

Evaluation of ORCHIDEE-MICT simulated soil moisture over China and impacts of different atmospheric forcing data

Zun Yin¹, Catherine Ottlé¹, Philippe Ciais¹, Matthieu Guimberteau^{1,2}, Xuhui Wang^{1,3,4}, Dan Zhu¹, Fabienne Maignan¹, Shushi Peng³, Shilong Piao³, Jan Polcher⁴, Feng Zhou³, Hyungjun Kim⁵, and other China-Trend-Stream project members*

¹Laboratoire des Sciences du Climat et de l'Environnement, CNRS-CEA-UVSQ, Gif-sur-Yvette 91191, France

²UMR 7619 METIS, Sorbonne Universités, UPMC, CNRS, EPHE, 4 place Jussieu, Paris 75005, France

³Sino-French Institute for Earth System Science, College of Urban and Environmental Sciences, Peking University, Beijing 100871, China

⁴Laboratoire de Météorologie Dynamique, UPMC/CNRS, IPSL, Paris 75005, France

⁵Institute of Industrial Science, The University of Tokyo, Tokyo, Japan

*A full list of the China-Trend-Stream project members and their affiliations appears at the end of the paper.

Correspondence to: vyin@lsce.ipsl.fr

Abstract. Soil moisture is a key variable of land surface hydrology and its correct representation in land surface models is crucial for local to global climate predictions. The errors may come from the model itself (structure and parameterization) but also from the meteorological forcing used. In order to separate the two source of errors, four atmospheric forcing datasets: GSWP3 (Global Soil Wetness Project Phase 3), PGF (Princeton Global meteorological Forcing), CRU-NCEP (Climatic Research Unit-National Center for Environmental Prediction), and WFDEI (WATCH Forcing Data methodology applied to ERA-Interim reanalysis data), were used to drive simulations in China by the land surface model ORCHIDEE-MICT. Simulated soil moisture was compared with in-situ and satellite datasets at different spatial and temporal scales in order to: 1) estimate the ability of ORCHIDEE-MICT (ORganizing Carbon and Hydrology in Dynamic EcosystEms: aMeliorated Interactions between Carbon and Temperature) to represent soil moisture dynamics in China; 2) demonstrate the most suitable forcing dataset for further hydrological studies in Yangtze and Yellow river basins; 3) understand the discrepancies of simulated soil moisture among simulations. Results showed that ORCHIDEE-MICT can simulate reasonable soil moisture dynamics in China, but the quality varies with forcing data. Simulated soil moisture driven by GSWP3 and WFDEI shows the best performance according to RMSE and correlation coefficient respectively, suggesting that both GSWP3 and WFDEI are good choices for further hydrological studies in the two catchments. The mismatch between simulated and observed soil moisture is mainly explained by the bias of magnitude, suggesting that the parameterization in ORCHIDEE-MICT should be revised for further simulations in China. Underestimated soil moisture in the North China Plain demonstrates possible significant impacts of human activities like irrigation on soil moisture variation, which was not considered in our simulations. Finally, the discrepancies of meteorological variables and simulated soil moisture among the four simulations are analyzed. The result shows that the discrepancy of soil moisture is mainly explained by differences in precipitation frequency and air humidity rather than differences in precipitation amount.

1 Introduction

Climate change strongly influences the hydrological cycle, which in turn affects ecosystems services, food security, and water resources (Bonan, 2008; Piao et al., 2010; Seneviratne et al., 2010; Zhu et al., 2016). **More importantly, the main mechanisms governing hydrological process vary across climate regimes under anthropogenic factors (Guimberteau et al., 2012; Wada et al., 2016, 2017). Covering different climate zones and most types of human activities (An et al., 2017; Basheer and Elagib, 2018; Bouwer et al., 2009; Feng et al., 2016; Rogers et al., 2016; Wu et al., 2018), China is a good test bed to investigate the hydrological complexity of climate-water-human interactions.** In China, annual precipitation increased in the South but declined in the North over the last several decades (Ye et al., 2013; Zhai et al., 2005). This dipole of precipitation trends is partly reflected in the discharge trends of Yangtze and Yellow rivers (Piao et al., 2010), but other factors than precipitation changes affect river discharge including changes in rainfall intensity, land surface state or condition, and water management (Ayalew et al., 2014; Grillakis et al., 2016; Williams et al., 2015). A prerequisite to understand how precipitation changes **transfer** into river discharge changes is to analyze and evaluate the various components of the surface water budget and especially the key variable relationships between precipitation and soil moisture (SM), result of the partition of precipitation **among** evapotranspiration, infiltration, and runoff.

SM indeed plays a crucial role in adjusting local climate (Seneviratne et al., 2013; Teuling et al., 2010), regulating productivity, and ecosystem dynamics (Schymanski et al., 2008; Yin et al., 2014) and affecting carbon budgets (Calvet et al., 2004). SM controls vegetation photosynthesis through transpiration, which in turn significantly influences surface temperature (Bonan, 2008; Dai et al., 2004). It also impacts the infiltration rate of precipitation in the soil and its state before rainfall events determines the ratio of surface runoff to precipitation (Grillakis et al., 2016). Therefore SM is not only of importance in understanding land surface processes, but also is a key indicator for predicting and addressing extreme events, such as heat waves, floods, and droughts (Hirschi et al., 2011; Teuling et al., 2010; Wanders et al., 2014).

