Development and preliminary validation of a land surface image assimilation system based on the common land model

Wangbin Shen[1], Zhaohui Lin[2], Zhengkun Qin[1], Juan Li[3]

5

[1]Center of Data Assimilation for Research and Application, Nanjing University of Information Science and Technology, Nanjing, 210044, China
[2]International Center for Climate and Environment Sciences, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, 100029, China
10 [3]CMA Earth System Modeling And Prediction Center, China Meteorological Administration, Beijing, 100086, China

Correspondence to: Zhaohui Lin (lzh@mail.iap.ac.cn); Zhengkun Qin (001771@nuist.edu.cn)

Abstract. Data assimilation is an essential approach to improve the predictions of land surface models. Due to the characteristics of single-column models, assimilation of land surface information has mostly focused on improving the assimilation of single-point variables. However, land surface variables affect short-term climate more through large-scale

15 anomalous forcing, so it is indispensable to pay attention to the accuracy of the anomalous spatial structure of land surface variables. In this study, a land surface image assimilation system capable of optimizing the spatial structure of the background field is constructed by introducing the curvelet analysis method and taking the similarity of image structure as a weak constraint. The ERA5_Land soil moisture reanalysis data are used as ideal observation for the preliminary effectiveness validation of the image assimilation system. The results show that the new image assimilation system is able to

20 well absorb the spatial structure information of the observed data and has a remarkable ability to adjust the spatial structure of soil moisture in the land model. The spatial correlation coefficient between model surface soil moisture and observation has increased from 0.39 to about 0.67 after assimilation. By assimilating the surface soil moisture data and combining with the model physical processes, the image assimilation system can also gradually improve the spatial structure of deep soil moisture, with the spatial correlation coefficient between model soil moisture and observation increased from 0.35 to about

25 0.57. The forecast results show that the positive assimilation effect could be maintained for more than 30 days. The results of this study adequately demonstrate the application potential of image assimilation system in the short-term climate prediction.

1 Introduction

30   Soil moisture not only affects surface processes such as dust (Lei et al., 2005), but also progressively influences climate change by altering surface albedo, heat capacity, and sensible heat and latent heat transported to the atmosphere (Lin et al., 2001; Li et al., 2019; Zhou et al., 2020a). Soil moisture changes slowly relative to the atmospheric variables, that is, the soil moisture has long-term memory. The initial soil moisture anomaly in the sub-seasonal to seasonal forecasting system can be transferred into the forecast, and thus it is an important source of sub-seasonal climate predictability (Koster et al., 2020).

35   Accurate initial land surface conditions can remarkably improve the accuracy of climate and hydrological projections in short-term climate prediction, especially in fully coupled numerical models (Zhan and Lin, 2011；Wang and Cui, 2018; Zheng et al., 2018; Crow et al., 2020; Reichle et al., 2021; Cui and Wang, 2022).

Based on the comprehensive consideration of observation and model errors, the land surface data assimilation method effectively integrates the model background field and various types of observational data with different spatio-temporal

40   distributions and error characteristics, so as to obtain the optimal initial conditions of soil moisture (Li et al., 2020a; Naz et al., 2020). Research on land surface data assimilation methods has gained the attention of meteorologists around the world. Initially, the European Centre for Medium-Range Weather Forecasts (ECMWF) used the nudging method to adjust land surface variables based on the relationship between the forecast errors of atmospheric variables (e.g. specific humidity at near-surface level) and soil moisture errors (Douville et al., 2000). In order to make the assimilated analysis fields better

45   coordinated with other variables of the model, Mahfouf (1991) proposed the optimal interpolation (OI) scheme to assimilate near-surface temperature and humidity observations. The four-dimensional variational data assimilation (VDA) method has also been applied in the study of land surface data assimilation (Reichle et al., 2001). However, for the complex land surface models with strong nonlinearity, it is difficult to compile adjoint models (Dunne and Entekhabi, 2005). Therefore, the VDA is barely used in land surface assimilation. The Kalman filter-like assimilation with no adjoint models is more widely used

50   for land surface data assimilation (Tian et al., 2008; Jin and Li, 2009; Jia et al., 2010; Shi et al., 2011; Sabater et al., 2019; Tangdamrongsub et al., 2020).

Soil moisture assimilation has been conducted at different spatial scales using a range of methods such as the VDA and Kalman filtering (Gruber et al., 2018; De Santis et al., 2021; Sabater et al., 2019). Stable assimilation also remarkably improves soil moisture prediction (Khaki et al., 2020). However, in order to accommodate the features of the single-column

55   land surface model, the current land surface assimilation system is a single-column assimilation system, which neglects the spatial continuity of soil variables. On the timescale of short-term climate change, soil moisture is commonly responsible for the abnormal changes in short-term climate through long-term and large-scale anomalous forcing (Lin et al., 2008; Zhong et al., 2020; Dirmeyer et al., 2021). Therefore, improving the accuracy of the anomalous spatial structure of land surface variables, which serve as the lower boundary conditions of numerical models, will help to better predict short-term climate

60   change caused by soil moisture anomalies. Ideally, a single-column assimilation system would also be able to reproduce the correct spatial-structure features of soil moisture anomalies, if the assimilation can obtain the closest result to the true value

at each column. However, in reality, the observation quality varies sharply across regions, and the strong spatial heterogeneity of soil variables also tends to cause large spatial variations in the accuracy of surface variables simulated by the land surface model (Li, 2013; Li et al., 2020b). This leads to the regional differences in the accuracy of the estimations of

65   observation error and background error in the single-column assimilation, and ultimately causes discontinuities in the spatial structure of the anomalies in the analyzed soil moisture fields. Zhou et al. (2020b) also pointed out that most current soil moisture assimilation methods eliminate the systematic biases between observation and simulation by applying the pixel-by-pixel scale transformation. This treatment discards the crucial spatial information contained in the observation, and affects the application of soil moisture in numerical weather prediction, flood forecasting and drought monitoring. Therefore, while

70   data assimilation improves the accuracy of single-point soil variables, appropriately adjusting the spatial structure of soil analysis variables is a critical development direction for land surface assimilation systems.

With the continuous development of meteorological observation techniques, more and more meteorological information can be displayed in the form of images with fine spatio-temporal resolutions, and their continuous dynamic changes generally allow us to better understand the observed variables. However, the huge amount of satellite observation images for

75   the earth system are not sufficiently utilized in the current data assimilation system (Vidard et al., 2008). Stroud et al. (2009) developed several assimilation schemes that combined the images obtained from the Sea-viewing Wide Field Sensor with a two-dimensional sediment transport model of Lake Michigan, which considerably improved the predictions of sediment concentrations in southern Lake Michigan. In order to quantitatively assimilate the structural information contained in images or image sequences into the numerical model, Le Dimet et al. (2015) extracted the key structural observation

80   information in the images by using curvelet transformation as the observation operator, and improved the prediction results of the shallow water equation model through the VDA approach. Titaud et al. (2010) also found that the direct VDA of image sequences is able to reconstruct initial vortices with highly correct positions, sizes and profiles by using the curvelet transform as the observation operator. Currently, direct assimilation of image sequences is primarily used to predict the evolution of geophysical fluids. If the structural information in the observed images can be introduced into the land surface

85   data assimilation system as observations, the accuracy of the spatio-temporal distribution structure of soil moisture in the model can be targetedly improved. The purpose of this study is to construct a land surface image assimilation system based on the theoretical frame of VDA, so as to realize the direct adjustment of the spatial structure of land surface variables and improve the accuracy of the initial soil moisture values, by integrating the observation information of image sequences and the priori knowledge from numerical models. In this study, an attempt is made to test the effectiveness of the image

90   assimilation module in improving the spatial structure of soil moisture at the land surface in a VDA framework, and the related research methodology can be implemented in the alternative assimilation framework as well.

The paper is organized as follows. Section 2 mainly introduces how to select appropriate image observation operators and establish an image assimilation system under the VDA framework. Then, the error characteristics of the image observation operators are systematically analyzed. The land surface model and the observational data used in the assimilation system are

95   also briefly described in section 2. Section 3 presents the experimental designs, and analyzes the error characteristics of

background field and observation data in detail. The results of the idealized experiments are shown in section 4 to verify the effectiveness of the image assimilation system in improving the predictions of the land surface model. Section 5 gives the summary and discussion.

100 **2 Construction of image assimilation system**

2.1 Land surface model

The common land model (CoLM) developed by Dai et al. (2003) was selected in this study. By considering biophysical, biochemical, ecological and hydrological processes, this model well describes the transfer processes of energy, water and

105 carbon dioxide among soil, vegetation, snow cover and atmosphere, allowing the simulation of soil temperature, soil moisture, runoff, heat flux, and other variables. In recent years, the CoLM has incorporated additional physical processes such as glaciers, lakes, wetlands and dynamic vegetation. It has also been successfully implemented in several global atmospheric models (Yuan and Liang, 2011; Ji et al., 2014; Zhang et al., 2020; Yuan and Wei, 2022).

The surface spatial heterogeneity of the CoLM is manifested as a nested sub-grid hierarchy, with the grid units consisting

110 of multiple land units and plant function types (PFTs). The bio-geophysical processes of the CoLM are simulated on a single soil-vegetation-snow column, and each sub-grid has its own predictor variables. Grid-averaged atmospheric forcing is used to force all sub-grid cells within a grid cell. The model used in this study has a horizontal resolution of about 1.4° × 1.4°. There are ten unevenly spaced soil layers and five snow layers in the vertical direction. In this study, soil moisture is calculated by using the equation of one-dimensional soil water vertical motion:

115

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q}{\partial z} - E - R_{fm} \, , \tag{1}$$

where θ is the volumetric water content of the soil (unit: m³·m⁻³), q the soil water flux calculated by the Darcy theorem, E the rate of evaporation (unit: mm·s⁻¹), R_{fm} the rate of thawing or freezing, and z the vertical distance from the soil layer to

120 the ground (q and z is downward-positive).

