SCIENCE CHINA
Earth Sciences

• RESEARCH PAPER •

September 2011 Vol.54 No.9: 1430–1440
doi: 10.1007/s11430-010-4160-3

China land soil moisture EnKF data assimilation based on satellite remote sensing data

SHI ChunXiang[1*], XIE ZhengHui[2], QIAN Hui[3], LIANG MiaoLing[4] & YANG XiaoChun[1]

[1] *National Satellite Meteorological Center, China Meteorological Administration, Beijing 100081, China;*
[2] *Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China;*
[3] *Institute of Geology, Chinese Academy of Geological Sciences; Key Geodynamics Laboratory, the Ministry of Land and Resources, Beijing 100037, China;*
[4] *National Meteorological Center, China Meteorological Administration, Beijing 100081, China*

Received December 8, 2009; accepted June 18, 2010; published online March 4, 2011

Soil moisture plays an important role in land-atmosphere interactions. It is an important geophysical parameter in research on climate, hydrology, agriculture, and forestry. Soil moisture has important climatic effects by influencing ground evapotranspiration, runoff, surface reflectivity, surface emissivity, surface sensible heat and latent heat flux. At the global scale, the extent of its influence on the atmosphere is second only to that of sea surface temperature. At the terrestrial scale, its influence is even greater than that of sea surface temperatures. This paper presents a China Land Soil Moisture Data Assimilation System (CLSMDAS) based on EnKF and land process models, and results of the application of this system in the China Land Soil Moisture Data Assimilation tests. CLSMDAS is comprised of the following components: 1) A land process model—Community Land Model Version 3.0 (CLM3.0)—developed by the US National Center for Atmospheric Research (NCAR); 2) Precipitation of atmospheric forcing data and surface-incident solar radiation data come from hourly outputs of the FY2 geostationary meteorological satellite; 3) EnKF (Ensemble Kalman Filter) land data assimilation method; and 4) Observation data including satellite-inverted soil moisture outputs of the AMSR-E satellite and soil moisture observation data. Results of soil moisture assimilation tests from June to September 2006 were analyzed with CLSMDAS. Both simulation and assimilation results of the land model reflected reasonably the temporal-spatial distribution of soil moisture. The assimilated soil moisture distribution matches very well with severe summer droughts in Chongqing and Sichuan Province in August 2006, the worst since the foundation of the People's Republic of China in 1949. It also matches drought regions that occurred in eastern Hubei and southern Guangxi in September.

EnKF land data assimilation, AMSR-E soil moisture, FY2C geostationary satellite, high-resolution precipitation, surface incident solar radiation

Citation: Shi C X, Xie Z H, Qian H, et al. China land soil moisture EnKF data assimilation based on satellite remote sensing data. Sci China Earth Sci, 2011, 54: 1430−1440, doi: 10.1007/s11430-010-4160-3

Soil moisture affects matter and energy balances between land and atmosphere through influencing moisture flux, sensible heat, and latent heat flux. Accurate estimates of the spatial and temporal distribution of soil moisture are of vital importance to further understanding the ecological and physical processes of land and land-atmosphere interactions, and are of great importance in the research and application of meteorology in environmental management, ecology, hydrology, and agriculture. Soil moisture information can currently be acquired with the following methods: regular and field observations, satellite remote sensing, and land

*Corresponding author (email: shicx@cma.gov.cn)

© Science China Press and Springer-Verlag Berlin Heidelberg 2011

(C)1994-2019 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

hydrological modeling and simulation. Soil moisture data acquired through regular observation has a low temporal frequency and few spatial points, and using some field observations is interim and limited to a small range. Regular and field soil moisture observations provide only point-based data and cannot represent spatial variations in the soil moisture profile. Synchronous field observation is usually expensive. Satellite remote sensing can provide global soil moisture data with a high spatial and temporal resolution, which cannot be achieved via the regular observation network. However, soil moisture observations based on satellite remote sensing are also problematic. Precision of soil moisture inversion is related to soil type, ground surface roughness, and vegetation coverage. The uncertainty in these parameters results in inversion errors and insufficient temporal and spatial resolutions. Soil moisture obtained through land hydrological modeling simulation and calculation has good temporal frequency and spatial distribution. However, its precision is influenced greatly by model structure and input data. Land assimilation technology serves as an effective solution to these problems [1, 2]. It is a technical approach that integrates optimally physical process information of the land model.

Data assimilation methods used currently in atmospheric, marine, and land data assimilation systems include mainly the optimal interpolation method, 3-D variation method, 4-D variation method, Kalman filter, and EnKF. EnKF is an assimilation method is used widely in land model assimilation systems, and uses the Monte Carlo method (overall integration method) to calculate predicted error covariance of the state. It was proposed by Evensen on the basis of Epstein's stochastic dynamic prediction theory [3]. The model state prediction is considered an approximately stochastic dynamic prediction. An overall state is used to represent the probability density function in a stochastic dynamic prediction. By integrating ahead, we can calculate easily the statistical characteristics (e.g., mean value and covariance) corresponding to probability density functions at various times of the overall state. EnKF key feature is that it does not require linearized model operators or observation operators as the Kalman filter does. Huge and low-efficiency computation requirements are its greatest drawbacks. As all current land models are single-column models, they contain much fewer model state variables than atmospheric and marine models and thus have a higher computational efficiency. Strong non-linear characteristics in land process models are significant.

Starting in 1998, the Department of Hydrological Sciences and Data Assimilation Office of US NASA Goddard Space Flight Center organized many organizations to initiate the research of Land Data Assimilation System (LDAS), and developed the Global Land Data Assimilation System (GLDAS). US National Oceanic and Atmospheric Administration also offers support for hydrology-based LDAS research at the scale of the North American continent (North America Land Data Assimilation System project). Land data assimilation uses many kinds of new satellite and ground-based observation data to generate optimal land state and flux data (http://www.knmi.nl/samenw/LDAS/). In 2001, Europe started research on the European Land Data Assimilation System. The main purpose of ELDAS research is to improve forecasting and monitoring of floods and droughts. The ELDAS project conducted to: assemble specialists in soil moisture assimilation to design and produce a unified, flexible, and practical data assimilation framework; use independent observation data to verify the assimilated soil moisture field; evaluate the seasonal water circulation forecast after soil moisture data assimilation; evaluate flood risk; establish an exemplary data set covering at least one season over the European region; use data from new satellite platforms MSG (MTEOSAT Second Generation) and SMOS (ESA Soil Moisture/Ocean Salinity Mission); and make European contributions to the Global Land Data Assimilation System. ELDAS has been applied in four numerical forecasting centers (ECMWF, DWD, CNRM and INM) and already been brought from the research stage into the service operation stage (http://www.knmi.nl/samenw/eldas/). The West China Land Data Assimilation System (WCLDAS) was researched and developed by the Cold and Arid Regions Environmental and Engineering Research Institute, Chinese Academy of Sciences and Department of Atmospheric Sciences, School of Resources and Environment, Lanzhou University. The institute established a single-point soil moisture assimilation system based on EnKF and SiB2 models, and completed preliminary analysis and evaluation of this assimilation system [4, 5]. Yang et al. [6] developed an automatic rating system used to estimate moisture and energy balances of soil by assimilating AMSR-E vertical polarization 6.9 and 18.7 GHz luminance temperature data. Tian et al. [7–13], Zhang et al. [14, 15], and Jia et. al. [16] researched the improvement of land data assimilation methods and the rating of microwave luminance temperature assimilation observation operator models.