To investigate the spatial and temporal SM variations, in-situ measurements (Dorigo et al., 2011; Liu et al., 2001; Piao et al., 2009; Robock et al., 2000) are too sparse and not always representative of larger scales. Although they can provide first hand records of SM fluctuations, the density of in-situ networks cannot meet the requirement for continental scale studies. And the different measurement techniques make it difficult to combine different datasets. Satellite-based SM products (Dorigo et al., 2015; Njoku et al., 2003; Su et al., 2003; Wagner et al., 2012) provide excellent spatial coverage and temporal sampling, but their accuracy varies between instruments and retrieval algorithms used (Liu et al., 2012). Moreover, these estimations concern only the first centimeters of soil but root zone SM cannot be directly assessed, unless a model simulating the water transfer processes is used. **To overcome the uneven coverage of raw data, data assimilation is widely applied to analyze SM from in-situ or satellite observations (Draper et al., 2012; Martens et al., 2016; Reichle et al., 2007). Analyzed products help us understanding SM variation and its relation to climate (Liu et al., 2015b, 2017; Taylor et al., 2012). However, to capture changes of hydrological mechanisms for future projection, such measurements are not enough.**

Land surface models (LSMs) are able to simulate the short- and long-term SM dynamics consistently with atmospheric forcing and surface information (Pierdicca et al., 2015; Rebel et al., 2012; Xia et al., 2014) by reproducing physical processes, and interactions with other climatic, hydrological and ecological factors (Seneviratne et al., 2010). The uncertainty of simulated SM depends on the accuracy of atmospheric forcing, in particular precipitation frequency and intensity, and radiation. However
5 LSMs complexity is a source of structural errors (missing processes) and biased parameters. Thus it is necessary to validate simulated SM by observations in order to diagnose the source of errors and estimate the ability of the chosen LSM to simulate SM dynamics in the area of interest.

In this study, the land surface model: ORCHIDEE-MICT (ORganizing Carbon and Hydrology in Dynamic EcosystEms: aMeliorated Interactions between Carbon and Temperature; Guimberteau et al. (2018)) is used to simulate SM over China. Be-
10 sides land surface hydrology, ORCHIDEE-MICT simulates energy budgets and vegetation dynamics (mechanistic phenology, photosynthesis, and ecosystem carbon cycling), which interact with the water cycle and climate (Guimberteau et al., 2012). Moreover, the evaluation of simulated SM controlled by natural processes is useful to identify human effects (e.g., crops, irrigation, dam operation, etc) on water budget in regions where there is a large misfit between model and observation.

Four global atmospheric forcing datasets are chosen to drive the simulations in China, including GSWP3 (Global Soil
15 Wetness Project Phase 3), PGF (Princeton Global meteorological Forcing), CRU-NCEP (Climatic Research Unit-National Center for Environmental Prediction), and WFDEI (WATCH Forcing Data methodology applied to ERA-Interim reanalysis data), due to their widely applications in numerous hydrological studies (Getirana et al., 2014; Guimberteau et al., 2014, 2017, 2018; Hirschi et al., 2014; Van Den Hurk et al., 2016; Polcher et al., 2016; Schmied et al., 2016; Tangdamrongsub et al., 2018; Yang et al., 2015; Zhao et al., 2017; Zhou et al., 2018). Although they provide gridded surface climate variables at global scale,
20 their uncertainties of representing regional climate are not clear. Through comparison of simulated SM to various datasets, our study also addresses which forcing has the best performance in SM simulation in China.

Our SM simulations are evaluated with different SM datasets including in-situ, remote sensing measurements, and reanalysis. In-situ measurements including ISMN (International Soil Moisture Network; Dorigo et al. (2011)) and PKU (in-situ SM from
25 Peking University; Piao et al. (2009); Xu (2014)) are used to evaluate temporal validation of simulated SM. To evaluate spatio-temporal variations of simulated SM, the satellite based dataset ESA CCI SM (European Space Agency Climate Change Initiative Soil Moisture; Wagner et al. (2012)) is applied in the comparison. Note that both in-situ and satellite SM datasets represent the ‘truth’ to some extent. This implies that real-world SM is influenced by processes that are not modeled such as irrigation and wetlands. Thus mismatches between measured and simulated SM may exist in some regions strongly affected by anthropogenic factors. Moreover, satellite instruments do not measure directly SM which is derived via a complex modelization
30 of the radiative transfer at the soil-vegetation interface calibrated with in-situ data.

Finally the GLEAM SM data (The Global Land Evaporation Amsterdam Model; Martens et al. (2017)) is compared to the simulated SM. Different from other SM datasets, GLEAM SM results from a land surface model constrained by a number of satellite and in-situ observations. This reanalysis product was shown to reproduce reasonable long period SM dynamics at global scale (Martens et al., 2017), which is valuable to evaluate ORCHIDEE-MICT simulations for both surface and root-zone

SM. Furthermore, GLEAM assimilates CCI SM data, so that evaluation of our model against root-zone SM from GLEAM is consistent with evaluation against surface SM from CCI.