Atmospheric forcing conditions provide constraints on land-surface models. The atmospheric forcing dataset used to drive the CoLM in this study includes the downward short-wave solar radiation at surface, downward long-wave radiation, near-surface air temperature, specific humidity, precipitation rate, surface atmospheric pressure, U-component wind speed, and V-component wind speed. It has a temporal resolution of three hours (at 0000 UTC, 0300 UTC, 0600 UTC, etc.) and the spatial

125 resolution is T62 (about 1.875°).

In this study, the CoLM is run in the offline mode cyclicly driven by the observation-based forcing data from 1948 to 2020 for 360 years. The water content of the deepest layer changes extremely slowly over the last 50 years, and the model can be considered to be in equilibrium.

2.2 Framework of image assimilation system based on variational data assimilation

130 This study is based on the framework of three-dimensional VDA (3D-VDA). The main principle of 3D-VDA is to simplify data assimilation to a quadratic functional minimization problem, which characterizes the deviations between analysis and observational fields as well as between analysis and background fields.

Assuming that x denotes the vector of analysis variables, x_b denotes the background field, and x_a denotes the analysis field, then the variation of x with time can be expressed as:

135

$$\begin{cases} x(t) = M\big(x(t_0)\big) + \eta(t) \\ \qquad x(t_0) = U \end{cases} , \tag{2}$$

where M denotes the numerical prediction model, and t and t_0 represents the prediction time and start time of the model, respectively. $\eta(t)$ is the model error at moment t, $U \in \mathbb{R}$ denotes the initial conditions of the model, and $\mathbb{R}$ represents the

140 space in which the state variables are located.

If a direct or indirect observation vector of length L is represented by $y^o \in \mathbb{Q}$ and the observation space is represented by $\mathbb{Q}$, then the relationship between observations and state variables can be expressed as follows :

$$y^o(t) = H\big(x(t)\big) + \mathcal{E} , \tag{3}$$

145

where $H: \mathbb{R} \to \mathbb{Q}$ is the observation operator that represents a mapping from the model space to the observation space. The observation operator is simplified to a simple interpolation operator when y^o and x are the same type of physical variable. If the two have different physical properties, the observation operator is a mapping operator with some complex structure that transforms the model space into the observation space. $\mathcal{E}$ represents observation error. The goal of variational assimilation is

150 to determine the model state at time t_0 , so that the sum of the deviation of state variable from background field and the deviation of simulated observation based on model variable from actual observation, is minimized under the premise of additional constraints, that is, to find an analysis field $x_a(t_0)$ which minimizes the following quadratic objective function J.

$$J(x) = \frac{1}{2}[x - x_b]^T B^{-1}[x - x_b] + \frac{1}{2}[H(x) - y^o]^T (O + F)^{-1}[H(x) - y^o] , \tag{4}$$

155

Where B , O and F are respectively the error covariance matrixes of background field, observation data and observation operator, which are known as prior information. B^{-1} are the inverse of the background error covariance matrix with order $N \times N$, and N is the freedom degree of the analysis field. $(O + F)^{-1}$ is the inverse of the observation error covariance matrix with order $L \times L$.

160 During minimizing the above objective function, the optimal analytical variable is $x = x_a$ when $\nabla_x J(x_a) = 0$. It represents the optimal estimate of the true atmospheric state under given background fields, observations, and their error information.

Images are generally characterized by the features of observation variables, such as geometry and distribution. From a "mathematical" point of view, images are usually considered to be real-valued functions of consecutive real variables, so

165 they can be processed by using mathematical tools. In this case, the "numerical image" is a discrete version of the final processed mathematical image (Le Dimet et al., 2014).

The so-called image assimilation refers to the introduction of a weak constraint on the similarity between the structure of the observed and simulated images in the VDA, so that the image observations are used together with the conventional observations to compute the optimal analysis variables. Thus, the cost function for the 3D-VDA can be written in the

170 following form:

$$J(x) = J_B + J_O + J_I = \underbrace{\frac{1}{2}(x_o - x_b)^T B^{-1}(x_o - x_b)}_{conventional\ cost\ of\ J_B} + \underbrace{\frac{1}{2}(H(x_o) - y^o)^T (O + F)^{-1}(H(x_o) - y_o)}_{conventional\ cost\ of\ J_O} +$$

$$\underbrace{\frac{1}{2}\left(H_{F \to s}(f_o) - H_{\mathbb{R} \to s}(x_b)\right)^T \left(H_{F \to s}(f_o) - H_{\mathbb{R} \to s}(x_b)\right)}_{Image\ cost\ J_I},$$

(5)

175

where J_B and J_O denote the background and observation terms in the conventional cost function, and J_I represents the added image observation term. In the image observation term $f(t) \in F$, f_0 represents the frame-wise observation image of the image dynamic observation system at moment t_0, which belongs to the image observation space F. s represents the image space under the mathematical definition. The image structure operator $H_{F \to s}$ represents the mapping from the image space to

180 the structure-defined mathematical space, that is, the structural information extracted from the image observations, which represents the multi-scale geometric features of the image. The operator $H_{\mathbb{R} \to s}$ implies a mapping from the space of state variables to the mathematical space where the structure resides, which represents obtaining the same type of structural information from the background field output by the model.

2.3 Curvelet multiscale analysis method

185   From the above equations, it is clear that the image structure operator is a key technique for the image assimilation system. The important information in an image is mainly located in the discontinuities of the image, which can be well described by the spectral space, and thus the image can be quantitatively described by the spectral transformation coefficients of the image. The curvelet transform is exactly a suitable multi-scale transform analysis method, which is not only capable of time-frequency analysis, but also has strong directional selection and discrimination capabilities (Titaud et al., 2010).

190   The curvelet transform takes the inner product of the basis function and the signal to achieve a sparse representation of the signal. In the two-dimensional space R^2, the spatial variable is denoted by x and the parent wave is denoted by $\varphi_j(x)$. Thus, the curve wave coefficient is the inner product of the function $f \in L^2(R^2)$ and the curve wave $\varphi_{j,l,k}$:

$$c(j,l,k) := \langle f, \varphi_{j,l,k} \rangle = \int_{R^2} f(x)\overline{\varphi_{j,l,k}(x)}d_x, \tag{6}$$

195

where j, l and k are the scale, direction and position parameters, respectively, and the biggest difference with other spectral analysis methods is that the curvelet transform takes into account the direction parameter. The curvelet coefficients are anisotropic, which can efficiently represent the image edges and fully exploit the image features. The reconstruction equation of the coefficients is:

200

$$f = \sum_{j,l,k}\langle f, \varphi_{j,l,k} \rangle \varphi_{j,l,k} = \sum_{j,l,k} c(j,l,k)\varphi_{j,l,k} . \tag{7}$$

An important fact in the practical application of data assimilation methods is the presence of observation errors, which in this case is represented by the noise in the observed images. In order to effectively remove the image noise and extract the main structural features from the image, we can choose the curvelet coefficients at different scales. A simple "hard threshold" approach can achieve this goal, by setting the curvelet coefficients below the threshold (represented by σ) to zero. Denoising and key-feature selection can be achieved by adjusting the threshold value.

Figure 1 gives the structural information of the soil moisture image extracted by the curvelet analysis method under different threshold conditions on May 1, 2016 in East Asia. From the raw soil moisture image (Fig. 1a), it can be seen that soil moisture is low in northwest China and high in the south and east. When the threshold is 0.1 (Fig. 1b), the reconstructed image reproduces the low-value areas of soil moisture in Northwest China and the high-value areas in eastern and southern China, but only represents the large-scale spatial structure features of the raw image. When the threshold value is increased to 0.5 (Fig. 1c), the reconstructed image is definitely close to the original image, and the critical features of the reconstructed image, such as the dry zone in Xinjiang-Inner Mongolia and the wet area in southeast China and Siberia, are basically consistent with those of the original image. Only some small-scale noise information, such as two dry zones located in northern Tibet and a wet zone near 59°N in northeastern Lake Baikal, has been filtered out. It is shown that the multi-scale