This paper explains the composition of the China Land Soil Moisture Data Assimilation System (CLSMDAS), processing methods and quality inspection of atmospheric driving data, the performance test of CLSMDAS with single-point observation data, error analysis of observed soil moisture data, and design, and analysis of the CLSMDAS experiment.

1 China Land Soil Moisture Data Assimilation System

CLSMDAS is comprised mainly of the following (Figure 1): 1) Land model: A NCAR-CLM3.0 land model used widely at present; 2) Driving data: High spatial and temporal resolution precipitation estimates and ground-incident solar radiation data acquired from the FY2 geostationary meteorological satellites, and surface air temperatures, humidity,

(C)1994-2019 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

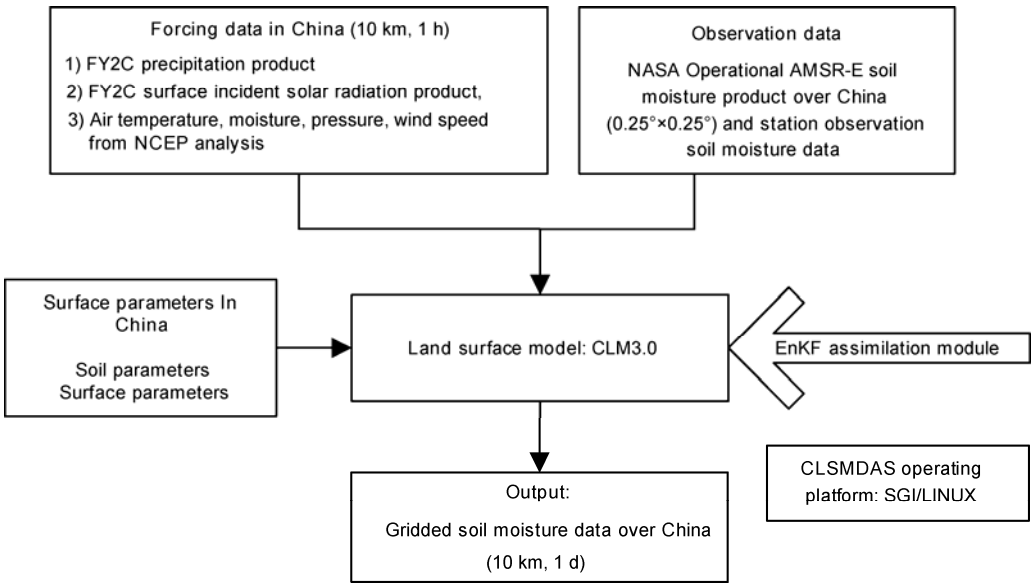

Figure 1 Structure of the China Land Soil Moisture Data Assimilation System.

atmospheric pressure and wind speed in the NCEP reanalysis data set are interpolated temporally and spatially to construct atmospheric driving data of the land model; 3) Data assimilation method: We have chosen the EnKF assimilation method which is suitable for land data assimilation; 4) Observation data: Soil moisture inversion data acquired from AMSR-E carried by the AQUA EOS satellite, and soil moisture data observed at ground level, are used; 5) Output data set: Assimilated soil moisture grid point data can be output. CLSMDAS works at present on a SGI-LINUX platform. Its individual components are described below.

1.1 Land surface model

The land surface model is the core of the land data assimilation system. The extent to which the land surface model describes accurately and reasonably the energy and material exchanges at the ground surface has a direct influence on the output of the land assimilation system, forecasting the state of the next iteration. A land process model—Community Land Model Version 3.0 (CLM3.0 for short)—developed by US National Center for Atmospheric Research (NCAR) [17] was used here.

The CLM3.0 model is designed mainly for coupling with the atmospheric numerical model and provides the surface albedo (direct and scattered light within the visible and infrared bands), upward long-wave radiation, sensible heat flux, latent heat flux, water vapor flux, and east-to-west and south-to-north surface stress needed by the atmospheric model. These parameters are controlled by many ecological and hydrological processes. The model simulates the phenology of leaves and physiological and water circulations of pores. Ecological differences between vegetation types and thermal and hydraulic differences between different soil types are also considered. Each grid cell can be covered by several types of land surfaces. The river transport model transfers downward to the sea. Because this land model has to be coupled with the climate model and numerical weather forecast model, compromise has to be made between effectiveness and complexity of the parametric computation of land processes. Land spatial nonuniformity is realized by nesting subgrids where one grid cell contains several land units, snow and soil cylindrical blocks, and different types of vegetation. Each cell contains several land units, each land unit contains a different number of soil and snow cylindrical blocks, and each cylindrical block may contain several types of vegetation functions [18]. Many researchers have tested and evaluated the application of the CLM3.0 model in China and proved the feasibility of its application in China [17].

The EnKF assimilation module is coupled mainly with the soil moisture module. Below is the equation of one-dimensional soil water vertical motion (horizontal flow is ignored) used in CLM3.0:

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

where θ is the volumetric water content of soil (m^3/m^3), q is vertical soil water flux (mm/s), E is evaporation rate (mm/s). R_{fm} is melting (negative) or freezing (positive) rate, and z is vertical distance to the ground surface (positive under q and z directions) [17].

1.2 Atmospheric driving data

One of the difficulties preventing us improving the land model simulation precision is lack of long-term, high-resolution, observational atmospheric driving data. Atmospheric

(C)1994-2019 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

driving data usually requires diurnal precipitation, atmospheric temperature, relative humidity, atmospheric pressure, wind speed and downward solar radiation data, which cannot be observed directly at the global scale. Much work has been done on processing atmospheric driving data for offline land surface models. Qian et al. [19] showed that there are false long-term variations with NCEP's reanalysis of precipitation and solar radiation. The errors in precipitation data will affect inevitably the effects of soil moisture prediction, simulation, and assimilation. Other researchers have also realized these problems, and as a result either directly use observation or adjust the reanalysis data.

In this paper, inverted precipitation and ground-incident solar radiation products acquired through high spatial and temporal resolution FY2C satellite data were introduced and combined with the NCEP reanalysis data to establish a set of high-quality atmospheric driving data which covers the Chinese region.

1.2.1 High spatial- and temporal-resolution precipitation

FY2C/D/E satellite precipitation estimate outputs can be downloaded free of charge from the website of the China Satellite Data Service Center (http://satellite.cma.gov.cn/). Precipitation estimate outputs include daily accumulative and 6-hour accumulative precipitation estimate outputs. A new hourly accumulative precipitation estimate output was made available in September 2008. The integral time step of the CLM3 land model is 30 min. Therefore, we need atmospheric forcing data with high temporal resolution. Compared with the 6-hour accumulative precipitation data, the hourly precipitation data is more suitable to drive simulation of the land model. To this end, Shi and Xie [20] developed a method to perform time-weighted interpolation of accumulative precipitation on the basis of hourly geostationary satellite cloud category information. This was applied to the FY2C 6-hour precipitation estimate output provided by the China Satellite Data Service Center, to derive a precipitation data set with 0.1°×0.1° spatial resolution and 1-h temporal resolution, and used automatic rainfall observation data from China to verify and evaluate the precipitation estimate data. The result indicated that the method was reasonable. See Figure 2 for the process flow of geostationary satellite 1-hour precipitation output. See ref. [21] for details on the processing and quality evaluation of precipitation data.

