**Case 4:** Study of Case 3 using a “bad” guess for the initial states of the temperature profile. This is used to study the effectiveness of the assimilation algorithm in the presence of errors in the initial conditions.

**Case 5:** Cases 1 and 2 rerun with a top layer of 1 cm. This is to study the functionality of the algorithm when a shallow top layer, analogous to a biomass layer seen by a satellite measuring surface temperature, is used.

4.1. Case 1: Quantifying model error

The EKF algorithm assumes that there is uncertainty associated with the model called model error (see Eq. (25)), and the current measurement, due to instrumentation error (see Eq. (46)). We define model error in a state variable as its departure from reality if the state variable is initialized with observed data. It is very difficult to quantify for several reasons. First, the model typically has many more degrees of freedom than observations, making it difficult to compare all model states with observations. Second, it is generally evolutionary or non-stationary, yet observations for quantifying the model error are not available as frequently as the model time step. Third, error in one particular state variable is correlated with several others, which results in unpredictable evolution of the errors. As described below, we have made the assumption of stationarity in
model error growth and ignored the correlation of the model error in temperature with other state variables.

For soil temperature the model error arises due to uncertainties in several things, including: representation of the temperature dynamics; parameter errors; and linearization and discretization of the dynamical equations. The model error plays a significant role in determining the performance of the assimilation and the uncertainty associated with the assimilated product. However, model errors are often specified arbitrarily with the assumption that their influence diminishes quickly during the process of incorporating the measurements (see Eq. (53)). However, this is largely dependent on the frequency of observations and the magnitude of the observational errors. Ideally, after a sufficiently long period of assimilation, the prediction errors should not be significantly larger than the observational error if model errors are small. However, when the observations are intermittent the prediction error is governed by the model error (see Eq. (50)).

One of the problems associated with the assimilation of remote sensing measurements is that data are often cloud contaminated and therefore cannot be used. This results in useful data being available at more sparse intervals than desired. This can be handled by assuming that those contaminated observations have large (infinite) errors. In these situations it is important to obtain proper estimates of the model error since it determines the error covariance.

This is achieved by comparing the deviations of the model prediction from the observations. Starting with an arbitrary initial condition the LSM was driven from May 10, 1997 (Julian Day 130) at a 5 min time step to provide a spin up period. The results of the last 24 days of the run, for which observations were available, are shown in Fig. 3 along with corresponding actual observations. As seen from this figure, the unmodified LSM gives predictions with an exaggerated amplitude in the top layer when compared to the observed data. While the model and the observations are in relatively close agreement for the nighttime phase of the diurnal cycle, the model estimates are considerably higher than observations during the daytime. Furthermore, the LSM over-predicts the soil temperature in lower layers compared to observations (note that observations corresponding to model layers 4–6 are not available).

In order to estimate the model error, the LSM was run again beginning on June 15, 1997 (Julian Day 166), the first day of the temperature profile measurements, and the temperature in the first three layers of the model were altered to correspond identically to the observed values for the first time step that these data were available. Fig. 4 shows the difference between model predictions and observations for a run of the last 24 days, at the beginning of which the initial temperature conditions in the top three layers were altered to correspond identically to the observed values, the period for which observations were available. In this comparison, a 5 cm observation was determined through linear interpolation of the 3 and 10 cm observations. For roughly the final 12 days of this run, the deviations remain relatively unchanged from day to day. Therefore, the deviations in

![Temperature profile from LSM](image-url)

Fig. 3. Temperature profile from LSM after the spin up period. Results for all six layers are shown with observed data for the first three layers.
the top three layers for the final 12 days of this run were used in quantifying the model error. These results were combined to create one average day's deviations. These deviations were then modelled as sine waves to capture the diurnal variation of the difference between model predictions and observations. Because the amplitude of the deviations decays markedly with depth, the bottom three layers are assumed to have constant deviation over time, similar to the magnitude of deviation in the third layer, for which data is available. This quantification of model error is then implemented as a time-variant standard deviation in the model error covariance (see Eq. (25)).

4.2. Case 2: Role of initial error covariance

In order to validate the EKF algorithm, studies using synthetic observations were performed. The top layer output values of soil temperature from the unmodified LSM run, as described in Case 1, were corrupted with 0.5% white noise. An error of 0.5%, on the order of 1.5 K, is similar to the magnitudes associated with using the split-window technique for surface temperature estimation. These were then used as top layer “observations” for the EKF mode. The purpose of using these synthetic observations was to see whether or not the assimilated values in the other five layers would converge to those in the unmodified LSM run.