structural information of the image could be efficiently extracted by the curvelet analysis method, which provides a basis for introducing the spatial structure information of the observed data into the assimilation.
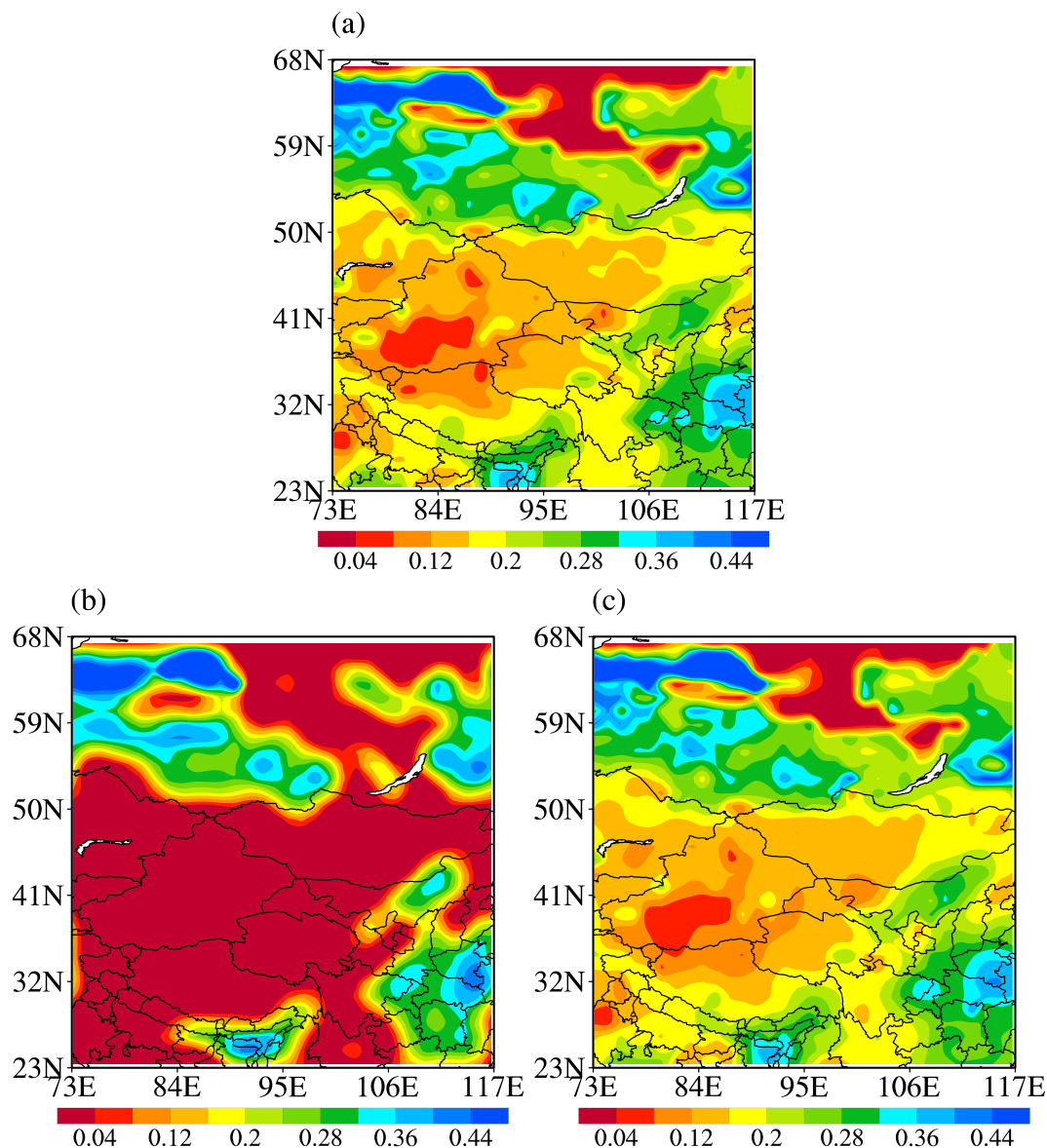
220

Figure 1: Soil moisture distributions from (a) the raw image and extracted by curvelet analysis under the thresholds of (b) 0.1 and (c) 0.5 on May 1, 2016 in western East Asia.

3 Experimental design and error analysis

225 ### 3.1 Ideal observation data

Taking the curvelet transform method as the image observation operator, we construct the land surface image assimilation system based on Eq. (5). To demonstrate the assimilation effect of the image assimilation system, the idealized data is used to examine the ability of the image assimilation method in adjusting the spatial structure of the land surface model variables. The soil moisture reanalysis data of the fifth generation ECMWF ReAnalysis Land (ERA5_Land) is chosen as the ideal

230 observation data in this study, which has a horizontal resolution of 31 km and a temporal resolution of one hour. The H-TESSEL (Hydrology Tiled ECMWF Scheme for Surface Exchanges over Land) land surface model is used in ERA5_Land to output the soil volumetric water content data at four depth layers (0–7, 7–28, 28–100 and 100–289 cm). ERA5-Land is more accurate for all land use types with a series of improvements and recalculations based on ERA5. The calculations are made without coupling the atmospheric and wave module of ECM-WF/IFS, allowing for a faster update frequency. The

235 horizontal and temporal resolutions are respectively increased to 9 km and one hour, and the stratification of output soil volumetric water content is consistent with that of ERA5 (Sabater et al., 2021).

3.2 Ideal experimental design

As shown in Fig. 1a, the selected region for the experiment (73°W-117°W, 23°N-68°N) covers most of the land area of China, and the model spatial resolution is 1.4° × 1.4°. The land-atmosphere coupling is the strongest in the western Qinghai-

240 Tibet Plateau, where soil moisture has a large impact on the climate change and is an essential precursor signal for the summer precipitation forecasts in eastern China (Yuan et al., 2021). The western arid zone has complex topography, with strong spatial heterogeneity in soil moisture. In this region, the surface energy and water vapor budgets also have a crucial impact on the climate (Yang et al., 2021).

The assimilation is run from May 1 to August 31, 2016, and the prediction is made from September 1 to September 30,

245 2016. Two sets of experiments are designed. The first sets of experiments perform data assimilation (DA) four times a day with an interval of 6 hours (at 0000 UTC, 0600 UTC, 1200 UTC and 1800 UTC), and the soil moisture in the surface layer of 0–7 cm from ERA5-Land is assimilated. The other group is the control (CTRL) experiment, which has no observations assimilated.

Since it takes a period of time for the model to integrate to adapt to the soil moisture after assimilation, the first 15-day

250 results of the experiment are discarded to ensure that the model can reach a new hydrological equilibrium state, which can make the evaluation of assimilation effect more objective. The analysis in this study mainly focuses on the period from May 16 to September 30, 2016. To highlight the effect of image assimilation, the other observations are not assimilated in this study, that is, J_O is zero.

3.3 Analysis of error characteristics

255

From the cost function shown in Eq. (4), it can be seen that the solution of the cost function also requires the estimation of the background field error covariance and the observation error in advance, and the elimination of observation noise in the image. To obtain more precise analysis results, we need to accurately estimate the characteristics of various error covariances.

260

3.3.1 The covariance matrix of background field error

According to the characteristic of single-column model that there is no correlation between the simulation errors at different grid points, the covariance matrix of background field error can be directly expressed as the variance of the simulation error
265 of the land-surface model at each grid point.

The depths of the top three soil layers in the CoLM model are 0.70 cm, 2.79 cm and 6.22 cm, which are close to the first-layer depth of the ideal observation data (0-7cm). In this study, the hourly soil moisture data at the top three layers output by the land surface model from 2014 to 2015 is used as the background field. The soil moisture reanalysis data of the first ERA5_Land layer (0-7 cm) at the same time is interpolated to the depth corresponding to the background field, and then is
270 used as the ideal observation data. Based on the difference between the background field and observations, the covariance matrices of background error are seprately obtained for the soil moisture at each of the top three layers. The error covariance between different levels is not considered here.

Figure 2 shows the distributions of background error covariance for the top three soil moisture layers. It can be seen that the soil moisture in the top layer (Fig. 2a) is affected by myriad factors, so its background error is larger than that at the other
275 two layers. The spatial characteristics of the background errors at different depths are similar, with relatively small errors in Xinjiang, northern Tibet and Inner Mongolia, which may be related to the drought in this region. In contrast, in humid areas such as Siberia, the soil moisture is affected by additional factors, which causes relatively larger simulation errors.
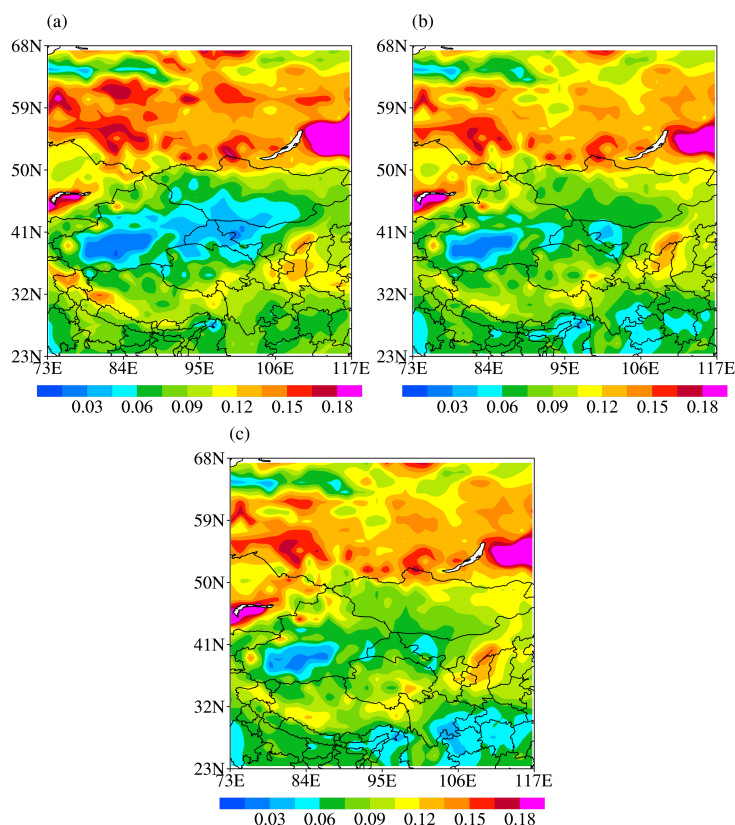
Figure 2: The spatial distributions of soil moisture simulation error variance in the (a) first, (b) second and (c) third layers of the common land model (CoLM) during the statistical period from May 16 to September 30, 2016.

3.3.2 Analysis of observation data error

In image assimilation, the observation noise can be efficiently eliminated by selecting an appropriate threshold value. To further objectively determine the threshold, the soil moisture of the ERA5 data at 100 instants is selected as the original image. Different threshold values are chosen for de-noising, and the reconstructed images with different degrees of de-noising are then obtained by inverse curvelet transformation. The threshold selection method is additionally discussed based on the statistical characteristics of the difference between reconstructed and original images.