1.2.2 Incident solar radiation data with high spatial and temporal resolution

FY2C/D/E incident solar radiation output can be downloaded free of charge from the website of China Satellite Data Service Center (http://satellite.cma.gov.cn/). The temporal resolution of the product is 1 day. The spatial resolution is 0.5°×0.5°. Due to its low spatial and temporal

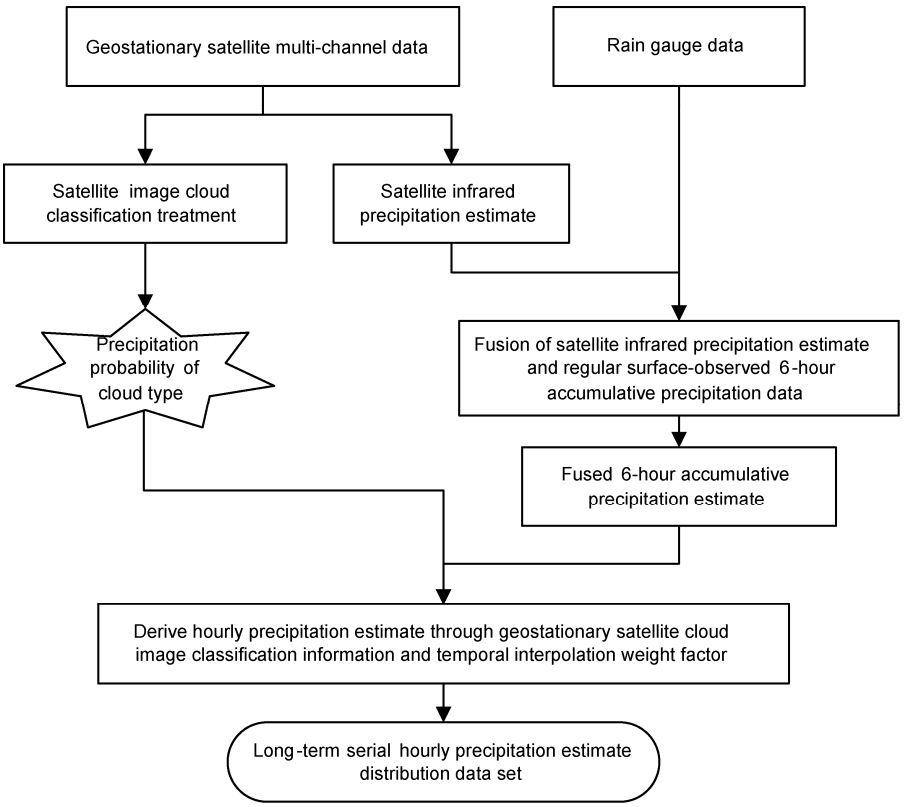

Figure 2 Precipitation process flow with high spatial and temporal distribution.

(C)1994-2019 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

resolution, the output does not meet the requirements of the land surface model for atmospheric forcing data. Therefore, an inversion algorithm was used here on the FY2C/D/E ground-incident solar radiation output of the China Satellite Data Service Center (http://satellite.cma.gov.cn/). Visible-band observations acquired through the FY2C geostationary meteorological satellite are used to generate a ground-incident solar radiation data set with a temporal resolution of 1 hour and a spatial resolution of 0.1°×0.1°. The quality of this data set was verified with the solar radiation data from five China Meteorological Administration climatic observation stations.

The discrete ordinate method proposed by Stamnes et al. [22] was used to calculate radiation transfer in the inversion algorithm for the ground-incident solar radiation output. This algorithm calculates the radiance of any direction and thus gives consideration to the anisotropy of solar radiation reflected at the top of the atmospheric layer. It first calculates the radiance of solar radiation reflected towards the satellite observation direction at the top of the atmospheric layer, and then converts it into the visible two-way albedo observed through the visible light channel of the satellite. The period in which the solar radiation incident at the top of the atmospheric layer penetrates the atmosphere and reaches the ground surface contains a series of physical interaction processes with the atmosphere and the ground surface. The inversion model considers mainly: 1) ozone absorption; 2) repeated Rayleigh scattering of molecules; 3) repeated scattering and absorption of cloud droplets; 4) water vapor absorption; 5) repeated scattering and absorption of aerosol; and 6) repeated reflection between the ground surface and the atmosphere. Similar to the work of Stuhlmann et al. [23], we designed a 5-layer planoparallel ideal atmospheric model which is not uniform in the vertical direction. It is divided into five solar spectral intervals (0.2–0.4, 0.4–0.5, 0.5–0.6, 0.6–0.7, 0.7–4.0 μm) to calculate the scattering, absorption, and reflection of solar radiation that take place within them. Judged by the extents to which the above- mentioned physical processes influence the ground-incident solar radiation, cloud has a greater effect than all other factors by one order of magnitude. Ground-incident solar radiation is determined mainly by cloud, which needs to be stressed in the inversion model.

Yuan Wanping performed many tests and verifications of the inversion algorithm of ground-incident solar radiation (http://satellite.cma.gov.cn/). In this paper, the ground-based observation data provided by newly-built climatic observation stations of China Meteorological Administration was used for a comparative verification of the FY2C ground-incident solar radiation data between July 2005 and June 2009. The five climatic observation stations were located in Xilinhot, Shouxian, Zhangye, Dali, and Dianbai. Figure 3(a) shows time-varying ground-incident solar radiation data observed at Xilinhot Observation Station and inverted from FY2C satellite data. The horizontal coordinate is time, the unit of measurement (UOM) is hours; the accumulative total counts from 00:00 August 1, 2007 (universal time). The vertical coordinate is ground-incident solar radiation, the UOM is W/m^2. Figure 3(a) shows that when the value of ground-incident solar radiation is high, the value observed at ground level is greater than the satellite-inverted ground-incident solar radiation. This is probably because satellite observation covers a relatively large area while ground-based observation covers a very limited area, which can almost be considered a point. Therefore, the result of satellite inversion is closer to the mean state. Figure 3(b) is a scatter diagram of ground-incident solar radiation data observed at ground level and inverted by the FY2C satellite. Figure 3(c) is a histogram of the difference between ground-incident solar radiation data inverted by the FY2C satellite and observed at ground level. Both Figure 3(b) and (c) show that satellite-inverted values are higher than ground-observed values. Figure 3(c) also shows that under most circumstances the difference between satellite-inverted data and ground-observed data of solar radiation is approximately <50 W/m^2.

