Improving the estimation of hydrothermal state variables in the active layer of frozen ground by assimilating in situ observations and SSM/I data

JIN Rui & LI Xin†

Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences, Lanzhou 730000, China

The active layer of frozen ground data assimilation system adopts the SHAW (Simultaneous Heat and Water) model as the model operator. It employs an ensemble kalman filter to fuse state variables predicted by the SHAW model with in situ observation and the SSM/I 19 GHz brightness temperature for the purpose of optimizing model hydrothermal state variables. When there is little water movement in the frozen soil during the winter season, the unfrozen water content depends primarily on soil temperature. Thus, soil temperature is the crucial state variable to be improved. In contrast, soil moisture is heavily influenced by precipitation during the summer season. The simulation accuracy of soil moisture has a strong and direct impact on the soil temperature. In this case, the crucial state variable to be improved is soil moisture. One-dimensional assimilation experiments that have been carried out at AMDO station show that land data assimilation method can improve the estimation of hydrothermal state variables in the soil by fusing model information and observation information. The reasonable model error covariance matrix plays a key role in transferring the optimized surface state information to the deep soil, and it provides improved estimations of whole soil state profiles. After assimilating the 4-cm soil temperature by in situ observation, the soil temperature RMSE (Root Mean Square Error) of each soil layer decreased by 0.96 °C on average relative to the SHAW simulation. After assimilating the 4-cm soil moisture in situ observation, the soil moisture RMSE of each soil layer decreased by 0.020 m³·m⁻³. When assimilating the SSM/I 19 GHz brightness temperature, the soil temperature RMSE of each soil layer during the winter decreased by 0.76 °C, while the soil moisture RMSE of each soil layer during the summer decreased by 0.018 m³·m⁻³.

The Qinghai-Tibet Plateau is the highest and largest mountain permafrost region in the low-middle latitudes. The area of permafrost is approximately 1.3×10⁶ km²[1], which covers 48% of the total Qinghai-Tibet Plateau area; the seasonally frozen ground is distributed throughout the entire plateau. The frozen ground, an indispensable component of the cryosphere, interacts closely with the climate system. Not only is the frozen ground widely distributed, but it also has special hydrothermal properties, which have a significant impact on the global climate system through the energy and water exchange between land surface and atmosphere, and on the hydrological cycle and monsoon circulation. On the other hand, the frozen ground on the Qinghai-Tibet Plateau is characterized by high ground temperature, thin thickness, shallow active layer, and instability, which displays a continuous degradation trend with climate warming and can be taken as a sensitive indicator for

Citation: Jin R, Li X. Improving the estimation of hydrothermal state variables in the active layer of frozen ground by assimilating in situ observations and SSM/I data. Sci China Ser D-Earth Sci, 2009, 52(11): 1732-1745, doi: 10.1007/s11430-009-0174-0

†Corresponding author (email: lixin@lzb.ac.cn)

Received December 26, 2008; accepted April 14, 2009
doi: 10.1007/s11430-009-0174-0

Supported by Knowledge Innovation Program of the Chinese Academy of Sciences (Grant No. KZCX2-YW-Q10-2-4), National Natural Science Foundation of China (Grant No. 40701113) and “Western Light” Program of Talent Cultivation of the Chinese Academy of Sciences (2008)
climate change\cite{2-4}. The active layer of the frozen ground is not only the boundary of interaction between the atmosphere and frozen ground but also the soil layer, wherein the majority of the land surface processes take place. In recent decades, many studies have shown that a significantly changed active layer thickness has a profound influence on surface runoff, soil moisture, evapotranspiration, growth period, and engineering stability\cite{5}.

The basic methods used in the present study on the interaction of frozen ground and climate include observation (i.e., \textit{in situ} observation and remote sensing observation) and model simulation. The frozen ground primarily develops in cold and inaccessible regions, especially on the Qinghai-Tibet Plateau. However, the \textit{in situ} stations on the Qinghai-Tibet Plateau are sparsely and unevenly distributed. Most of these stations are located at the eastern margin of the plateau, or along the Qinghai-Tibet road and railway. There has been no long-term continuous observation at the northwestern plateau until now. Even though the \textit{in situ} stations are densified and strengthened, they remain sparse relative to the heterogeneity of land surface and the extensive region of frozen ground, which are difficult to apply in the spatial distribution and change research of the frozen ground.

Remote sensing provides a potential technique for regional frozen ground monitoring and research. The direct application of remote sensing is to detect the surface soil freeze/thaw status by the corresponding microwave frequency\cite{6,7}. However, the microwave is only sensitive to the surface soil state and cannot provide quantitative information about deep soil. On the other hand, the model simulation is necessary to understand the dynamic response processes of frozen ground to climate change, especially by using physically based land surface models. It is forced by the lower boundary conditions of the atmosphere and describes the heat exchange and water transfer at the land surface and in the soil based on energy and mass balance. It is important to consider the frozen soil parameterization scheme in the land surface model. The lack of soil freeze/thaw process results in a high degree of uncertainty in soil moisture simulations, overestimated evapotranspiration, overestimated daily amplitude of soil temperature, and amplified cooling effects of soil during the winter\cite{8}. Significant attention has been paid to the frozen soil parameterization and the influence of soil freeze/thaw processes on energy and mass transfer in land surface models since the 1990s, including CoLM (Common Land Model) and the SHAW model. Limited by computational efficiency and process understanding, the existing land surface models have simplified and abstracted the true physical process. Furthermore, the model parameters and forcing data bear uncertainties. As such, the model error will be accumulated, and the uncertainty will be enhanced as the model is run over time.