Two situations are considered. First, an arbitrary initial error covariance is specified in the diagonal elements of the matrix and zero elsewhere. The simulation results are shown in Fig. 5. One can see that while the

![Graph](image-url)
top layer behaves the same as in the unmodified run with a slight equilibration period in which some degree of fluctuation occurs in the predictions, the deeper into the column one goes, the greater the magnitude of the initial fluctuations and the longer the equilibration time. The reason for the initial fluctuations in assimilated temperatures is inadequate specification of the error covariance. The diagonal matrix would indicate that errors in one layer are correlated only with errors in that same layer from one time step to the next. However, because the temperature profile propagates through the layer with time, it is reasonable to expect that the correlation structure is in fact more complicated. The fluctuations occur while this correlation structure is being numerically established during the assimilation period.

In order to test this proposition, in the second simulation we specified the equilibrium error covariance, arbitrarily scaled by a factor of 100, from the previous run as the initial guess. The equilibrium values were obtained by using the values of the error covariance after 60 days of model run when all the fluctuations had died out. It was also necessary to specify a climatological mean temperature for the bottom layer when computing the equilibrium error covariance. When this constraint was not used, the performance of the assimilation algorithm was severely degraded. The use of the equilibrium error covariance enables us to numerically capture the correlation structure between the different layers while the scaling factor of 100 indicates a large initial uncertainty. The results are displayed in Fig. 6. We see that the use of the “equilibrium” error covariance did not produce any initial fluctuations in the assimilated temperatures. This equilibrium error covariance is used as the initial guess for the assimilation of the observed temperature as described next. When this error covariance is used as the initial guess for the assimilation of observations, the constraint of climatological mean temperature for the bottom-most layer was not required.

4.3. Case 3: Assimilation at different temporal resolutions

Next, the EKF assimilation was run using actual top layer observations (observed temperature values at 3 and 10 cm depth interpolated to 5 cm depth). An assumed estimation error of 0.5% was specified. This corresponds to about 1.5 K, roughly the same magnitude as errors associated with AVHRR-derived surface temperature measurements over the SGP97 site [9].
First, observations at half-hour increments were used. Results are shown in Figs. 7 and 8. The results show noticeable improvement in the assimilated temperatures in the second and the third layer. However, the difference in the top layer between the unmodified LSM (no assimilation) and EKF assimilation is nearly undetectable. This may at first seem unexpected, since the observational information is being injected into the top layer. However, the fact that the top layer shows minimal response to the assimilation is most likely a function of the way the LSM is setup. In this model, the behavior of the top layer is tightly constrained by the surface energy balance component of the model, which does not seem to be significantly impacted by the assimilation. Despite the minimal response in the top layer, there is considerable change in the response of the lower layers when using the EKF scheme. This can be seen in Fig. 7 where lower layer predictions using the EKF differ from the unmodified LSM predictions and are closer to observed values.

Second, the observed data was pared down to a 24 h time step in order to simulate the best-case temporal scale of a polar orbiting satellite such as AVHRR. Results for the 1-day data assimilation are shown in Figs. 9 and 10. Here we again observe that the temperatures in the lower layers converge to the observations, albeit more slowly than the half-hour assimilation scenario. However, the top layer response seems to be still constrained by the surface energy balance computations.

4.4. Case 4: Impact of initial errors

It was found that the performance of the assimilation was not sensitive to the initial state of the temperature profile. A detailed study showed that the evolution of the temperature was relatively insensitive to the initial conditions. In fact the mean temperature variances tend to decrease over a period of time. The LSM was driven starting with a range of initial conditions (30 values within a range of $\pm 9$ K about the observed values on the first day) and the standard deviation of daily mean temperature in different layers was computed. Fig. 11 shows the result from this analysis. It is seen that the standard deviation of the model forecast decreases in time indicating the errors in in the initial conditions do not grow. This property makes the scheme quite attractive as the stable dynamics of the temperature evolution makes the assimilation scheme more robust. It is possible that errors in soil-moisture or other related parameters might have a different impact, but this was not investigated.