1.2.3 Atmospheric driving data set

Temperature, humidity, atmospheric pressure, and wind speed that drive operation of the CLM3.0 land model were derived from interpolation of NCEP reanalysis data with 1°×1° resolution and 6-h interval. These parameters were combined with precipitation and ground-incident solar radiation data, with high spatial and temporal distribution inverted through geostationary satellite to form an atmospheric forcing data set that drives the land model.

In this paper, the data between July 2005 and June 2009 were processed and then subject to quality control and verification. An atmospheric driving data set is generated (UOM of one month). Temporal resolution of the data set is 1 hour, the horizontal resolution is 0.1°×0.1°, the spatial coverage is 15°–55°N, 75°–135°E, and the data format is NETCDF. This atmospheric driving data set was used in the China Soil Moisture Land Data Assimilation experiment described below. A 0.1°×0.1° atmospheric driving data set was used as the basic data. Data corresponding to assimilation experiments with other spatial resolutions were derived through spatial interpolation of this atmospheric driving data.

1.3 Land data assimilation method

Evensen proposed a new solution to square root analysis [24] on the basis of standard EnKF [25]. It does not require observation disturbance in the calculation of analytic field sets and will thus reduce or eliminate sample error brought by observation disturbance. This algorithm also does not require additional assumptions or approximations in calculation of the analytic field structures. For example, it neither

(C)1994-2019 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

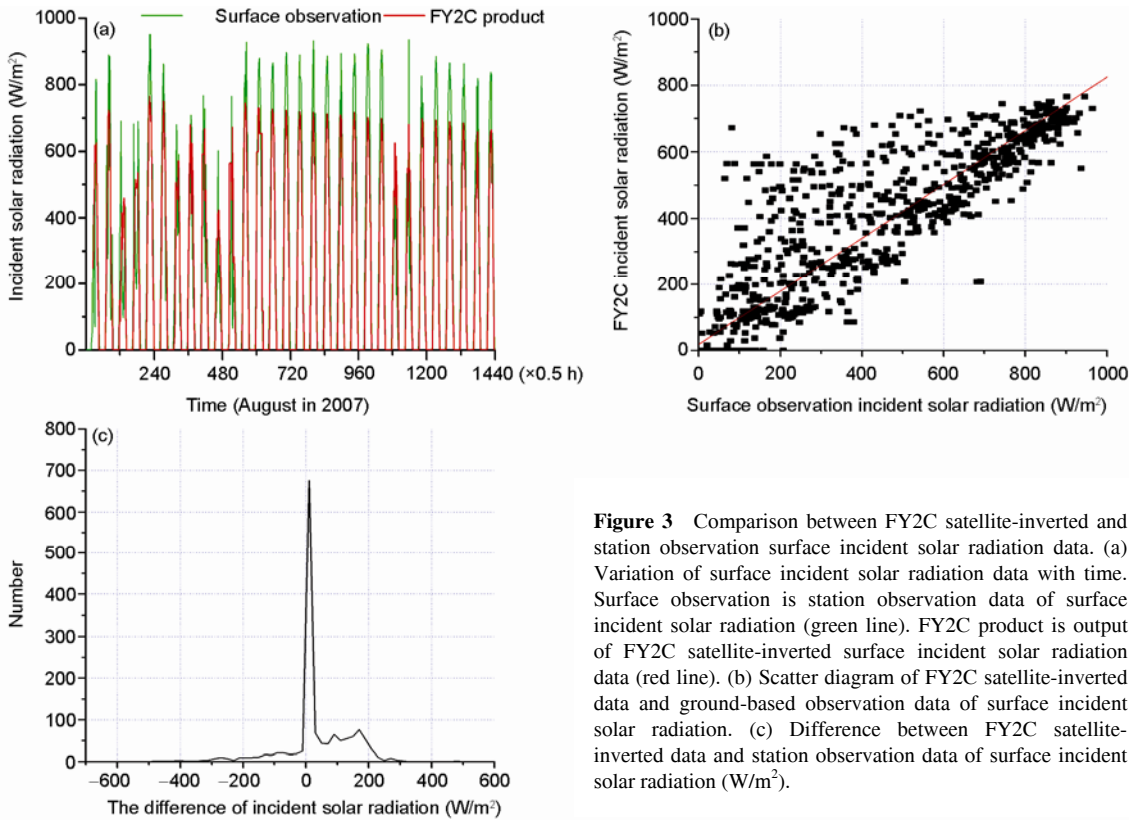

Figure 3 Comparison between FY2C satellite-inverted and station observation surface incident solar radiation data. (a) Variation of surface incident solar radiation data with time. Surface observation is station observation data of surface incident solar radiation (green line). FY2C product is output of FY2C satellite-inverted surface incident solar radiation data (red line). (b) Scatter diagram of FY2C satellite-inverted data and ground-based observation data of surface incident solar radiation. (c) Difference between FY2C satellite-inverted data and station observation data of surface incident solar radiation (W/m^2).

requires the assumption that an observation is non-related to the disturbance of the state variable set, nor requires the inversion of the observation error covariance matrix. This algorithm simplifies calculation.

Below are the traditional EnKF equations [24, 25]:

(i) State error covariance matrix. Matrix defining the state of the set:

$$A = (\psi_1, \psi_2, \cdots, \psi_N) \in R^{n\times N}, \tag{2}$$

where ψ_i (i=1, ⋯, N) is the member of the sample set, n is dimensionality of the state variable, N is the number of samples in the set.

Matrix defining the disturbance of the set:

$$A' = A - \overline{A} = A(I - 1_N), \tag{3}$$

wherein $\overline{A} = A1_N$.

State error covariance matrix:

$$P^f = \frac{A'A'^T}{N-1}. \tag{4}$$

Each element in matrix 1_N is $1/N$.

(ii) Observation error covariance matrix. Given an observation $d \in R^m$, observation vectors of N disturbances are defined as follows:

$$d_j = d + \varepsilon_j, j = 1, \cdots, N, \tag{5}$$

$$D = (d_1, d_2, \cdots, d_N) \in R^{m\times N}, \tag{6}$$

$$E = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_N) \in R^{m\times N}, \tag{7}$$

$$R = \frac{EE^T}{N-1}, \tag{8}$$

where m is dimensionality of the observation vectors.

(iii) Analysis equation.

$$A^a = A + P^f H^T (HP^f H^T + R)^{-1}(D - HA), \tag{9}$$

where H is the operator (H may be non-linear, in which case HA=$H(A)$).

Based on the update of the covariance matrix of the traditional Kalman filter analysis equation, a square root algorithm was used to calculate the update of disturbance in the state variables of the set:

$$P^a = P^f - P^f H^{\mathrm{T}}(HP^f H^{\mathrm{T}} + R)^{-1} HP^f, \tag{10}$$

where H is the observation operator, P and R are the state error covariance matrix and observation error covariance matrix, respectively. The superscripts a, f and T represent analysis field, forecast field, and matrix transposition, respectively. Suppose A and A' are the state matrix and disturbance matrix of the set, respectively. The state variable in this paper is soil volumetric moisture. Therefore, $A = (\theta_1, \theta_2, \cdots, \theta_N) \in R^{n\times N}$. In addition, the observation vector $d \in R^m$ is set as the satellite-inverted soil moisture.