Both observation and model simulation have their own advantages and disadvantages. Model imperfections and model parameter uncertainties conflict with the requirement for high spatial-temporal resolution and high accuracy land surface datasets required by scientific research. Land data assimilation was started in the 1990s, and it offers innovative methods with which to integrate model simulation and observation in order to make the best of multi-source and multi-resolution observation to control model error accumulation and obtain optimized state variables estimation\cite{9,10}.

The current operational land data assimilation systems include GLDAS (Global Land Data Assimilation System), NLDAS (North American Land Data Assimilation System), ELAS (European Land Data Assimilation System) and CLDAS (Chinese Land Data Assimilation System)\cite{11,12}. The land data assimilation method was primarily applied to soil moisture estimation in the thawed soil according to the published papers: (1) comparing and analyzing the efficiency and optimization results of various assimilation algorithms\cite{11-15}, (2) retrieving the soil moisture profile by assimilating the \textit{in situ} surface soil moisture observation and the sensitive analysis of parameters used in the assimilation system\cite{16-19}, (3) retrieving the soil moisture profile by assimilating the passive microwave brightness temperature, such as carrying out the ideal experiment using artificial data produced by the forward microwave radiative transfer model\cite{20,21}, and assimilating real airborne or satellite-borne remote sensing observations\cite{22-24} or (4) retrieving the soil moisture profile by assimilating the active microwave backscattering coefficient\cite{25}. All of the above studies deal with one-dimensional experiments. There are few publications about land data assimilation used in frozen ground research. Additionally, there are a few ongoing projects, such as that of the group led by England and DeRoo\cite{26}, which has devel-
oped a land data assimilation system, adopting SVAT (Soil-Vegetation-Atmosphere Transfer) as the system framework and assimilating 1.4 and 6.9 GHz brightness temperatures to obtain the daily thickness and soil moisture of the active layer in the arctic tundra region. Tian and Xie has developed a land surface soil moisture data assimilation framework that considers the model subgrid-scale heterogeneity and soil water thawing and freezing. However, this research has only realized the assimilation of in situ observations.

This paper develops a land data assimilation system for the active layer of frozen ground (ALDAS), which adopts the SHAW model as the system dynamical framework to simulate coupled water-thermal-solution processes in both frozen and thawed soil. Furthermore, it aims to improve the estimation accuracy of model state variables by assimilating in situ observations and the SSM/I 19 GHz brightness temperature through the Ensemble Kalman Filter (EnKF) algorithm. The one-dimensional assimilation experiments have been carried out to test the performance of ALDAS, and they provide a foundation on which to produce soil temperature, soil moisture, and ice content datasets over the Qinghai-Tibet Plateau with spatial-temporal consistency and physical consistency by assimilating multi-source remote sensing observations. By merging the model simulation and remote sensing observation, the advantages of ALDAS can be identified as following:

1. Introducing the microwave remote sensing observation that is sensitive to surface soil moisture and soil temperature into the land surface model optimizes the surface soil state variables. Furthermore, a reasonable model error covariance matrix describing the correlation between state variables of each layer combined with energy and mass transfer process between surface soil and deep soil also improves the state variable estimations of the deep soil.

2. Microwave remote sensing can only monitor instantaneous surface states. However, the soil freeze/thaw cycle is a continuous process and may occur many times in several days. The revisit period of on-orbit radiometers makes it difficult to capture the soil state transition, especially in the spring and autumn seasons. Depending on the time integration process of the land surface model, the instantaneous remote sensing observation can be converted to continuous information.

3. Microwave remote sensing can only penetrate several centimeters into the frozen soil and even less for thawed soil due to the absorptions effect of liquid water. By means of a model error covariance matrix, the optimized surface soil information can be transferred to deep soil to achieve an overall improvement in soil temperature and soil moisture profiles.

4. Microwave remote sensing can currently only detect surface soil freeze/thaw states and moisture. By integrating with physically based land surface models, the quantitative information about unfrozen water content and ice content in the frozen soil can be obtained, and the accuracy of the soil freeze/thaw state dataset can be improved.

1 Structure of the active layer data assimilation system

The ALDAS is composed of a model operator, observation operator, assimilation algorithm, and model dataset (Figure 1). The land surface model is adopted as the model operator, which provides a physically based dynamic framework for the active layer data assimilation system and predicts the soil temperature, soil moisture, and ice content of each soil layer at hourly or daily time intervals. The role of the observation operator is to convert the state variable prediction of the land surface model to remote sensing observation variables, such as brightness temperature for passive microwave remote sensing. The role of the assimilation algorithm is to fuse the model simulation information and remote sensing observation information according to the weights of model error and observation error, and then to update the model state variables using the optimized values of current time steps.

1) The SHAW (Version 2.3.6) is a state-of-art land surface model, which performs well in simulating soil freeze/thaw process. A key feature of SHAW, distinct
from other land surface models, is its ability to take into account the influence of residue and solution on soil freeze/thaw processes, except for the influences of canopy, snow, and soil, which are suitably accounted for in the hydrothermal simulation in cold regions\textsuperscript{[28–29]}. Through forward integration, the SHAW model predicts the soil temperature, soil moisture content, and ice content of each soil layer at each time interval.