![Layer 1 Temperatures](image1)

![Layer 2 Temperatures](image2)

![Layer 3 Temperatures](image3)

Fig. 7. Comparison of model, assimilated and observed temperatures (K) for the first three layers. The assimilation is performed using the top layer observations at 30 min interval.
4.5. Case 5: Shallow top layer

In order to test the theoretical functionality of using a very shallow top layer, analogous to the shallow biomass layer visible to a satellite, Cases 1 and 2 were rerun with a 1 cm thick top layer. The results indicate that the assimilation scheme performs in the same manner as when the top layer were thicker. When the assimilation scheme was implemented with an arbitrary error covariance, there were undesirable fluctuations in the predictions, but when the resulting equilibrium error covariance was used, the fluctuations were minimal (on the order of 0.1 K in the lowest layer). Note that using this new configuration necessitated determination of a new equilibrium error covariance. Since there were no observations available suitable to assimilation with this 1 cm top layer, no further analysis of this situation was performed.

5. Summary and conclusions

In order to improve the soil temperature profile predictions in LSMs, an assimilation scheme using the EKF is developed. This formulation is based on the discretized diffusion equation of heat transfer through the soil column. The scheme is designed to incorporate the knowledge of the uncertainties in both the model and the measurement. Model uncertainty, which plays an important role in the evolution of the error covariance when there are missing observations, is estimated by quantifying the model drift from observations when the model is initialized using the observed values.

In addition to the model error, the initial error covariance has a significant influence on the performance of the assimilation scheme. It is shown that an inaccurate initial value for the error covariance can actually diminish the predictive capabilities of the model. When an arbitrary diagonal matrix was used as the initial error covariance in this study, the results showed unreasonable amounts of fluctuations before equilibrating. This is clearly an undesirable effect. However, when the model was permitted to run for a certain amount of time and the resulting equilibrium error covariance, scaled to account for larger initial error, was used as the initial condition for a restarted run, the fluctuations were almost non-existent. This is a consequence of the fact that the correlation structure between different layers is more
complicated than can be described by a simple diagonal matrix.

In the case where measurements are available every 30 min, the lower layers respond much more rapidly to the infusion of observed data than in the case where the measurements occur every day. For the case of 24 h observations, the lower layers respond considerably slower. The ability of the assimilation to improve lower layer predictions based upon top layer observation is consistent with the experiences in the assimilation of soil-moisture observations [3,5,16]. The success of the assimilation in constraining the model temperatures and preventing a drift has significant implications for both short-term and seasonal-to-interannual predictions. While 24 h observations are the best-case scenario from polar orbiting satellites, due to factors such as cloud contamination, a longer interval between observations is probably more likely. In a study by Walker [16], researchers found that a 5 day interval between soil moisture observations yielded poor results in improving model predictions. If actual AVHRR surface temperature estimates were to be used as observations, it is likely that the frequency of good observations would be less than 1 every 24 h. Alternatively, data from a geostationary satellite such as GOES, which provides more or less continuous coverage, could be used. The tradeoff, however, is that the spatial resolution of geostationary satellites currently in wide use is lower than that of AVHRR.

The choice of the first layer soil temperature for the assimilation of surface temperature is a moot issue. It is argued that the ground temperature, which is computed by iteratively solving the surface energy balance, is not the most appropriate choice. This is because the errors associated with this state variable can be quite large due to the errors in several empirical parameters associated with the energy balance components. In contrast it is found that the soil temperature diffusion equation exhibits stable dynamics where the errors do not grow indefinitely. We believe that the proposed scheme can be used more effectively if the thickness of the first layer is made even shallower, computational cost permitting.

We have not dwelled on the computational aspects of the algorithm. The emphasis has been on gaining insight about the various factors impacting the temperature assimilation. Operational implementation for short term weather and climate prediction will involve coupling the scheme with those for several other variables and the computational issues should be discussed in this broader context.
Fig. 10. Same as Fig. 9 but comparison of model and assimilated temperatures for the bottom three layers. No observations are available for these layers.

Fig. 11. Daily standard deviation of model predictions for each layer based upon 30 different initial conditions ranging from 9 to –9 K about the observations on the first day.
Acknowledgements

Support for this project has been provided by the National Science Foundation Graduate Student Fellowship and NASA Grants NAGW-5247 and NAG5-7170. We also thank the SGP97 team for collecting and providing the observations used in the research.

References