(C)1994-2019 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

E is the observation disturbance set. Matrices are defined as $S=HA'$ and $C=SS^{\mathrm{T}}+(N-1)R$. The algorithm was performed through the following steps:

1) Calculate matrix C and decompose the characteristic values of C, i.e., $Z\Lambda Z^{\mathrm{T}}=C$;

2) Update the mean value of the state variable set $\overline{\theta^a}=\overline{\theta^f}+A'S^{\mathrm{T}}Z\Lambda^{-1}Z^{\mathrm{T}}(d-H\overline{\theta^f})$;

3) Calculate Matrix $X_2=\Lambda^{-1/2}Z^{\mathrm{T}}S$;

4) Perform SVD decomposition $U_2\Sigma_2V_2^{\mathrm{T}}=X_2$;

5) Solve the disturbance of the analytic set of state variables $A^{a'}=A'V_2\sqrt{I-\Sigma_2^{\mathrm{T}}\Sigma_2}\Theta$, Θ is any orthogonal matrix) and then add the mean value of the state variable set $\overline{\theta^a}$ calculated with eq. (2) to derive the analysis field A^a of the state variables.

During the establishment of CLSMDAS, we first used a simple soil water module to establish a land near-surface soil moisture assimilation model based on the EnKF method, constructed a set of ideal near-surface observation data, and then performed an ideal simulation verification test to prove the accuracy and feasibility of the EnKF assimilation module because only near-surface observation was available [21]. On the basis of this analysis, an EnKF assimilation module was coupled with the CLM3.0 model to establish CLSMDAS.

1.4 AMSR-E soil moisture output

The advanced microwave scanning radiometer (AMSR-E) carried by EOS/Aqua is the world's first sensor capable of providing soil moisture service outputs at the global scale. These outputs have been applied widely in hydrological, meteorological, and climatic studies. The working frequencies of AMSR-E are 6.925, 10.65, 18.7, 23.8, 36.5, and 89 GHz.

The AMSR-E land parameters inversion algorithm is based on the radiation transfer model. Three major geophysical parameters are obtained through inversion: soil moisture m_e, vegetation moisture content w_e, and surface temperature T_e. According to the relationship between the observed brightness temperatures and the geophysical variables related to the atmosphere, the model equation can be simplified to:

$$T_{Bi}=\Phi_i(x), \tag{11}$$

where $x=\{x_j\}$, x_j are geophysical variables i.e., soil moisture, vegetation moisture content, and surface temperature; T_{Bi} is the brightness temperature observed through channel i; $\Phi_i(x)$ are standards for the functional relationship between parameters and the brightness temperature.

AMSR-E brightness temperatures (T_{bs}) were first subjected to projection treatment. Projected brightness T_{bs} is classified to make ensure it matches pixel points of the inversion conditions. It is matched with auxiliary data. The soil moisture inversion process comprises: 1) Quality control of input data; 2) project re-sampling; 3) surface classification; 4) eliminating data that do not meet inversion requirements; 5) inversion; and 6) obtaining soil moisture [26].

We downloaded and processed AMSR-E daily output data between 2004 and 2007 and used regular nationwide soil moisture observation data, detailed soil moisture observation data from Inner Mongolia and Henan Province, and AMSR-E inverted soil moisture data for a comparative analysis.

Comparative analysis results showed that: 1) Both spatial and temporal variation in AMSR-E inverted soil moisture data was small; 2) the difference between AMSR-E inverted soil moisture and ground-based observed soil moisture was related significantly to level of soil moisture. The inverted soil moisture is more accurate in arid and semi-arid regions than in wet regions. This coincides with the theoretical analysis results; and 3) the difference between AMSR-E inverted soil moisture and ground-level data observed at Station 117 in Inner Mongolia was generally smaller than the differences at Station 115 in Henan. This is because observation points of Station 117 in Inner Mongolia were located mostly in grasslands and observation points in Henan were located mostly in farmlands. When there is high vegetation coverage, the soil moisture inversion capacity of microwaves is reduced. See ref. [27] for details.

2 China Land Soil Moisture Data Assimilation experiment and results verification

Single-point ecological station observations capture key atmospheric driving observation data, and surface and soil data. After CLSMDAS was established, first we used single-point observation data acquired from ecological stations for the land soil moisture data assimilation experiment and analyzed the performance of this assimilation system. After the single-point soil moisture assimilation experiment, we prepared Chinese atmospheric driving field data with temporal resolution of 1 hour and spatial resolution of 0.25°×0.25° and AMSR-E inverted soil moisture data. We then performed the China soil moisture assimilation experiment. In this section, the processes of the single-point and regional soil moisture assimilation experiments are described and the results analyzed.

2.1 Single-point land soil moisture data assimilation experiment

Shouxian National Climatic Observation Station was located by Huaihe River in the north of Anhui Province at 116°47′E, 32°33′N. It has an average elevation of 23.5 m ASL. It is 25000 m^2 in extent and belongs to Huanghuai Agricultural Ecological Observation Area. Data from Shouxian Observation Station were acquired between April

(C)1994-2019 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

and June 2004. The temporal resolution of the observation data is 30 min. The atmospheric driving data include: atmospheric temperature, relative humidity at 2 m altitude, wind speed, atmospheric pressure and precipitation; surface short-wave radiation and 10 cm soil moisture data observed at ground level. The CLSMDAS established in this paper was used to carry out a soil moisture data assimilation sensitivity experiment. In CLSMDAS, the state error covariance matrix formed a state variable set by randomly disturbing the state variables. Then eq. (4) was used to calculate the state error covariance matrix. The observation error covariance matrix was calculated with eq. (8). The following four soil moisture assimilation experiments were designed (Table 1).

Experiment 1: Alternate the model background error and observation error to assess effects of on assimilation results. The model background error was set at 0.03. In other words $\theta_i=\theta_b(1+\gamma)$ was used to generate a sample set. γ was a random value within [−0.03, 0.03]. Figure 4(a) shows time-variation curves of modeled, observed, and assimilated soil moisture when the observation error was set at 0.01 and 0.03, respectively (empirical value is used in this paper). The assimilated soil moisture is closer to the observation data when the observation error is set at 0.01.

Experiment 2: Change the size of the EnKF sample set and assess the effects of samples in the set on assimilation results. Figure 4(b) shows time-variation curves of modeled, observed, and assimilated soil moisture data when the number of samples in the set was set at 10, 20, and 50. The more samples the set contained, the closer the assimilation result became to observed data. Differences between the assimilation results of 10, 20, and 50 samples were not very significant.

Experiment 3: Assess the effects of soil moisture depth on the assimilation results. Because the first layer of ground-observed soil moisture is the mean value of a 10 cm-deep soil layer which covers the first four soil moisture layers in CLM3.0, we designed two experiments. The observed soil moisture of the first layer was assimilated with the model first-layer soil moisture. The first-layer observed soil moisture (10 cm) was then divided into four layers and then assimilated with soil moisture data of the first four layers of the model for comparison. Figure 4(c) shows time-variation curves of the modeled, observed, and assimilated soil moisture. There were major differences between the results of both assimilation schemes. The more the observed soil moisture information, the closer to the observed results the assimilated soil moisture was.