2) When assimilating in situ soil temperatures or soil moisture observations, the observation operator is the identity vector; when assimilating the passive microwave remote sensing observation, the observation operator is the forward microwave radiative transfer model, which is used to convert the state variables prediction into the simulated brightness temperature. The microwave radiative transfer model in the present paper assumes the soil to be a semi-infinite medium. The emitted soil brightness temperature was calculated by multiplying the surface emissivity by the effective emission temperature (eq. (1)). The soil surface emissivity $e$ was obtained by the Land Surface Process/Radiobrightness (LSP/R)\textsuperscript{[33]} according to eq. (2):

$$T_b = e \cdot T_{\text{eff}},$$

$$T_{\text{eff}}(t) = T_g(0, t) + \frac{1}{\kappa_e \sec \theta_t} \left( \frac{\partial T_g(z, t)}{\partial z} \right)_{z=0},$$

where $\kappa_e$, $\theta_t$ and $T_g(z, t)$ represent the soil extinction coefficient, Snell refraction angle, and soil temperature at time $t$ and depth $z$, respectively. The soil temperatures of the two layers required by eq. (2) were derived from the 0-cm and 4-cm soil temperature predictions of the SHAW model. The Dobson model was used to calculate the soil complex dielectric constant\textsuperscript{[34]} through the 4-cm soil temperature, moisture content, and ice content prediction of the SHAW model.

3) The Ensemble Kalman Filter algorithm (EnKF), as a sequential data assimilation method based on Monte Carlo integration, has the advantage of handling nonlinear models, and it is widely used in land data assimilation research\textsuperscript{[35,36]}. The EnKF is model independent, so there is no need to develop TLM (Tangent Linear Model) or model adjoint, and all the dynamic characteristics in the land surface model can be maintained. In addition, the computational cost of EnKF is comparable to other highly efficient approaches such as variational methods if the ensemble member is between 0 and 100\textsuperscript{[11]}. The calculation of EnKF includes an ensemble forecast step and ensemble analysis step.

Ensemble forecast step:

$$X^f_i(t_k) = M(X^a_i(t_{k-1})), \quad i \in [1, N],$$

where $X^i$ is the $i$th element of model state ensemble $X^a$, $M$ is model operator that is the SHAW model in this paper, superscripts $f$ and $a$ represent forecast field and analysis field, $t_k$ is time and $k$ is the subscript of time, $n$ is the dimension of model state vector, and $N$ is the number of ensembles. The error covariance matrix of forecast field is calculated as

$$P^f(t_k) = MP^a(t_{k-1})M^T, \quad P \in \mathbb{R}^{n \times n},$$

where, $P^f$ is the error covariance matrix of analysis field at last time interval.

Ensemble analysis step:

First, application of the observation operator $H$ converts the model state vector to the estimated observation vector

$$Y_i(t_k) = H_i[X^f_i(t_k)], \quad Y_i \in \mathbb{R}^{m \times 1},$$

where, $Y_i$ is the $i$th element of estimated observation vector $Ye$, and $m$ is the dimension of observation vector. Next, the following equations can be obtained according to the definition of ensemble\textsuperscript{[35,36]}:

$$P^f H^T = X^a (Y^e)^T 
\left( \begin{array}{c} X^a \end{array} \right) N^{-1} \in \mathbb{R}^{m \times m},$$

$$HP^f H^T = Y^e (Y^e)^T 
\left( \begin{array}{c} Y^e \end{array} \right) N^{-1} \in \mathbb{R}^{m \times m},$$

where, $X^e$ and $Y^e$ represent the increments of $X^a$ and $Ye$, respectively. Then, the Kalman gain can be obtained:

$$K = P^f H^T (HP^f H^T + R)^{-1} \in \mathbb{R}^{m \times m},$$

The analysis equations for the ensemble of the updated model state and covariance matrix of analysis field are

$$X^a = X^f + K(Y^e - Y^e),$$

$$P^a = \left( X^a - X^a \right) \left( X^a - X^a \right)^T N^{-1},$$

where $Ye$ is the ensemble of measurement vectors.

4) The dataset comprises the atmospheric forcing data, land surface parameters, and observation data. The atmospheric data used to force the SHAW model include air temperature, wind speed, relative humidity, precipitation, and shortwave radiation, which are all obtained from in situ automatic weather station observations. Due to a lack of precipitation data during the winter, the 25-km resolution dynamic downscaling result of
the NCEP (National Centers for Environmental Prediction) reanalysis data by the Newton relaxation method was adopted as a replacement[11]. The validation of air temperature, humidity, and wind field indicated that the downscaling result was better than both the objective analysis result and the NCEP reanalysis data. However, the precipitation accuracy has not been evaluated due to a lack of ground-based observations. On the other hand, the AMDO region is controlled by the monsoon climate, and the precipitation primarily occurs from July to September. There were only two snowfall events during the simulation period, which rapidly melted in 2 to 3 days according to surface albedo data. Furthermore, due to the shallow snow at the AMDO region, the downscaling result of the NCEP data did not have a large influence on the model simulation and assimilation results.

The land surface parameters include soil texture, canopy height, leaf area index, effective root depth, saturated water potential, and saturated hydraulic conductivity so on. All parameters were collected from published references and then manually calibrated by comparing model simulations with in situ observations.

The in situ Soil Moisture and Temperature Measurement System (SMTMS) observations were obtained by the GEWEX Asia Monsoon Experiment on the Qinghai-Tibet Plateau (GAME-Tibet) during the summer of 1998 and by the Coordinated Enhanced Observation Period (CEOP) during the winter of 2002/2003. The SMTMS includes ten layers of soil temperature observation at 4, 20, 40, 60, 80, 100, 130, 160, 200, and 279 cm depth, and six layers of soil moisture observation at 4, 20, 60, 100, 160, and 258 cm. In order to make sure that the lower boundary conditions were available, the lowest soil layer in SHAW model was set to 258 cm, and the soil temperature at 258 cm was assumed to equal to that at 279 cm. The SMTMS data are not only the assimilated observation variables but also the “true value” to validate and evaluate the model simulation and assimilation results. The Level 3 daily SSM/I brightness temperature products were provided by the National Snow and Ice Data Center (NSIDC) and formatted in an EASE-Grid with 25 km resolution.