Figure 3 shows the spatial distributions of the mean value of 100 reconstructed fields and the mean value of reconstructed errors, based on the raw soil moisture images every 6 hours (0000 UTC, 0600 UTC, 1200 UTC and 1800 UTC) from May 1 to May 25, 2016, with thresholds being 0.1 and 0.5 separately. As can be seen from the original image (Fig. 3a), in terms of large-scale structural features, soil moisture is relatively low in the central part of the selected region (from Xinjiang and northern Tibet to Mongolia), while it is relatively high in the surrounding of low-value areas. High-value centers of soil moisture are found in the southern Siberian Plain, east of Lake Baikal, eastern China, and south of the Tibetan Plateau. When the threshold is 0.1, the average distribution of the reconstructed field (Fig. 3b) reproduces the large-scale characteristics of

the original field that the low soil moisture is located in the middle and surrounded by high value centers. However, there are

295   large errors between the reconstructed field and the original field (Fig. 3c). In particular, the spatial distribution of errors is similar to that of the large-scale original field, which indicates the loss of spatial structure information of the observations. When the threshold increases to 0.5, the spatial correlation coefficient (SCC) between reconstructed field (Fig. 3d) and original field is greater than 0.99, and the multi-scale features of the original field are properly reflected. As can be seen from Fig. 3e, the errors between the reconstructed and original fields are basically within 0.02, and the error distribution presents

300   no obvious spatial structure characteristics.
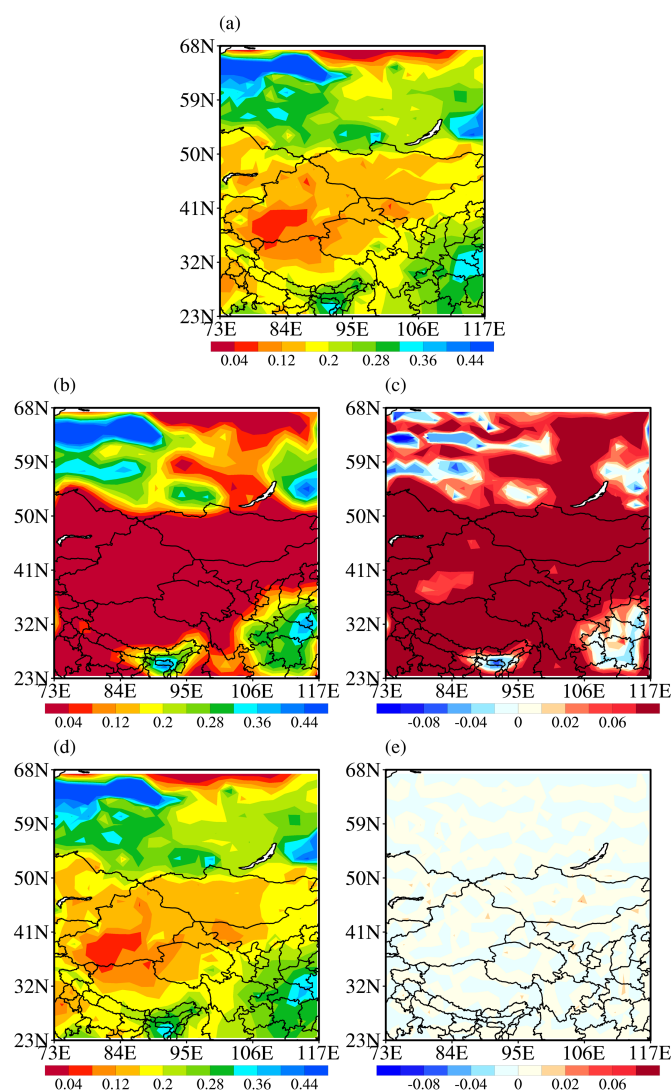
Figure 3: The spatial distributions of (a) original soil moisture, (b, d) reconstructed soil moisture and (c, e) reconstruction errors under the threshold values of 0.1 (b, c) and 0.5 (d, e) averaged from May 1 to May 25, 2016.

To better clarify the statistical characteristics of the reconstruction errors under different thresholds, Figure 4 shows the

305   probability density distribution curves of the reconstruction errors for 100 reconstructed fields at different thresholds. For the error at the threshold of 0.5, the skewness coefficient of the probability density distribution curve is 0.00 and the kurtosis coefficient is 0.38, indicating the curve is close to the standard normal distribution curve (the skewness and kurtosis coefficients are all 0). With the gradual increase of the threshold value, although the reconstruction error decreases, the residual error is mainly concentrated in the range of smaller values, and the curve shows a "sharp peak" distribution.

310   Considering that the observation errors are mostly random errors, it is reasonable to believe that the reconstruction errors at the threshold of 0.5 are mainly observation errors, which also implies that this threshold is good for the purpose of denoising the observation images.
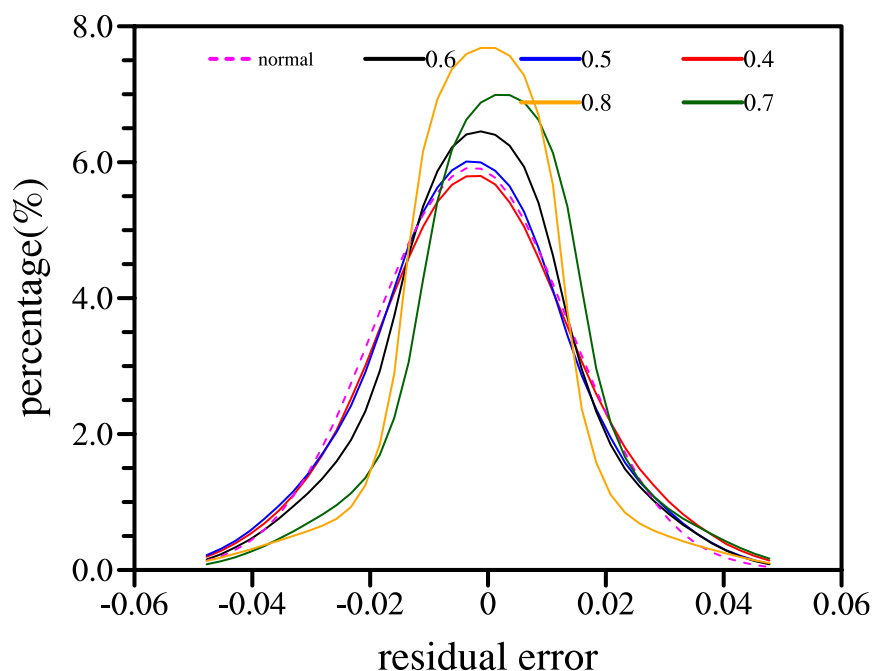
Figure 4: Probability density distributions of 100 reconstructed errors under different thresholds. The magenta dashed line
315   **represents standard normal distribution; the red, blue, black, green and orange solid lines represent threshold values of 0.4, 0.5, 0.6, 0.7 and 0.8, respectively.**

4 Results

Figure 5 shows the variation of the cost function values with the number of iterations when assimilating ERA5-Land surface
320   soil moisture using the image assimilation system at 0000 UTC on May 16, 2016. The criterion of convergence is that the gradient of the cost function values is less than 10^{-9}. It can be seen that the initial value of the cost function is 1121.0, which

has been reduced to 863.1 by the second iteration. The convergence speed of the cost function is relatively fast, and it only needs 22 iterations, which also demonstrates the validity and rationality of introducing the image operator term in the cost function. The fast convergence speed of the cost function caused by the constraint of the image operator also indicates that the assimilation can effectively absorb the spatial structure information of the observation.
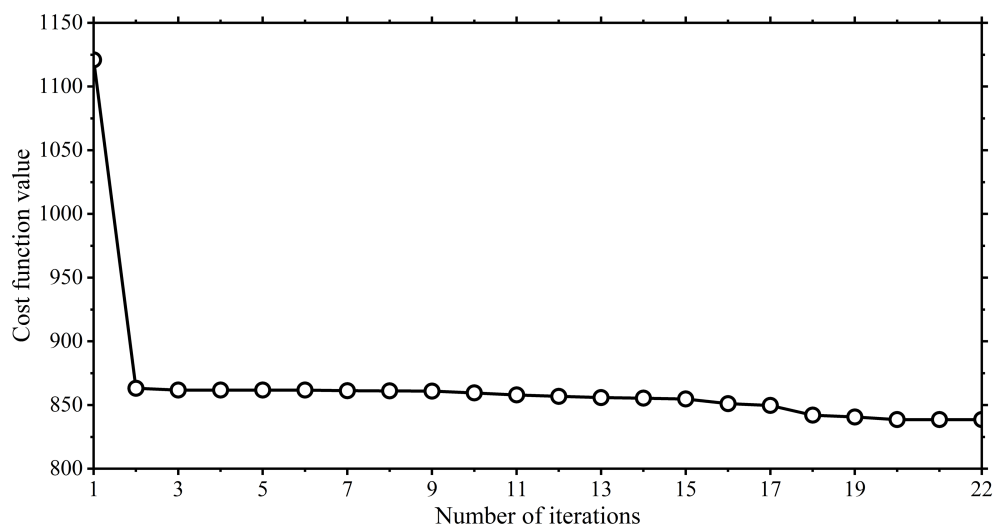

Figure 5: Variation of the cost function value with the number of iterations using the image assimilation system at 0000 UTC on May 16, 2016.