Experiment 4: Assess the effects of soil moisture frequency on assimilation results. The frequency of the observation data was set at once an hour, once a day or every 3 days. Figure 4(d) shows time-variation curves of the modeled, observed, and assimilated soil moisture data. The higher the frequency of the observation data, the closer the assimilation data became to the observed data. The lower the observation frequency, the closer the assimilation result became to the modeled result. Figure 4(e) and (f) are the time-variation curves of temperature and precipitation, respectively, used during the above-mentioned experiments.

2.2 China Land Soil Moisture Data Assimilation experiment

CLSMDAS was used to perform a number of experiments. The spatial extent of the assimilation experiment was 15°–55°N, 75°–135°E. CLM3.0 was used as the land surface model. FY2C satellite precipitation estimate data and FY2C ground-incident solar radiation data described in 1.2 were used as the precipitation and radiation in the atmospheric driving data. Temperature, moisture, atmospheric pressure, and wind speed data were derived through interpolation of NCEP reanalysis data. The observation data were AMSR-E inverted soil moisture data downloaded from NASA's official website (http://nsidc.org/data/amsre/). EnKF assimilation was used as the assimilation method. The spatial resolution of soil moisture data was 0.25°×0.25° after assimilation. There were 10 layers in the vertical direction. The classification of soil moisture layers was consistent with that of the CLM3.0 model. The temporal integral step of the model was 30 min. The experiment in 2.1 showed that the sum of the soil moisture values of the topmost four layers in CLSMDAS was consistent in the physical significance of the 10 cm soil moisture data observed at ground level. Therefore, the output soil moisture data of the topmost four layers were combined and then processed into grid-point daily mean and monthly mean soil moisture data for further analysis.

Analysis of the assimilation experiment between June and September 2006 shows that the CLSMDAS soil moisture distribution is consistent with soil moisture observed on the ground. Due to insufficient available ground-level observation soil moisture data, the large error in observed soil moisture, and the inherent high variability of the spatial distribution of soil moisture, it is difficult to verify the assimilated soil moisture data. The Drought and Flood Climate Bulletin is an authoritative output issued regularly by

Table 1 Four soil moisture assimilation experiment schemes

Experiment	Background error (mm^3/mm^3)	Observation error (mm^3/mm^3)	Number of samples	Observation frequency (h)	Number of assimilation layers
1	0.03	0.01/0.03	30	0.5	1
2	0.03	0.01	10/20/50	0.5	1
3	0.03	0.01	30	0.5	1/4
4	0.03	0.01	30	0.5/24/72	1

(C)1994-2019 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

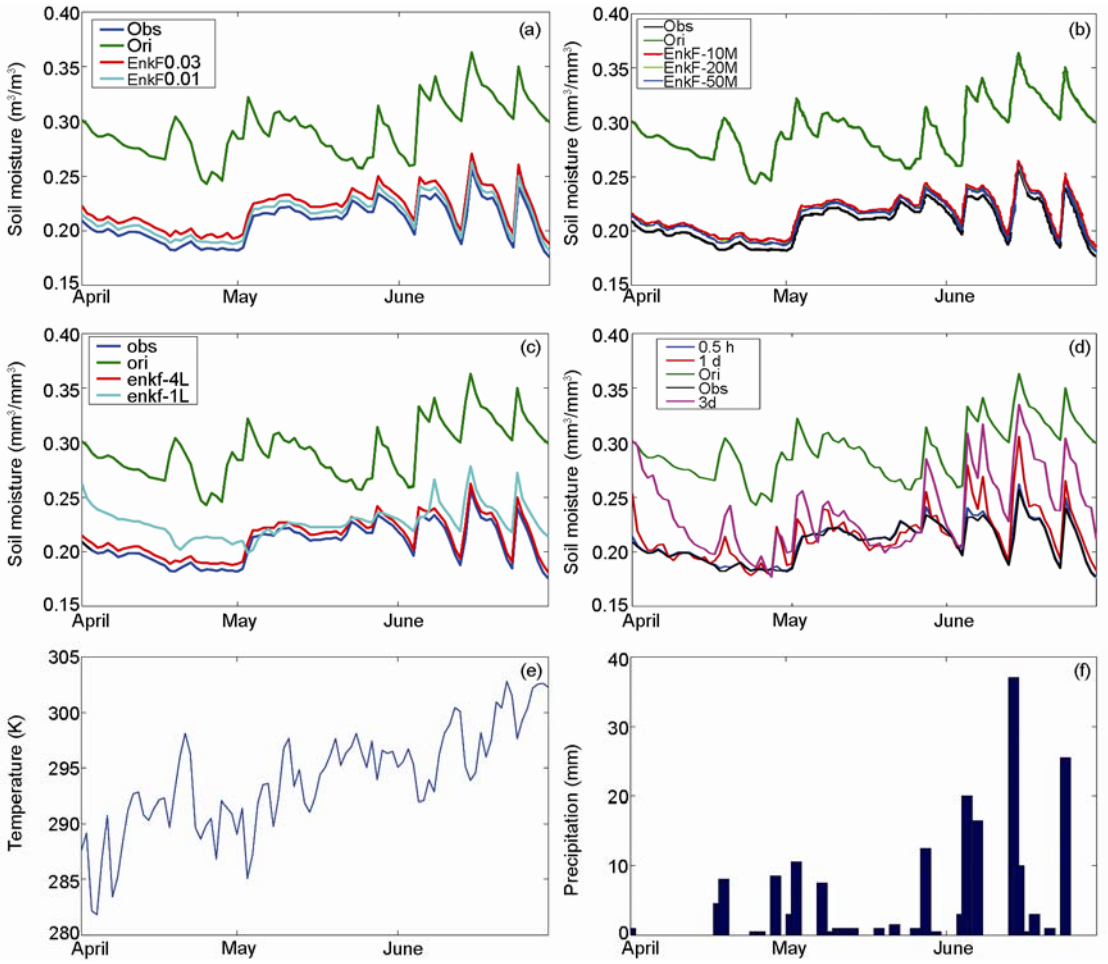

Figure 4 Results of assimilation experiment at Shouxian in April–June, 2004.

the National Climate Center. The climatic drought and flood distribution maps of this are drawn based on comprehensive analysis of station observation data observed. In this paper, the distribution map of ground soil moisture observation stations and the climatic drought and flood distribution map published by the National Climate Center were used to analyze the CLSMDAS soil moisture data (Figures 5 and 6).

Figure 5(a) shows the CLSMDAS soil moisture on July 18, 2006 and Figure 5(b) shows Chinese 10-cm deep soil moisture observations on July 18, 2006 (Figure 5(b)). According to the CLSMDAS soil moisture distribution map, soil moisture is relatively high in most southern regions. In Figure 5(b), there are few observation stations in the south, but all of these have high soil moisture. In the low soil moisture regions in eastern and central of Inner Mongolia, the observed and assimilated soil moisture were consistent. In the lower-soil-moisture regions in central China, the observed and assimilated soil moisture data were also quite consistent.