2 One-dimensional assimilation experiment scheme

2.1 Research region

The one-dimensional assimilation experiments were conducted at the AMDO station (32.241°N, 91.635°E; 4700 m), which is located in the seasonally frozen ground region on the Qinghai-Tibet Plateau. The AMDO region is characterized by the east-Asian monsoon season and a semi-humid climate. The land surface is dominated by the alpine meadows with 5-cm height; the vegetation coverage is near 70% during the summer, and the soil texture is sandy[17]. There is permafrost or island permafrost in the region. The depth of the permafrost is over 3 m, and the maximum frozen depth is approximately 3.5 m[38]. The annual average air temperature is −2.9°C, while the annual precipitation is 420 mm and generally occurs during the period from July to September and comprises 85% of total annual precipitation.

2.2 Design of assimilation experiments

First, the observation errors of the SMTMS, SSM/I, and the SHAW model simulation error were estimated to provide prior information for the ALDAS. Next, the one-dimensional experiments assimilating SMTMS 4-cm soil temperature and soil moisture observation were carried out respectively to test the performance of the ALDAS and to determine the reasonable model error covariance matrix for further assimilating the SSM/I microwave brightness temperature. The land data assimilation method not only improved the surface soil state variable estimation but also optimized the state variable profiles through a reasonable model error covariance matrix[15,17,18]. We have designed an experiment to compare the assimilation results of different model error covariance matrices for the soil temperature profile. The covariance was set equal to 0 in one case and was obtained by correlation analysis in another.

The in situ observation has a high degree of accuracy, but it has a limited degree of representativeness due to its sparse distribution, which results in difficulties for use in frozen ground research at a regional scale. Remote sensing has the advantage of monitoring regionally distributed land surface states, which play an important role in the operational land data assimilation system[21,23]. Based on the experiments of assimilating in situ observations, the final objective of this paper was to realize the assimilation of SSM/I brightness temperature in ALDAS.

In winter, the low soil hydraulic conductivity in the frozen soil and minimal precipitation resulted in a weak correlation between the soil moisture of each layer. Furthermore, there was no reliable unfrozen water observation to be assimilated, and the soil temperature is the major factor to determine the unfrozen water content. As
such, the soil temperature is the key variable to be improved during the winter. The summer monsoon precipitation has a significant influence on the soil moisture, and the simulation error of soil moisture would also result in obvious deviations from the soil temperature simulation.\textsuperscript{[19, 20]} Thus, the soil moisture is the key variable to be improved during the summer.

Considering both the assimilation efficiency and performance, we set the number of ensembles to 50\textsuperscript{[19, 40]}. The RMSE (Root Mean Square Error) and MBE (Mean Bias Error) were used to evaluate the SHAW model simulation accuracy and assimilation results.

### 2.3 Observation error

The observation error includes instrument error and representative error. It is hypothesized that the SMTMS data have a high degree of observation accuracy. The variance of soil temperature observation error obtained by the Pt-100 temperature sensor is 0.5°C\textsuperscript{2}; the variance of soil moisture observation error obtained by Trime MUX TDR is 0.0004 (m\textsuperscript{3}·m\textsuperscript{-3})\textsuperscript{2} in thawed soil. However, the variance of soil moisture observation error in the frozen soil is difficult to estimate because the unfrozen water content is systemically overestimated by TDR; this error is closely dependent upon the initial liquid water content.\textsuperscript{[41, 42]} Assuming that the observation error obeys a Gaussian distribution, the observation was randomly disturbed according to the observation error, which was also constrained by the physical ranges to produce the observation vector ensemble.

### 2.4 Model simulation error

The model error originating from uncertainties of the model parameterization scheme, model parameters, and atmospheric forcing data is important in the land data assimilation system. The SHAW model errors were analyzed by comparing the model simulation with SMTMS observations. The periods of July 1, 1998 to August 31, 1998 and October 1, 2002 to March 31, 2003 were selected to evaluate the SHAW model error of thawed soil during the summer and frozen soil during the winter, respectively. The error statistics are listed in Tables 1 and 2.

Generally, the SHAW model accurately depicts the soil hydrothermal dynamics. The simulation accuracy during the summer was better than that during the winter due to imperfections of the frozen soil parameterization scheme in the SHAW model. The soil moisture of each layer were underestimated to different degrees (MBE was a negative value), whereas the soil temperatures were overestimated (MBE was a positive value). The same simulation result was also found when running the SHAW model at the alpine meadow of the Heihe River basin during the period from March to August.\textsuperscript{[29]} Likely, the overestimated soil temperature was due to a simulation error of soil moisture, and the underestimated soil moisture resulted from an uncertainty of evaporation, interception, and infiltration. Tables 1 and 2 indicate that both RMSEs and MBEs of the soil temperature and soil moisture simulations in the shallow soil layer were larger than those in the deep soil layer. The soil temperature RMSE is 4−9°C\textsuperscript{2} for the 0−4-cm soil layer, 1−3°C\textsuperscript{2} for the 10−100 cm soil layer, and less than 1°C\textsuperscript{2} for 100−160 cm soil layer. The soil moisture RMSE is between 0.0005−0.0064 (m\textsuperscript{3}·m\textsuperscript{-3})\textsuperscript{2}.