Figure 6 shows the 0-7 cm soil moisture distributions of the observation, the analysis field from the image assimilation system and the output from CoLM at 0000 UTC on May 16, 2016. It can be seen that the surface soil moisture of ideal observations (Fig. 6a) is drier along Mongolia and Xinjiang, but wetter in southern China. The observed soil moisture is also relatively high in the vicinity of the Tianshan Mountains, the eastern part of Qinghai-Tibet Plateau and the southern part of Lake Baikal, as well as in the eastern parts of Henan Province and Inner Mongolia. However, the results of the CTRL experiment (Fig. 6c) show a "high-low-high" distribution from south to north, which is extremely different from the spatial structure of the observation. It can be seen from the spatial distribution of the analysis field that the soil moisture structures are all remarkably improved at different scales (Fig. 6b). In view of large-scale structure, the structural wet bias from northwest China to Mongolia can be thoroughly corrected, and the observed structural features of low soil moisture from southern Qinghai to southeastern China are nicely reproduced in the analysis field. In addition, some relatively small-scale structures, such as the relatively wet soil in the Tianshan Mountains region of Xinjiang, the central part of Qinghai Province, and the northeastern part of Lake Baikal, are also well represented in the analysis field.
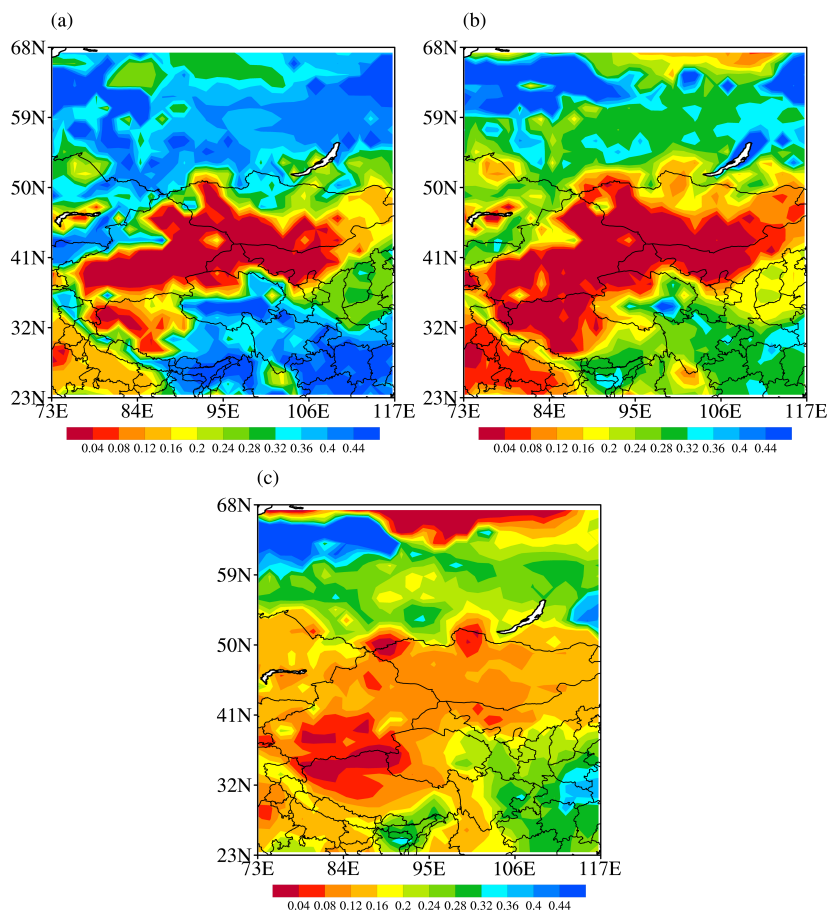
Figure 6: Spatial distributions of soil moisture at the depth of 0-7 cm for (a) the observation, (b) the analysis field from the image assimilation system and (c) the background field from the CoLM at 0000 UTC on May 16, 2016.

345

Figure 7 shows the spatial distributions of soil moisture at 0-7 cm at 0000 UTC on September 1, 2016 from the observation, DA experiment and CTRL experiment. It can be seen from Fig. 7 that, after four consecutive months of cyclic assimilation, the large-scale structure of surface soil moisture in the analysis field is much closer to the observation than that in the CTRL experiment. The improvement is mainly concentrated in the dry zones along Xinjiang-Mongolia, as well as the

350 wet centers in the Tianshan Mountains and central-northern Qinghai Province. The dry tongue from Ningxia Province to Shanxi Province in China has also been reproduced in the analysis field. In addition, the distribution structure of wet areas in the region north of 60°N and east of 111°E has also been thoroughly improved in the analysis field. Overall, the structural characteristics of soil moisture at different scales in the analysis field are in better agreement with the observations.
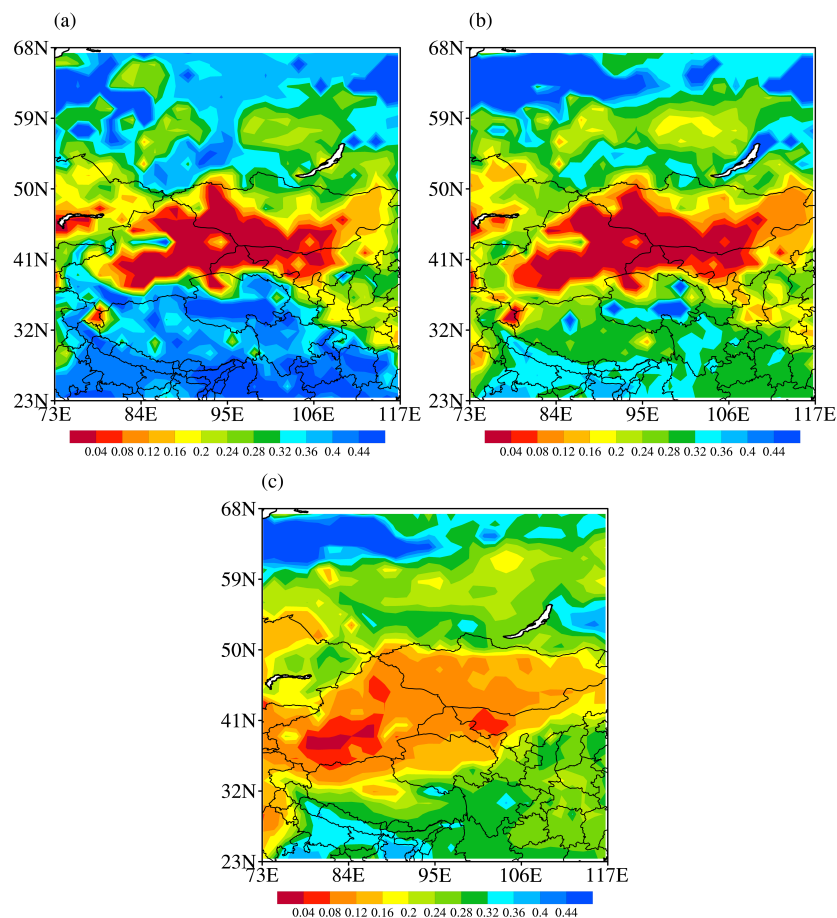
Figure 7: The soil moisture at 0-7 cm from (a) the observation, (b) the data assimilation (DA) experiment after 4 months of continuous assimilation and (c) the control (CTRL) experiment at 0000 UTC on September1, 2016.

Moisture condition of the underlying surface may have some influence on short-term climate anomalies, but whether the effect is significant mainly depends on the duration of the underlying surface features. The surface soil moisture is considerably affected by external high-frequency perturbations, so the retained anomalous signals are susceptible to interference, making the anomalous signals difficult to maintain. However, the deep-layer soil moisture has an excellent and persistent ability to maintain the abnormally strong signals,, which may have a certain impact on the later climate anomalies (Xu et al., 2021). Therefore, it is necessary to conduct further analysis on soil moisture improvement at a deeper level through image assimilation system.

By assimilating the surface soil moisture through the image assimilation system, the deep-layer soil moisture is simultaneously adjusted under the soil hydrodynamic and thermodynamic constraints of the land surface process model. Figure 8 shows the spatial distributions of the ideal soil moisture observation and the soil moisture predictions from the DA and the CTRL experiments at the depth of 7-28 cm at 0000 UTC on September 1, 2016 after the final assimilation.

16

It can be seen that the distribution pattern of deep-layer soil moisture observation (Fig. 8a) is relatively consistent with that

370   of the surface-layer soil moisture. This essentially shows that the region from Xinjiang to Mongolia is an arid region, while the relatively high soil moisture regions are located in southern-southeastern China, and the Siberian Plain. The high-value centers for the surface and deep soil are practically the same, but overall, the deep soil is wetter than the surface. The analysis field of image assimilation (Fig. 8b) shows drought in the southern part of Mongolia, which is consistent with the observation. At the same time, the high-value centers near Novosibirsk, the Tianshan Mountains and the central part of