Figure 6 compares the Chinese assimilated soil moisture distribution and a Chinese drought and flood monitoring map published by the National Climate Center [27]. According to the China Drought and Flood Climate Bulletin published by the National Climate Center, the most serious summer drought since 1949 took place in Sichuan Province and Chongqing in August 2006. National climatic drought monitoring results from August 18 showed that the west of Chongqing and the east of Sichuan still remained in severe to extremely severe drought. The great majority of Sichuan, eastern Tibet, southwest Hubei, northwest Hunan, northern Guizhou, northern Xinjiang and southern Gansu and eastern Inner Mongolia remained in medium to severe drought. The east of Jiangnan, the great majority of northern China, and the east of northwest China had mild to medium droughts (http://climat.cma.gov.cn/). The CLSMDAS soil moisture distribution on August 18, 2006 showed significantly lower soil moisture values in drought regions of Chongqing and Sichuan than in surrounding regions. According to the China Drought and Flood Climate Bulletin issued by the National Climate Center, in September 2006 droughts persisted or developed in northern China, southern Chongqing, northern Guizhou and northern Xinjiang, and eastern Inner Mongolia. Different degrees of drought occurred in eastern Hubei and southern Guangxi. In the CLSMDAS monthly

(C)1994-2019 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

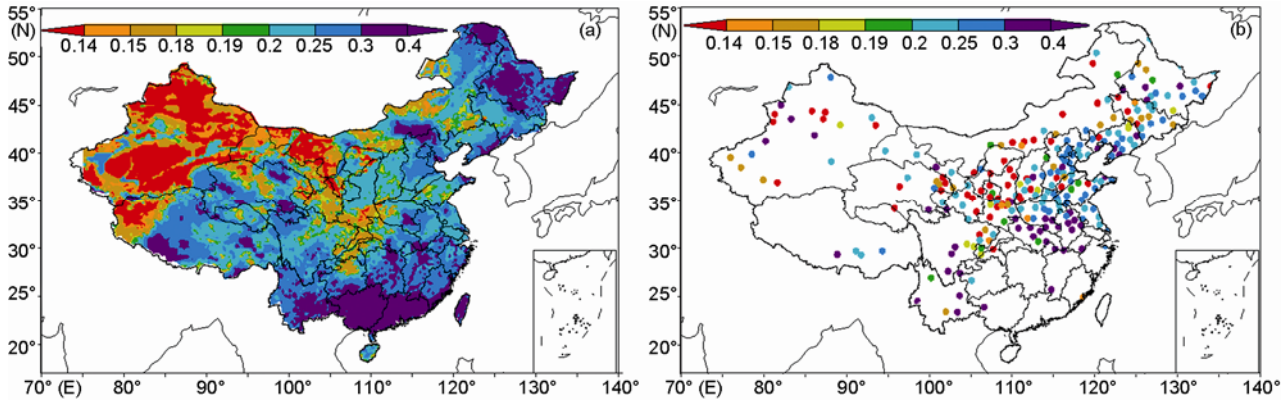

Figure 5 Assimilated soil moisture distribution in China (a) and station observation soil moisture distribution (b). Unit in m^3/m^3.

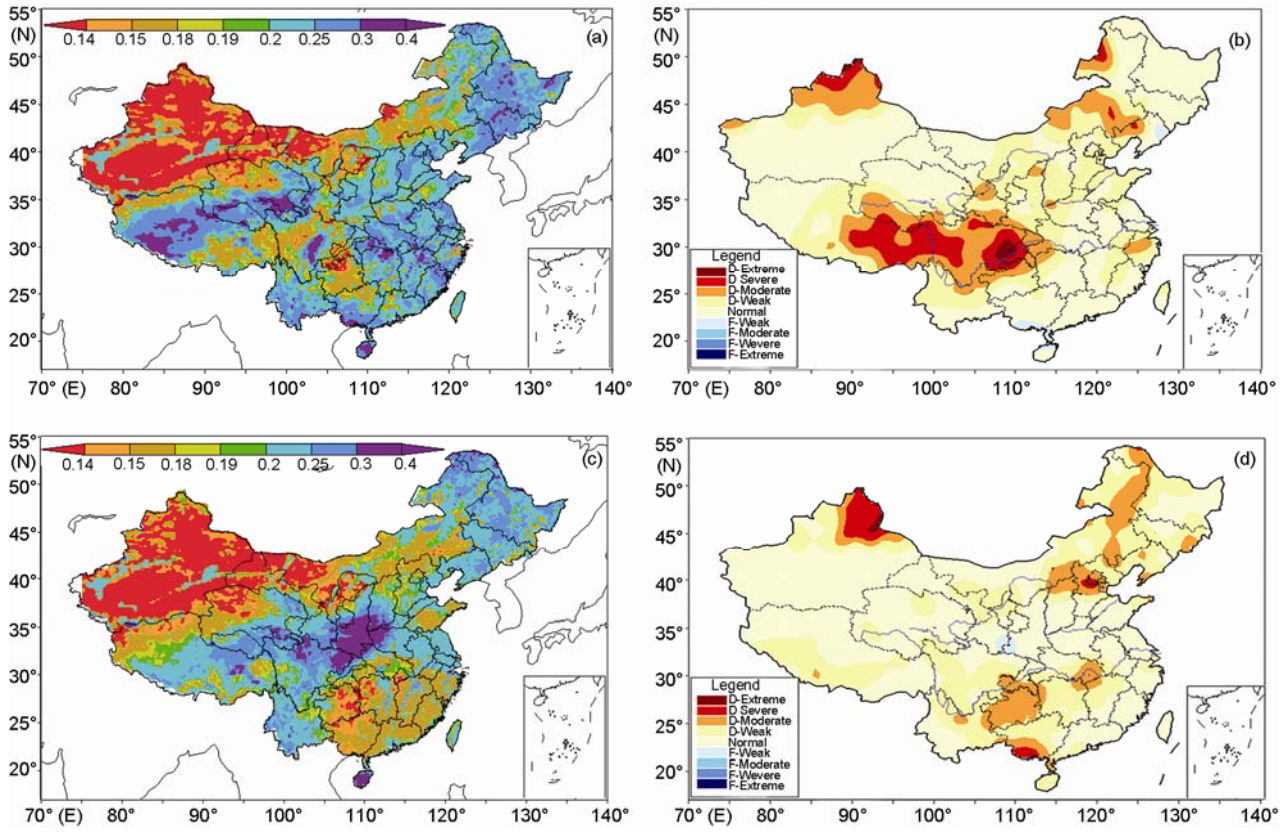

Figure 6 A comparison between Chinese assimilated soil moisture distribution ((a) August 2006, (c) September 2006; in m^3/m^3) and Chinese drought and flood monitoring map ((b) August 2006, (d) September 2006).

mean soil moisture map of September 28, 2006, the low value centers of soil moisture are consistent with drought regions reported in the China Drought and Flood Climate Bulletin. It should be noted that the comparison between the Chinese assimilated soil moisture distribution map of China and the national climatic drought and flood distribution map published by National Climate Center is only qualitative because these two maps do not represent the same phenomenon. The climatic drought and flood distribution map contains relative information integrating many conditions and climatic states. Although soil moisture is a physical quantity that most directly reflects droughts and floods, it does not appropriately reflect the drought and flood characteristics of a certain region unless compared with climate data. For example, although soil moisture is very low in the great majority of southern Xinjiang, it is not defined as an arid region in the climatic drought and flood distribution map. Soil moisture is relatively high in normal years in Sichuan and Chongqing. Therefore, when it decreases significantly, it will be considered a sign of severe drought.