### 3 Results of one-dimensional assimilation experiments

#### 3.1 One-dimensional experiments of assimilating the SMTMS in situ observation

When assimilating the in situ SMTMS data, the 4-cm
observations were used by considering: 1) The 0-cm soil temperature observation is influenced by many factors, i.e., solar radiation, cloud cover, and land surface type, resulting in its large fluctuations; there is no 0-cm soil moisture observation. 2) Compared to the 4-cm soil temperature, the 0-cm soil temperature has lower correlations with soil temperatures of other soil layers. It is difficult to transfer the surface observation information into the deep soil and improve the state variable estimations of deep soil. 3) The 4-cm soil temperature represents the effective emission temperature of soil at microwave frequency, which can provide useful prior information for assimilating the passive microwave brightness temperature.

3.1.1 One-dimensional experiments of assimilating the in situ 4-cm soil temperature observations

The one-dimensional experiments of assimilating the in situ SMTMS observations with time intervals of 1, 3, 6, 12 and 24 hours were carried out to test the performance of ALDAS. The simulation period was from October 1, 2002 to March 31, 2003. The assimilation results matched well with the SMTMS observation because the direct in situ observations had little error. Additionally, the shorter the assimilation time interval, the better the assimilation results. When the time interval reached one hour, the assimilation result of the 4-cm soil temperature was nearly consistent with observation. The 4-cm soil temperature RMSE significantly decreased from 1.86 ℃ by the SHAW simulation to 0.16 ℃ after hourly assimilation. When the assimilation time interval was increased, the choice of time to assimilation became even more important. For example, assimilating the observation at 12:00 is more effective than doing so at 24:00 for the same 24-hour time interval. This is because the overestimated thermal conductivity and underestimated soil moisture in the SHAW model result in an overestimated heating effect of solar radiance on the 4-cm soil temperature. Assimilating the observation at 12:00 adjusts the simulation error correctly (Table 3).

The land data assimilation method can not only improve the surface state variable estimations but also optimize the state variable profiles through the model error covariance matrix combined with the energy and mass transfer process between soil layers. A key consideration is the reasonable model error covariance matrix. We assumed that the correlation between state variable errors of each layer could be represented by the correlation between state variables of each layer, which can be calculated by the SMTMS observation of each soil layer and is the normalized form of the model error covariance matrix (Table 4).

<table>
<thead>
<tr>
<th>Time interval of assimilation (h)</th>
<th>RMSE (℃)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.16</td>
</tr>
<tr>
<td>3</td>
<td>0.83</td>
</tr>
<tr>
<td>6</td>
<td>1.19</td>
</tr>
<tr>
<td>12</td>
<td>1.35</td>
</tr>
<tr>
<td>24 (12 h)</td>
<td>1.44</td>
</tr>
<tr>
<td>24 (24 h)</td>
<td>1.74</td>
</tr>
</tbody>
</table>

Table 4 Correlation coefficients of soil temperature between each soil layer

<table>
<thead>
<tr>
<th>Soil layer (cm)</th>
<th>0</th>
<th>4</th>
<th>20</th>
<th>40</th>
<th>60</th>
<th>80</th>
<th>100</th>
<th>130</th>
<th>160</th>
<th>200</th>
<th>258</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0.440</td>
<td>0.186</td>
<td>0.210</td>
<td>0.21</td>
<td>0.190</td>
<td>0.168</td>
<td>0.144</td>
<td>0.133</td>
<td>0.128</td>
<td>0.114</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0.840</td>
<td>0.781</td>
<td>0.753</td>
<td>0.718</td>
<td>0.678</td>
<td>0.618</td>
<td>0.582</td>
<td>0.568</td>
<td>0.544</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>1</td>
<td>0.982</td>
<td>0.956</td>
<td>0.927</td>
<td>0.892</td>
<td>0.836</td>
<td>0.799</td>
<td>0.785</td>
<td>0.764</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>1</td>
<td>0.991</td>
<td>0.973</td>
<td>0.948</td>
<td>0.902</td>
<td>0.868</td>
<td>0.855</td>
<td>0.839</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>1</td>
<td>0.994</td>
<td>0.980</td>
<td>0.945</td>
<td>0.917</td>
<td>0.904</td>
<td>0.892</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>1</td>
<td>0.995</td>
<td>0.972</td>
<td>0.949</td>
<td>0.936</td>
<td>0.927</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>0.989</td>
<td>0.972</td>
<td>0.961</td>
<td>0.953</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>130</td>
<td>1</td>
<td>0.995</td>
<td>0.986</td>
<td>0.981</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>160</td>
<td>1</td>
<td>0.996</td>
<td>0.991</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td></td>
<td>0.996</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>258</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>
deep soil layer can only be optimized slowly through the hydrothermal transfer process in the SHAW model. The improvement of soil temperature profiles is limited (Figure 2(a)). When the covariance terms in the model error covariance matrix are given according to the correlation analysis, the optimized 4-cm soil temperature can be transferred rapidly to the deep soil layer, achieving an adjustment of soil temperatures below the 4-cm depth and an improvement in the overall soil temperature profile (Figure 2(b)).

Table 5 indicates the soil temperature RMSEs of the SHAW simulation versus assimilation results with different model error matrices. The RMSE with covariance equal to 0 was on average 0.30°C lower than that of the SHAW simulation, while the RMSE with reasonable covariance was on average 0.96°C lower than that of the SHAW simulation. The reasonable non-diagonal elements in the model error covariance matrix can efficiently and significantly improve the estimation of soil temperature profiles. Furthermore, due to the interactions of soil layers, the estimation accuracy of the 4-cm soil temperature with reasonable non-diagonal elements was better than that with non-diagonal elements equal to 0.