375   Qinghai Province are also well reproduced. It indicates that the image assimilation also has excellent results in improving the spatial structure of soil moisture at deeper layers.
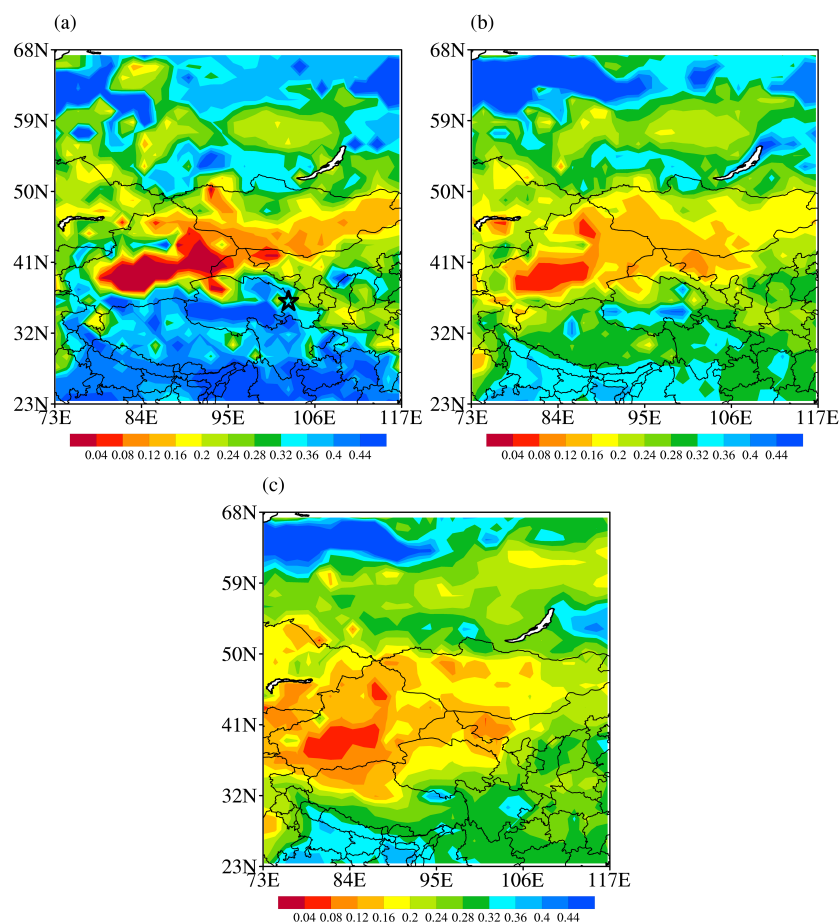
Figure 8: Same as Fig. 7, but for the soil moisture at the depth of 7-28 cm. The pentagram shows the station location for the single-site analysis in Fig. 10.

380

To quantitatively assess the effect of image assimilation in improving the spatial structure of model soil moisture, the SCCs between the ideal observation and model outputs with and without assimilation are calculated and shown in Fig. 9. It can be seen that the SCC of the DA experiment is already much higher than that of the CTRL experiment after the first

17

assimilation, with the SCC increased from 0.44 to 0.76. This proves that the image assimilation can quickly and effectively
385   adjust the spatial structure of soil moisture. During the period of cyclic assimilation period, the SCC can be maintained
above 0.6 with an average value of 0.67, which is steadily higher than that of the CTRL experiment. This indicates the image
assimilation effectively improves the spatial structure of soil moisture, making it more consistent with the observation. The
SCC of the DA experiment is also higher than that of the CTRL experiment in the one-month prediction stage after
September 1, indicating that the optimization of the soil moisture spatial structure by image assimilation could have a
390   obvious positive impact on the prediction in the following month.

Soil moisture is relatively more stable at subsurface depths. The SCC of subsurface soil moisture between observation and
CTRL experiment is 0.31 at the initial time, which increases to 0.53 by introducing image assimilation. After cyclic
assimilation, the mean value of the SCC between observation and DA experiment increases from 0.41 to 0.57, which is
higher than that between observation and the CTRL experiment throughout the entire assimilation period. This indicates that
395   by optimizing the spatial structure of the soil moisture in the surface layer, the soil moisture in the deeper layers is also
favorably improved. In the prediction stage, the SCCs between the DA experiment and the observation are always higher
than those between observation and the CTRL experiment. The mean value reaches 0.63, suggesting that optimization of
surface soil moisture could lead to an excellent improvement in the prediction of deep-layer soil moisture.
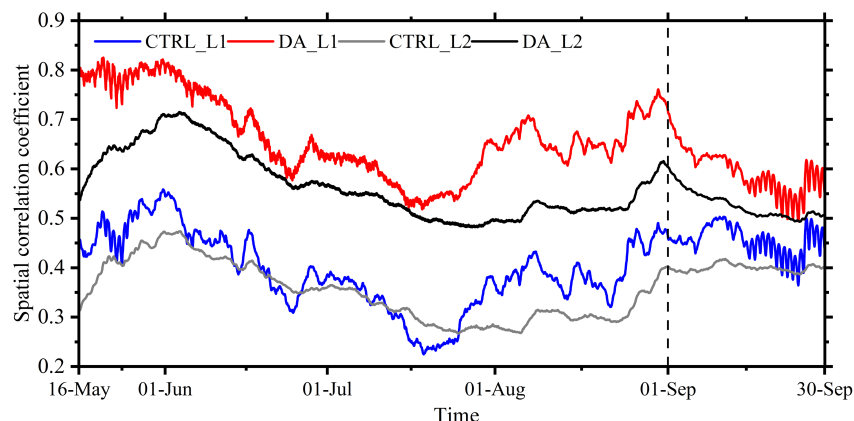

400   **Figure 9: Hourly variations of the spatial correlation coefficient of the surface (red and blue solid lines) and subsurface (black and
gray solid lines) soil moisture between the observations and the experiments with (black and red solid lines) and without (black
and gray solid lines) image assimilation. After the vertical dashed line, it is the prediction period.**

In order to further show the variation characteristics of soil moisture during assimilation, a single-point analysis is also
405   performed by using the hourly soil moisture data from observation, DA and CTRL experiments at a single station in the
Tianshan Mountains region of Xinjiang Province. From the hourly variation of 0-7 cm soil moisture (Fig. 10a), it can be seen
that the observed soil moisture fluctuates around 0.40 while the soil moisture of the CTRL experiment has an obvious
deviation from the observation, and they have different variation trends. However, the soil moisture slowly adjusts during the
image assimilation period and gradually approaches the observation from mid-May to mid-June. By late June, the surface

18

410 soil moisture gradually increases to more than 0.33 m³·m⁻³, which is closer to the observation. At the prediction period in

September, the soil moisture in the DA experiment is also closer to the observation than that in the CTRL experiment.

From the hourly variations of soil moisture at 7-28 cm (Fig. 10b), it can be seen that the observed soil moisture in the deep

layer is more stable than that in the surface layer, and the variation range is smaller, but the trend in the deep layer is

approximately the same as that in the surface layer. In late May, when the surface soil moisture becomes wetter, the deep-

415 layer soil moisture in the DA experiment gradually responds, and its value gradually increases and approaches the observed

value. Both in the assimilation and the prediction periods, the soil moisture in the DA experiment is closer to the observation

than in the CTRL experiment.
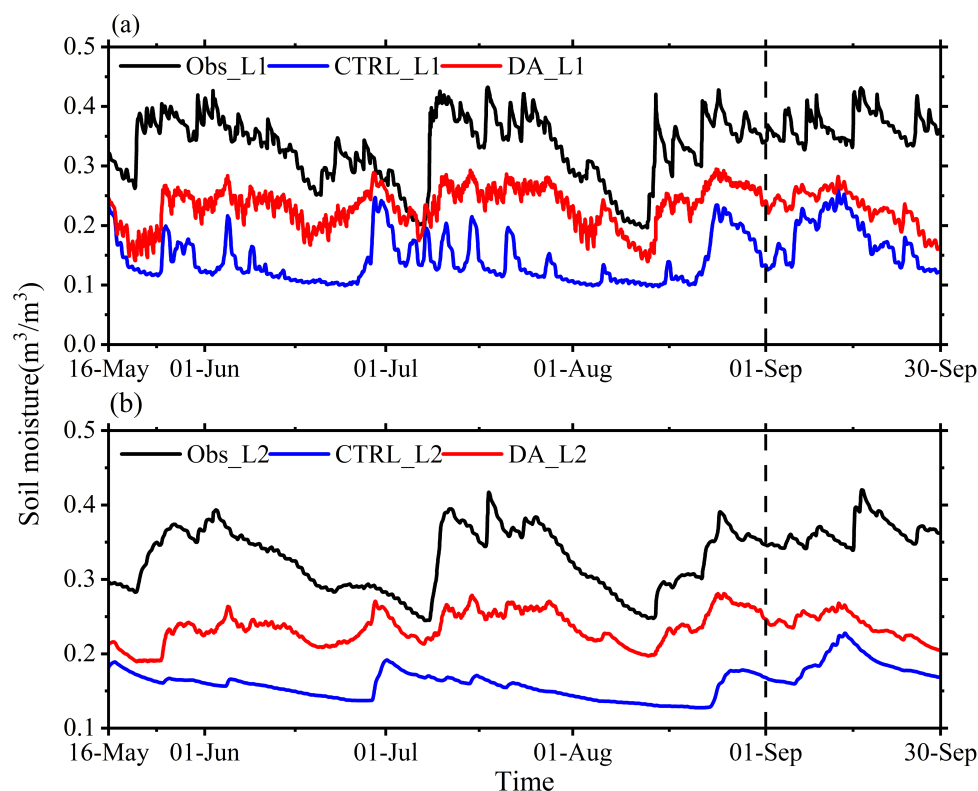
Figure 10: Hourly variations of the (a) 0-7 cm and (b) 7-28 cm soil moisture in the observation (black solid lines), DA experiment

420 **(red solid lines) and CTRL experiment (blue solid lines) at a single station in the Tianshan Mountains region of Xinjiang Province**

from May 16 to September 30, 2016. The prediction period is after the vertical dashed line.