(C)1994-2019 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

3 Summary and discussion

The establishment of the CLSMDAS, especially the application of high temporal and spatial resolution precipitation data acquired from the Chinese geostationary meteorological satellite FY2C, and ground-incident solar radiation data inverted with FY2C visible channal data in CLSMDAS, characterized the spatial and temporal distribution characteristics of atmospheric forcing variables that drive the operation of the land model, improved the simulation precision of the land model, and improved the soil moisture assimilation results. The assimilated high-quality soil moisture grid-point data serve as important basic information for monitoring climate changes, including droughts.

One set of satellite-inverted soil moisture data was assimilated. The inversion error introduced in the soil moisture inversion was large and thus reduced the precision of the soil moisture assimilation results. In future research we will use the surface microwave radiation transfer model to conduct direct assimilation of satellite microwave channel radiation brightness temperature data to improve the precision of soil moisture assimilation. With regard to processing atmospheric driving data, we are currently using multisource data fusion to interface the numerical model with surface and sounding regular observation data, and data recorded automatically at weather stations. We can then use this system to obtain surface atmospheric temperatures, atmospheric pressure, humidity, and wind speed grid-point data with high precision and high spatial and temporal distribution. By combination with precipitation and ground-incident solar radiation data obtained through geostationary meteorological satellite inversion, we can construct a more reasonable atmospheric driving data set and further improve the precision of land model simulation and assimilation.

This work was supported by National High Technology Research and Development Program of China (Grant Nos. 2007AA12Z144, 2009AA12Z129), Chinese COPES Project (Grant Nos. GYHY200706005, GYHY200806014), China Meteorological Administration New Technology Promotion Project (Grant No. CMATG2008Z04).

1 Houser P R. Remote-sensing soil moisture using four-dimensional data assimilation. Doctoral Dissertation. Arizona: University of Arizona, 1996
2 Houser P R, Shuttleworth W J, Famiglietti J S, et al. Integration of soil moisture remote sensing and hydrologic modeling using data assimilation. Water Resour Res, 1998, 34: 3405–4320
3 Evensen G. Sequential data assimilation with a non-linear geostrophic model using Monte Carlo methods to forecast error statistics. J Geophys Res, 1994, 99: 10143–10162
4 Li X, Toshio K, Cheng G D. A land data assimilation algorithm based on simulated annealing. Adv Earth Sci, 2003, 18: 632–636
5 Huang C L, Li X. A summary of researches of land data assimilation systems. Remote Sens Technol Appl, 2004, 19: 424–430
6 Yang K, Takahiro W, Toshio K, et al. Auto-calibration system developed to assimilate AMSR-E data into a land surface model for estimating soil moisture and the surface energy budget. J Meteorol Soc Japan, 2007, 85A: 229–242
7 Tian X J, Xie Z H, Zhang S L, et. al. A subsurface runoff parameterization with water storage and recharge based on the Boussinesq-Storage equation for a land surface model, Sci China Ser D-Earth Sci, 2006, 49: 622–631
8 Tian X J, Xie Z H. A land surface soil moisture data assimilation framework in consideration of the model subgrid-scale heterogeneity and soil water thawing and freezing. Sci China Ser D-Earth Sci, 2008, 51: 992–1000
9 Tian X J, Xie Z H, Dai A G. An ensemble-based explicit four-dimensional variational assimilation method, J Geophys Res, 2008, 113: D21124
10 Tian X J, Xie Z H, Dai A G. A land surface soil moisture data assimulation system based on the dual-UKF method and the Community Land Model. J Geophys Res, 2008, 113: D14127
11 Tian X J, Xie Z H. An ensemble-based three-dimensional variational assimilation method for land data assimilation. Atmos Oceanic Sci Lett, 2009, 2: 125–129
12 Tian X J, Xie Z H. An explicit four-dimensional variational data assimilation method based on the proper orthogonal decomposition: theoretics and evaluation, Sci China Ser D-Earth Sci, 2009, 52: 279–286
13 Tian X Z, Xie Z H, Dai A G, et al. A dual-pass variational data assimilation framework for estimating soil moisture profiles from AMSR-E microwave brightness temperature. J Geophys Res, 2009, 114: D16102
14 Zhang S L, Xie Z H, Tian X J, et al. A soil moisture assimilation method based on soil water model and station data. Adv Earth Sci, 2006, 21: 1350–1362
15 Zhang S L, Xie Z H, Shi C X, et al. Application of collective Kalman filter in soil moisture assimilation. Atmos Sci, 2008, 32: 1419–1430
16 Jia B H, Xie Z H, Tian X J, et al. A soil moisture assimilation scheme based on microwave luminance temperature and collective Kalman filter. Sci China Ser D-Earth Sci, 2009, 52: 1835–1848
17 Oleson K W, Dai Y J, Bonan G, et al. Technical description of the Community Land Model (CLM), NCAR/TN-461+STR. 2004
18 Lu Q F. Influence of land surface boundary conditions (such as LAI) on the regional climate simulation and preparation of land surface remotelysensed parameters for regional climate model (CWRF+CLM). Doctoral Dissertation. Nanjing: Nanjing University of Information Science and Technology, 2006
19 Qian T T, Dai A G, Trenberth K, et al. Simulation of global land surface conditions from 1948 to 2004, part I: Forcing data and evaluations. J Hydrometeorol, 2006, 7: 953–975
20 Shi C X, Xie Z H. A time downscaling scheme of precipitation by using geostationary meteorological satellite data. Adv Earth Sci, 2008, 27: 15–22
21 Shi C X. A study of satellite soil moisture data assimilation based on EnKF algorithm. Doctoral Dissertation. Beijing: Graduate University of Chinese Academy of Sciences, 2008
22 Stamnes K, Tsay S C, Wiscombe W, et al. Numerically stable algorithm for discrete ordinate method radiative transfer in multiple scattering and emitting layered media. Appl Optics, 1988, 27: 502–509
23 Stuhlmann R, Paschke E, An improvement of the IGMK model to derive total and diffuse solar radiation at the surface from satellite data. J Appl Meteorol, 1990, 29: 596–603
24 Evensen G. The Ensemble Kalman Filter: Theoretical formulation and practical implementation. Ocean Dynamics, 2003, 53: 343–367
25 Evensen G. Sampling strategies and square root analysis schemes for the EnKF. Ocean Dynamics, 2004, 54: 539–560
26 Njoku E G, Jackson T J, Lakshmi V, et al. Soil moisture retrieval from AMSR-E. IEEE Trans Geosci Remote Sensing, 2003, 41: 215–229
27 Zhu C H, Zhang Q. A brief introduction of the drought and flood climate monitoring service system of China. Meteorol Sci Technol, 1996, 2: 33–35

(C)1994-2019 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net