![Figure 2](image_url)  
*Figure 2* Soil temperature assimilation results with covariance of 0 (a) and with covariance obtained by correlation analysis (b).
Table 5  
soil temperature RMSEs of the SHAW simulation versus assimilation results after fusing the in situ 4-cm soil temperature observation  
(Unit: ℃)
<table>
<thead>
<tr>
<th>Soil layer (cm)</th>
<th>SHAW simulation</th>
<th>Assimilation (covariance equaling to 0)</th>
<th>Assimilation (reasonable covariance)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10.09</td>
<td>9.84</td>
<td>8.05</td>
</tr>
<tr>
<td>4</td>
<td>3.03</td>
<td>2.39</td>
<td>1.61</td>
</tr>
<tr>
<td>10</td>
<td>2.43</td>
<td>1.86</td>
<td>0.91</td>
</tr>
<tr>
<td>20</td>
<td>2.07</td>
<td>1.52</td>
<td>0.60</td>
</tr>
<tr>
<td>40</td>
<td>1.65</td>
<td>1.17</td>
<td>0.99</td>
</tr>
<tr>
<td>60</td>
<td>1.75</td>
<td>1.34</td>
<td>0.79</td>
</tr>
<tr>
<td>80</td>
<td>1.67</td>
<td>1.34</td>
<td>0.63</td>
</tr>
<tr>
<td>100</td>
<td>1.61</td>
<td>1.36</td>
<td>0.38</td>
</tr>
<tr>
<td>130</td>
<td>1.20</td>
<td>1.08</td>
<td>0.42</td>
</tr>
<tr>
<td>160</td>
<td>0.81</td>
<td>0.79</td>
<td>0.51</td>
</tr>
<tr>
<td>200</td>
<td>0.36</td>
<td>0.35</td>
<td>0.31</td>
</tr>
<tr>
<td>258</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
</tr>
</tbody>
</table>

a) There is no 10-cm soil temperature observation; it is replaced by the average of 4-cm and 20-cm soil temperatures.

3.1.2 One-dimensional experiment of assimilating in situ 4 cm soil moisture observation

The one-dimensional experiment of assimilating the soil moisture observation was carried out during the period from July 1 to August 31, 1998 in order to improve the estimation of soil moisture profile. Similar to the soil temperature assimilation, the correlation coefficients between soil moisture of each soil layer were calculated based on in situ TDR observations (Table 6). The correlation of soil moisture was not like that of soil temperature, which decreased with distance between soil layers. There is a weak correlation between soil moisture of 60-cm and 100-cm soil layers with other soil layers. The model state vector only includes the 0-cm, 4-cm, 10-cm, and 20-cm soil moisture, which have higher correlations between one another.

Figure 3 shows the daily soil moisture profile after assimilating the 4-cm soil moisture observations. The improvements of 4-cm and 20-cm soil moisture were remarkable due to higher correlations. The RMSEs of 4-cm and 20-cm soil moisture decreased by 0.033 m$^3$·m$^{-3}$ and 0.027 m$^3$·m$^{-3}$, respectively. Although the estimation of 60-cm and 100-cm soil moisture were improved overall, their assimilation results were larger than the observation after July 24, 1998. This may be attributed to the weak correlation between 4-cm soil moisture and 60-cm and 100-cm soil moisture. After assimilating the 4-cm soil moisture observation, the soil moisture RMSE of each soil layer decreased by 0.02 m$^3$·m$^{-3}$ on average compared to the SHAW simulation (Table 7). Furthermore, the overestimated soil temperatures were indirectly adjusted by being influenced by the increased soil moisture assimilation results. The soil temperature RMSE decreased 0.08 ℃ on average; however, the improvement was still smaller than that of directly assimilating the soil temperature observations (Table 8).

Table 6 Correlation coefficients of soil moisture between each soil layer
<table>
<thead>
<tr>
<th>Soil layer (cm)</th>
<th>0</th>
<th>4</th>
<th>20</th>
<th>40</th>
<th>60</th>
<th>80</th>
<th>100</th>
<th>130</th>
<th>160</th>
<th>200</th>
<th>258</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0.924</td>
<td></td>
<td>0.403</td>
<td>0.411</td>
<td>0.702</td>
<td>0.854</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>1</td>
<td></td>
<td>0.450</td>
<td>0.441</td>
<td>0.598</td>
<td>0.928</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>1</td>
<td></td>
<td></td>
<td>0.784</td>
<td>0.459</td>
<td>0.395</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>1</td>
<td></td>
<td></td>
<td>0.535</td>
<td>0.410</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>0.562</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>130</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>160</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>258</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Table 7 Soil moisture RMSE of the SHAW simulation versus assimilation result after fusing the in situ 4-cm soil moisture observation
<table>
<thead>
<tr>
<th>Soil layer (cm)</th>
<th>4</th>
<th>20</th>
<th>60</th>
<th>100</th>
<th>160</th>
<th>258</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHAW simulation (m$^3$·m$^{-3}$)</td>
<td>0.082</td>
<td>0.047</td>
<td>0.033</td>
<td>0.026</td>
<td>0.068</td>
<td>0.0003</td>
</tr>
<tr>
<td>Assimilation result (m$^3$·m$^{-3}$)</td>
<td>0.049</td>
<td>0.020</td>
<td>0.019</td>
<td>0.017</td>
<td>0.039</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