The improvement of soil moisture after image assimilation is further evaluated based on the root mean square error

(RMSE). Figure 11 shows the hourly RMSE variations of surface and subsurface soil moisture in DA and CTRL

425 experiments. As can be seen, the RMSE of surface soil moisture in the CTRL experiment is larger with mean value of 0.16

m³·m⁻³, and the RMSE fluctuates considerably due to the influence of additional factors. The RMSE is fundamentally

reduced by about 0.04 m³·m⁻³ after image assimilation, which also indicates that the image assimilation not only optimizes the spatial distribution structure of soil moisture, but also has a certain improvement effect on the soil moisture values. In the prediction period, the surface layer is more disturbed by atmospheric forcing, so the RMSE at surface layer gradually

430   increases with time, but the RMSE of DA experiment is also consistently smaller than that of the CTRL experiment.

The RMSEs of subsurface soil moisture in both experiments are smaller than that of the surface soil moisture. Although the initial error is larger, but it gradually decreases with time and shows a stable variation. The mean RMSE of the subsurface soil moisture in the CTRL experiment is about 0.15 m³·m⁻³, while it reduces to 0.12 m³·m⁻³ after assimilation. Similarly, the RMSEs of the DA experiment are steadily less than those of the CTRL experiment in both cyclic assimilation

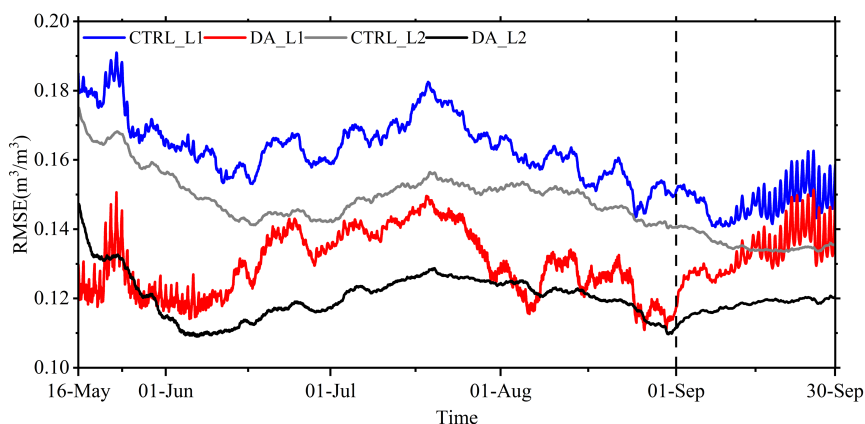435   and prediction periods.

Figure 11: Hourly variations of the soil moisture RMSEs for the DA (red and black solid lines) and CTRL experiments (gray and blue solid lines) at the surface (blue and red solid lines) and subsurface layers (black and gray solid lines) in the study area.

440   **5 Discussion and conclusions**

Currently, the dominant land surface assimilation systems are mostly single-column assimilation systems. Due to the strong spatial heterogeneity of the land surface, the continuity of the spatial structure of soil variables is frequently overlooked. Improving the spatial structure accuracy of soil moisture anomalies will help to improve the predictability of short-term climate change caused by soil moisture anomalies.

445   Nowadays, more and more information is presented in the form of images or sequences of images. In this study, we developed an image assimilation system to quantitatively utilize the observed structural information of the variables presented in the image, so as to improve the spatial structure accuracy of the initial soil moisture field. This image assimilation system is mainly based on the 3D variational assimilation framework with the image term added as a constraint, and it is constructed by using the anisotropic and multiscale curvelet transform methods as image structure operator. By

450 efficiently extracting valuable structural information from observed images, the spatial structure of the soil moisture analysis field is remarkably optimized, and the accuracy of soil moisture quantities is also improved.

The performance of the image assimilation system is verified based on the ideal experiment. It is found that the image assimilation system can effectively and reasonably improve the spatial structure of the analysis field by assimilating the observed images of surface soil moisture. The SCC between the analysis field and observation has increased from 0.39 to

455 0.67, and the RMSE has reduced from 0.16 $m^3 \cdot m^{-3}$ to 0.12 $m^3 \cdot m^{-3}$. At the same time, the spatial structure of the subsurface soil moisture is further improved after efficiently optimizing the surface soil moisture structure, under the reasonable constraints of the model dynamics and thermal processes. The SCC increases from 0.35 to 0.57, and the RMSE is reduced from 0.15 $m^3 \cdot m^{-3}$ to 0.13 $m^3 \cdot m^{-3}$. During the entire assimilation and prediction period, the SCCs and RMSEs of the surface and subsurface layer in the DA experiment are better than in the CTRL experiment.

460 It should be pointed out that the image assimilation system constructed in this study can remarkably improve the spatial structure similarity between the analysis field and the observed image by assimilating the effective structure observations in the image, and improve the simulation accuracy of soil moisture in the land surface process model as well.

However, at present, this study is only based on ideal experiments to verify the performance of image assimilation. In the subsequent study, the direct assimilation of geostationary and polar-orbiting satellite brightness temperature observation

465 images based on the image assimilation system will be considered. Then, we can obtain a higher precision and more accurate spatially-structured soil-moisture initial field, and further improve the short-term climate prediction level. In this way, the great application potential of image assimilation can be better demonstrated.

Author contributions. Conceptualization, Z.Q. and W.S. and Z.L.; methodology, W.S.; software, W.S. and Z.Q.; validation, Z.Q., W.S. and Z.L.; formal analysis, Z.L.; resources, Z.L. and J.L.; data curation, W.S.; writing—original draft preparation,

480 W.S.; writing—review and editing, Z.Q. and J.L.; visualization, W.S.; supervision, Z.L. and J.L.; project administration, Z.Q.; funding acquisition, Z.Q. and Z.L. All authors have read and agreed to the published version of the manuscript.

References

490   Crow, W. T., Gomez, C. A., Sabater, J. M., Holmes, T., Hain, C. R., Lei, F., Dong, J., Alfieri, J. G., and Anderson, M. C.: Soil moisture–evapotranspiration overcoupling and L-Band brightness temperature assimilation: Sources and forecast implications, J Hydrometeorol, 21, 2359-2374, https://doi.org/10.1175/JHM-D-20-0088.1, 2020.

Cui, Z., and Wang, C.: Improvement of summer precipitation simulation with indirect assimilation of spring soil moisture over the Tibetan Plateau, Q J R Meteorol Soc, 148, 3231-3251, https://doi.org/10.1109/JSTARS.2021.3118595, 2022.

495   Dai, Y., Zeng, X., Dickinson, R., Baker, I., Bonan, G. B., Bosilovich, M. G., Denning, S., Dirmeyer, P., Houser, P., Niu, G., Oleson, K., Schlosser, A., and Yang, Z.: The Common Land Model, Bull Am Meteorol Soc, 84, 1013–1023, https://doi.org/10.1175/BAMS-84-8-1013, 2003.

De Santis, D., Biondi, D., Crow, W. T., Camici, S., Modanesi, S., Brocca, L., and Massari, C.: Assimilation of satellite soil moisture products for river flow prediction: An extensive experiment in over 700 catchments throughout Europe, Water Resour Res, 57, e2021WR029643, https://doi.org/10.1029/2021WR029643, 2021.

500   Dirmeyer, P. A., Balsamo, G., Blyth, E. M., Morrison, R., and Cooper, H. M.: Land-atmosphere interactions exacerbated the drought and heatwave over Northern Europe during summer 2018, AGU Adv, 2, e2020AV000283, https://doi.org/10.1029/2020AV000283, 2021.

Douville, H., Viterbo, P., Mahfouf, J. F., and Beljaars, A. C. M.: Evaluation of the optimum interpolation and nudging
505   techniques for soil moisture analysis using FIFE data, Mon Weather Rev, 128: 1733-1756, https://doi.org/10.1175/1520-0493(2000)128<1733:EOTOIA>2.0.CO;2, 2000.

Dunne, S., and Entekhabi, D.: An ensemble-based reanalysis approach to land data assimilation, Water Resour Res, 41, https://doi.org/10.1029/2004WR003449, 2005.

Gruber, A., Crow, W., and Dorigo, W.: Assimilation of spatially sparse in situ soil moisture networks into a continuous
510   model domain, Water Resour Res, 54, 1353-1367, https://doi.org/10.1002/2017WR021277, 2018.

Ji, D., Wang, L., Feng, J., Wu, Q., Cheng, H., Zhang, Q., Yang, J., Dong, W., Dai, Y., Gong, D., Zhang, R.-H., Wang, X., Liu, J., Moore, J. C., Chen, D., and Zhou, M.: Description and basic evaluation of Beijing Normal University Earth System Model (BNU-ESM) version 1, Geosci. Model Dev., 7, 2039–2064, https://doi.org/10.5194/gmd-7-2039-2014, 2014.

515 Jia, B., Xie, Z., Tian, X., and Shi, C.: A soil moisture assimilation scheme based on the ensemble Kalman filter using microwave brightness temperature. Sci China Ser D-Earth Sci, 52: 1835–1848, https://doi.org/10.1007/s11430-009-0122-z, 2010.

Jin, R., and 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, 52, 1732-1745,
520   https://doi.org/10.1007/s11430-009-0174-0, 2009.

Khaki, M., Hendricks, F., and Han, S.: Multi-mission satellite remote sensing data for improving land hydrological models via data assimilation, Sci Rep, 10, 18791, https://doi.org/10.1038/s41598-020-75710-5, 2020.

Koster, R. D., Schubert, S. D., DeAngelis, A. M., Molod, A. M., and Mahanama, s. p.: Using a simple water balance framework to quantify the impact of soil moisture initialization on subseasonal evapotranspiration and air temperature
525   forecasts. J Hydrometeorol, 21: 1705-1722. https://doi.org/10.1175/JHM-D-20-0007.1, 2020.

Le Dimet, F.-X., Souopgui, I., Titaud, O., Shutyaev, V., and Hussaini, M. Y.: Toward the assimilation of images, Nonlin. Processes Geophys., 22, 15–32, https://doi.org/10.5194/npg-22-15-2015, 2015.