Table 8 Soil temperature RMSE of the SHAW simulation versus assimilation result after fusing the in situ 4-cm soil moisture observation
<table>
<thead>
<tr>
<th>Soil layer (cm)</th>
<th>0</th>
<th>4</th>
<th>10</th>
<th>20</th>
<th>40</th>
<th>60</th>
<th>80</th>
<th>100</th>
<th>130</th>
<th>160</th>
<th>200</th>
<th>258</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHAW simulation (℃)</td>
<td>3.826</td>
<td>2.176</td>
<td>1.190</td>
<td>1.101</td>
<td>0.829</td>
<td>1.130</td>
<td>1.033</td>
<td>1.023</td>
<td>0.911</td>
<td>0.757</td>
<td>0.618</td>
<td>0.045</td>
</tr>
<tr>
<td>Assimilation result (℃)</td>
<td>3.842</td>
<td>2.104</td>
<td>1.134</td>
<td>1.013</td>
<td>0.729</td>
<td>1.000</td>
<td>0.897</td>
<td>0.881</td>
<td>0.791</td>
<td>0.651</td>
<td>0.570</td>
<td>0.045</td>
</tr>
</tbody>
</table>
Figure 3  Daily soil moisture profiles after assimilating in situ 4-cm soil moisture observations.

3.2 One-dimensional experiments of assimilating the SSM/I brightness temperature

Although the in situ observation has a reliable accuracy, its sparse distribution, limited regional representativeness and availability result in difficulties in its application to regional land data assimilation systems. However, the spatial resolution of passive microwave remote sensing is tens of kilometers, and microwave is extremely sensitive to surface soil temperature and soil moisture, which make it suitable for the large-scale land data assimilation applications.

The most obvious difference between assimilating the SMTMS in situ observation and remote sensing observation is the observation operator. When assimilating the SMTMS in situ observation, the observation operator is the identity vector. While assimilating the passive microwave remote sensing observation, the observation operator is the forward microwave radiative transfer model used to convert the soil state variables predicted by the SHAW model into the simulated brightness temperature. The LSP/R model coupled with AIEM was used as the observation operator in ALDAS. The surface emissivity was calculated by the surface scattering model AIEM. The standard deviation of surface height and correlation length describing the surface roughness were manually adjusted to 0.215 and 8.0 cm according to the model parameterization.

Compared to thawed soil, the volume scattering effect in the frozen soil is evident with the increased microwave penetration depth. The higher the microwave frequency is, the stronger the volume scattering. This phenomenon has been previously defined as “volume scattering darkening”\cite{43}. However, the LSP/R model does not consider the volume scattering effect. As such, the 19.35 GHz vertical and horizontal polarization brightness temperatures were assimilated as observation variables; 19.35 GHz is the lowest frequency within the four SSM/I frequencies (19.35, 22.235, 37.0 and 85.5 GHz), and it is less influenced by the volume scattering effect.

The observation error of SSM/I brightness temperature primarily originated from radiometer performance and atmospheric effects. The error variance of SSM/I brightness temperature was estimated as 4–10 K\cite{21} and the model error covariance matrix was the same as in section 3.1.

3.2.1 One-dimensional experiment for assimilating the SSM/I brightness temperature during the winter

Figure 4 shows the soil temperature assimilation results of each soil layer. The assimilation results matched well with the SMTMS observation apart from the 160-cm soil temperature. The error statistics indicated that the soil temperature RMSE of each soil layer decreased by 0.76°C on average after assimilating the SSM/I 19 GHz brightness temperature compared to the SHAW simulation (Table 9). The effect of assimilating the microwave brightness temperature was not better than that of assimilating the in situ observation due to the indirect observation of land surface states by the passive microwave radiometer; the error in the microwave radiative transfer model and scale do not match between the SHAW model and the SSM/I resolution. The soil tem-
perature assimilation results of each soil layer were underestimated prior to January 2003 and improved subsequently. This is likely because the soil was stratified when it started to freeze from the top toward the bottom. The top frozen soil layer has an extinction effect on the microwave radiance from the underlying thawed soil layer. However, the LSP/R model does not simulate the microwave radiative transfer process of layered soil. On the other hand, the assimilation effect was influenced by the SHAW simulation, especially during the transition period between soil freezing and thawing. The low and unstable simulation accuracy of the SHAW model during soil state changes disturbed the assimilation result and resulted in the assimilation result even being worse than the SHAW simulation.

### 3.2.2 One-dimensional experiment for assimilating the SSM/I brightness temperature during the summer

The error statistics indicated that the soil moisture RMSE of each soil layer was on average decreased by 0.018 m$^3$·m$^{-3}$ after assimilating the SSM/I 19 GHz brightness temperature, compared to the SHAW simulation. The estimation of soil moisture profiles was improved, especially when a precipitation event occurred (Figure 5). For example, the 6.2 mm of precipitation on August 31, 1998 resulted in an improvement in a soil moisture assimilation result. This is because the microwave penetration depth is shallow due to the absorption

![Figure 4](image)

**Figure 4** Soil temperature profiles after assimilating SSM/I 19 GHz brightness temperatures.

![Figure 5](image)

**Figure 5** Soil moisture profiles after assimilating SSM/I 19 GHz brightness temperatures.

<table>
<thead>
<tr>
<th>Soil layer (cm)</th>
<th>0</th>
<th>4</th>
<th>10</th>
<th>20</th>
<th>40</th>
<th>60</th>
<th>80</th>
<th>100</th>
<th>130</th>
<th>160</th>
<th>200</th>
<th>258</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHAW simulation (℃)</td>
<td>10.09</td>
<td>3.03</td>
<td>2.43</td>
<td>2.07</td>
<td>1.65</td>
<td>1.75</td>
<td>1.67</td>
<td>1.61</td>
<td>1.20</td>
<td>0.81</td>
<td>0.36</td>
<td>0.04</td>
</tr>
<tr>
<td>Assimilation result (℃)</td>
<td>8.60</td>
<td>1.89</td>
<td>1.23</td>
<td>1.15</td>
<td>1.44</td>
<td>1.06</td>
<td>0.73</td>
<td>0.61</td>
<td>0.62</td>
<td>0.64</td>
<td>0.34</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Table 9  Soil temperature RMSEs of the SHAW simulation versus assimilation results after fusing the SSM/I 19 GHz brightness temperatures.
effect of liquid water in the soil, which means that the surface scattering effect dominates the radiated energy. The LSP/R model only considers the surface scattering and can accurately simulate microwave emission of thawed soil, especially moist soil after precipitation. Thus, the simulation accuracy of microwave radiative transfer models has important influences on the assimilation result.

4 Conclusions and discussion

One-dimensional assimilation experiments indicated that ALDAS can improve the estimation of soil temperature and soil moisture profiles. The reasonable model error covariance matrix can play a key role in transferring the optimized state information to the deep soil, thereby achieving an improvement in the overall soil state profile estimation. After assimilating the 4-cm soil temperature observations, the soil temperature RMSE decreased by 0.96 °C on average compared to the SHAW simulation. After assimilating the 4-cm soil moisture observation, the soil moisture RMSE decreased by 0.020 m³·m⁻³ on average compared to the SHAW simulation.

After assimilating the SSM/I 19 GHz brightness temperature, the soil temperature RMSE decreased by 0.018 m³·m⁻³ during the winter, while the soil moisture RMSE decreased by 0.018 m³·m⁻³ during the summer.

The current version of ALDAS is primarily concentrated with system integration and one-dimensional assimilation experiments, which provide prototype for regional active layer research. Each part of the ALDAS still requires improvement. These include:

1) Synchronously assimilating soil temperature and soil moisture. Both the soil temperature and moisture are important state variables in the land surface model and interact with one another. The current version of ALDAS can only assimilate soil temperature or soil moisture separately, and it directly updates the model state variables. It is important to assimilate soil temperature and soil moisture synchronously. The key is to determine the covariance between soil temperature and soil moisture, which is of essential importance because both variables have significantly different orders of magnitude.

2) Developing a dual-pass land data assimilation system. The parameters in the SHAW model have a high degree of uncertainty due to limited understanding and availability of soil texture, soil hydrothermal properties, and land surface properties. For example, soil water flux is sensitive to the pore size distribution index and soil porosity. Furthermore, the saturated hydraulic conductivity has a large range, from 6×10⁻⁶ to 1.2×10⁻⁴ m·s⁻¹. The inaccurate assignment of these model parameters leads to a large error in the simulation result. This has great potential toward developing the dual-pass land data assimilation system to optimize the estimation of state variables and model parameters[44,45].

3) Improving microwave radiative transfer models for frozen soil. The forward microwave radiative transfer model links state variables predicted by land surface models and the brightness temperature observed by microwave radiometers. The simulation performance of microwave radiative transfer model directly influences the assimilation result. Therefore, it is necessary to consider absorption and scattering effects in the frozen soil, the energy redistribution at the boundary of frozen soil, snow cover, thawed soil and air, and the vertical stratification of physical properties in the frozen soil. The advanced microwave radiative transfer model should thus be developed for three cases of frozen soil: (1) the whole soil layer is frozen; (2) the top soil layer is frozen, and the underlying soil layer is thawed; (3) snow covers the frozen soil.

Furthermore, the total volume fraction of the soil matrix and ice in the frozen soil is considerable. If the coherent effect of scattered electromagnetic waves is not considered, it would result in an overestimation of scattering extinction effect. As such, the dense media radiative transfer (DMRT) theory[46] should be introduced. The improved microwave radiative transfer model can be used to assimilate high frequency microwave remote sensing data and therefore extend the observation resource.

4) Estimation of errors in the land data assimilation system. The errors play an important role in the land data assimilation system. If the one-dimensional data assimilation system is extended to the regional application, the horizontal spatial correlation between state variables should be considered[47].

5) Corresponding relationship between the state variables in the SHAW model and the input variables into the observation operator. The soil layers defined in the SHAW model are fixed. However, the microwave sensible depth depends primarily on the microwave frequency and soil moisture. As such, the corresponding relationship between the state variables in the SHAW
model and the input variables into the observation operator should be found according to the microwave frequency, frozen depth, and unfrozen water content.

We thank the reviewers for their helpful suggestions, which have improved the overall manuscript. We are also grateful to the CEDOP project for generously providing SMTMS data, and the NSIDC for providing SSM/I brightness temperatures.

26 England A W, DeRoo R. Active layer thickness and moisture content of arctic tundra from SVAT/Radiobrightness models and assimilated 1.4 or 6.9 GHz brightness. Final Report of NSF Award, ID 0240747, 2006


Evensen G. Sampling strategies and square root analysis schemes for the EnKF. Ocean Dyn, 2004, 54: 539－560[doi]


Song K C. The hydrothermal transfer observation and simulation research of typical vegetation Landscapes zone in Heihe River Basin. Doctoral Dissertation. Beijing: Graduate University of Chinese Academy of Sciences, 2005


England A W. The effect upon microwave emissivity of volume scattering in snow, in ice and in frozen soil. Proc URSI Special Meeting on Microwave Scattering and Emission from the Earth, Berne, Switzerland, 1974. 273－287