Lei, H., Lin, Z., and Sun, J.: An Improved Dust Storm Prediction System and Its Simulation Experiments. Climatic and Environmental Research (in Chinese), 10, 669-683, https://doi.org/10.3878/j.issn.1006-9585.2005.03.33, 2005.

530 Li, K., Zhang, J., Yang, K., and Wu, L.: The role of soil moisture feedbacks in future summer temperature change over East Asia, J Geophys Res Atmos, 124: 12034-12056, https://doi.org/10.1029/2018JD029670, 2019.

Li, P., Zha, Y., Shi, L., Tso, C., Zhang, Y., and Zeng, W.: Comparison of the use of a physical-based model with data assimilation and machine learning methods for simulating soil water dynamics, J Hydrol, 584: 124692, https://doi.org/10.1016/j.jhydrol.2020.124692, 2020b.

535 Li, X.: Characterization, controlling, and reduction of uncertainties in the modeling and observation of land-surface systems, Sci China Earth Sci, https://doi.org/10.1007/s11430-013-4728-9, 2013.

Li, X., Liu, F., and Fang, M.: Harmonizing models and observations: Data assimilation in Earth system science, Sci China Earth Sci, 63, 1059–1068, https://doi.org/10.1007/s11430-019-9620-x, 2020a.

Lin, Z., Liu, H., Xie, Z., Wang, A., and Liu, S.: Recent Progress in the Land-Surface and Hydrological Process
540   Studies, Chinese Journal of Atmospheric Sciences, 32, 935-949, https://doi.org/10.3878/j.issn.1006-9895.2008.04.19, 2008.

Lin, Z., Yang, X., and Guo, Y.: Sensitivity of Land Surface Model to the Initial Condition of Soil Moisture, Climatic and Environmental Research (in Chinese), 6: 240-248, https://doi.org/10.3878/j.issn.1006-9585.2001.02.17, 2001.

Mahfouf JF.: Analysis of Soil Moisture from Near-Surface Parameters: A Feasibility Study, J Appl. Meteorol, 7: 506-526,
545   https://doi.org/10.1175/1520-0450(1991)030<1534:AOSMFN>2.0.CO;2, 1991.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo,

C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth Syst. Sci. Data, 13, 4349–4383, https://doi.org/10.5194/essd-13-4349-2021, 2021.

550   Muñoz-Sabater, J., Lawrence, H., Albergel, C., Rosnay, P., Isaksen, L., Mecklenburg, S., Kerr, Y., and Drusch, M.: Assimilation of SMOS brightness temperatures in the ECMWF Integrated Forecasting System, Q J R Meteorol Soc, 145, 2524-2548, https://doi.org/10.1002/qj.3577, 2019.

Naz, B. S., Kollet, S., Franssen, H-J. H., Montzka, T., and Kurtz, W.: A 3 km spatially and temporally consistent European daily soil moisture reanalysis from 2000 to 2015, Sci Data, 7, 111, https://doi.org/10.6084/m9.figshare.11993547, 2020.

555   Qian, T., Dai, A., Trenberth, K. E., and Oleson, K. W.: Simulation of global land surface conditions from 1948 to 2004. Part I: Forcing data and evaluations, J. Hydrometeorol., 7, 953– 975, doi:10.1175/JHM540.1., 2006.

Reichle, R. H., McLaughlin, D. B., and Entekhabi, D.: Variational data assimilation of microwave radiobrightness observations for land surface hydrology applications, IEEE Trans Geosci Remote Sens, 39: 1708–1718, https://doi.org/10.1109/36.942549, 2001.

560   Reichle, R. H., Zhang, S. Q., Liu, Q., Draper, C. S., Kolassa, J., and Todling, R.: Assimilation of SMAP brightness temperature observations in the GEOS land-atmosphere data assimilation system, IEEE J Sel Top Appl Earth Obs Remote Sens, 14: 10628-10643, https://doi.org/10.1109/JSTARS.2021.3118595, 2021.

Shi, C., Xie, Z., Qian, H., Liang, M., and Yang, X.: China land soil moisture EnKF data assimilation based on satellite remote sensing data, Sci China Earth Sci, 41, 375-385, https://doi.org/10.1007/s11430-010-4160-3, 2011.

565   Stroud, J. R., Lesht, B. M., Schwab, D. J., Beletsky, D., and Stein, M. L.: Assimilation of satellite images into a sediment transport model of Lake Michigan, Water Resour Res, 45, https://doi.org/10.1029/2007WR006747, 2009.

Tangdamrongsub, N., Han, S-C., Yeo, I-Y., Dong, J., Steele-Dunne, S., Willgoose, G., and Walker, J.: Multivariate data assimilation of GRACE, SMOS, SMAP measurements for improved regional soil moisture and groundwater storage estimates, Adv Water Resour, 135, 103477, https://doi.org/10.1016/j.advwatres.2019.103477, 2020.

570   Tian, X., Xie, Z., and Dai, A.: A land surface soil moisture data assimilation system based on the dual-UKF method and the Community Land Model, J Geophys Res Atmos, 113(D14), https://doi.org/10.1029/2007JD009650, 2008.

Titaud, O., Vidard, A., and Souopgui, I.: Assimilation of image sequences in numerical models, Tellus A, 62, 30-47, https://doi.org/10.1111/j.1600-0870.2009.00416.x, 2010.

Vidard, A., Le Dimet, F. X., Souopgui, I., and Titaud, O.: Assimilation of Images in Numerical Models in Geophysics, in:
575   EngOpt 2008 - International Conference on Engineering Optimization, COPPE/UFRJ, Rio de Janeiro, Brazil, June 2008, (inria-00319972).

Wang, C., and Cui, Z.: Improvement of short-term climate prediction with indirect soil variables assimilation in China, J Clim, 31, 1399-1412, https://doi.org/10.1175/JCLI-D-17-0210.1, 2018.

Xu, Z., Wu, Z., He, H., Guo, X., and Zhang, Y.: Comparison of soil moisture at different depths for drought monitoring
580   based on improved soil moisture anomaly percentage index, Water Sci Eng, 14, 171-183, https://doi.org/10.1016/j.wse.2021.08.008, 2021.

Yang, L., Wei, W., Wang, T., and Li, L.: Temporal-spatial variations of vegetation cover and surface soil moisture in the growing season across the mountain-oasis-desert system in Xinjiang, China, Geocarto Int., 37, 3912-3940, https://doi.org/10.1080/10106049.2021.1871666, 2022.

585   Yuan, X., and Liang, X.-Z.: Evaluation of a Conjunctive Surface-Subsurface Process model (CSSP) over the contiguous United States at regional-local scales, J Hydrometeorol, 12, 579–599, https://doi.org/10.1175/2010JHM1302.1, 2011.

Yuan, Y., Lai, X., Gong, Y., and Chen, J.: The impacts of late spring soil moisture in the Tibetan Plateau on summer precipitation in eastern China, Int J Climatol., 41: 862-877, https://doi.org/10.1002/joc.6692, 2021.

Yuan, Z., and Wei, N.: Coupling a New Version of the Common Land Model (CoLM) to the Global/Regional Assimilation

590   and Prediction System (GRAPES): Implementation, Experiment, and Preliminary Evaluation, *Land, 11*, 770, https://doi.org/10.3390/land11060770, 2022.

Zhan, Y., and Lin, Z.: The Relationship Between June Precipitation over Mid-Lower Reaches of the Yangtze River Basin and Spring Soil Moisture over the East Asian Monsoon Region, Acta Meteorologica Sinica, 25: 355-363, https://doi.org/10.1007/s13351-011-0310-6, 2011.

595   Zhang, H., Zhang, M., Jin, J., Fei, K., Ji, D., Wu, C., Zhu, J., He, J., Chai, Z., Xie, J., Dong, X., Zhang, D., Bi, X., Cao, H., Chen, H., Chen, K., Chen, X., Gao, X., Hao, H., Jiang, J., Kong, X., Li, S., Li, Y., Lin, P., Lin, Z., Liu, H., Liu, X., Shi, Y., Song, M., Wang, H., Wang, T., Wang, X., Wang, Z., Wei, Y., Wu, B., Xie, Z., Xu, Y., Yu, Y., Yuan, L., Zeng, Q., Zeng, X., Zhao, S., Zhou, G., and Zhu, J.: CAS-ESM 2: Description and climate simulation performance of the Chinese Academy of Sciences (CAS) Earth System Model (ESM) version 2, J Adv Model Earth Syst, 12,

600   e2020MS002210, https://doi.org/10.1029/2020MS002210,2020.

Zheng, W., Zhan, X., Liu, J., and Ek, M.: A preliminary assessment of the impact of assimilating satellite soil moisture data products on NCEP Global Forecast System, 2018, Adv Meteorol, https://doi.org/10.1155/2018/7363194, 2018.

Zhong, J., Lu, B., Wang, W., Huang, C., and Yang, Y.: Impact of soil moisture on winter 2-m temperature forecasts in northern China, J Hydrometeorol, 21, 597-614, https://doi.org/10.1175/JHM-D-19-0060.1, 2020.

605   Zhou, J., Wu, Z., Crow, W. T., Dong, J., and He, H.: Improving spatial patterns prior to land surface data assimilation via model calibration using SMAP surface soil moisture data, Water Resour Res, 56, e2020WR027770, https://doi.org/10.1029/2020WR027770, 2020b.

Zhou, J., Zuo, Z., Rong, X.: Comparison of the effects of soil moisture and El Niño on summer precipitation in eastern China, Sci China Earth Sci, 63, 267-278, https://doi.org/10.1007/s11430-018-9469-6, 2020a.

610