Through the simulations and comparisons, three questions will be addressed:

- 5 – Is the model able to provide a reasonable estimation of SM dynamics in China, as a prerequisite for further hydrological studies?
- Which atmospheric forcing gives the best SM simulation according to the comparisons with available observations?
- Which meteorological variable drives the differences of SM among the simulations?

The study area, atmospheric forcing, and SM datasets used in this study are described in Section 2. Section 3 presents the model experiments. Evaluation of simulated SM and discussion are given in Section 4 and 5 respectively.

10 2 Study area, forcing, and evaluation datasets

2.1 Study area

China has multiple climate regimes, which makes hydrological situations influenced by different variables in different regions. The land water budget in China is affected by anthropogenic factors, such as irrigation (Puma and Cook, 2010), afforestation (Liu et al., 2015a; Peng et al., 2014), deforestation (Wei et al., 2018), polders (Yan et al., 2016), dams (Deng et al., 2016), and inter-basin water transfer (Li et al., 2015). Two main river basins are of interest: the Yangtze River Basin (YZRB) and the Yellow River Basin (YLRB) (red and magenta contours respectively in Fig. 1), which cover the main regions of industry and agriculture (gray regions in Fig. 1). The Yangtze River originates in the Qinghai-Tibetan Plateau and flows through two wetted traditional agricultural zones: Sichuan Basin and the plain at the downstream of the Yangtze River (Fig. 1). The Yellow River originates in the Qinghai-Tibetan Plateau as well, but it flows through another two agricultural regions (the Loess Plateau and the North China Plain) under semi-arid and semi-humid zones (Kottek et al., 2006). Our simulations cover the main part of China ($[85-124^{\circ}\text{E}] \times [20-44^{\circ}\text{N}]$) including these two watersheds to assess SM dynamics at catchment scale. Note that in the analysis, the specific regions of the two river basins are coarser than the exact basin contours shown in Fig. 1 due to the interpolation of routing files at the resolution of our simulations.

2.2 Atmospheric forcing

25 Four gridded atmospheric forcing datasets are used to force the model over China: GSWP3, PGF, CRU-NCEP, and WFDEI. All input variables needed are the air temperature at 2 m (T_a), rainfall and snowfall rates, atmospheric specific humidity at 2 m (Q_a), surface pressure, downward short/long wave radiation (R_s and R_l), and wind speed (W). The four forcing datasets are combinations of reanalysis and observation data. These datasets, although built by different methods, are not independent from each other since they share some common inputs. Detailed descriptions are listed below and general information is summarized in Table 1. Preprocessing of the datasets for ORCHIDEE-MICT is described in Sect. 3.3.

GSWP3

The GSWP3 v0 (<http://hydro.iis.u-tokyo.ac.jp/GSWP3>); Kim (2017)) provides a 3-hourly climate data at 0.5° resolution from 1901 to 2010. It is based on the 20th Century Reanalysis (20CR; Compo et al. (2011)), which is downscaled from 2° to 0.5° by a spectral nudging technique in a Global Spectral Model (Yoshimura and Kanamitsu, 2008), in order to maintain both
5 low and high frequency signals at high spatio-temporal scale. Single ensemble correction and vertically weighted damping are applied to remove known artifacts in high latitude regions (Hong and Chang, 2012; Yoshimura and Kanamitsu, 2013). Moreover, observation data are used for bias-correction, such as GPCC v6 (Global Precipitation Climatology Centre; Becker et al. (2013)) for precipitation, SRB (Surface Radiation Budget; Stephens et al. (2012)) for radiation and CRU TS v3.21 (Climate Research Unit; Harris et al. (2014)) for temperature.

10 **PGF**

The PGF (<http://hydrology.princeton.edu/data.pgf.php>, latest version released on 13 Jul, 2014) provides 3-hourly data at 1° resolution from 1901 to 2012 (Sheffield et al., 2006). It is constructed by combining the NCEP-NCAR (National Centers for Environmental Prediction-National Center for Atmospheric Research) reanalysis of Kalnay et al. (1996) with several observation datasets. Precipitation is corrected by downscaled CRU TS v3.1, GPCP (Global Precipitation Climatology Project; Adler
15 et al. (2003)), and TRMM (Tropical Rainfall Measuring Mission; Huffman et al. (2007)) data. SRB and CRU TS data are used in the assimilation of radiation and air temperature, respectively. Other variables (e.g., specific humidity, surface air pressure, wind speed) are just spatially downscaled from NCEP-NCAR according to the local elevation.

CRU-NCEP

The CRU-NCEP v6.1 (ftp://nacp.ornl.gov/synthesis/2009/frescati/modeldriver/cru_ncep/analysis/readme.htm) provides 6-hourly
20 0.5° data. It combines the coarse temporal resolution (monthly) CRU TS dataset with the NCEP reanalysis, which has a higher time interval (6-hourly) but is only available at 2.5°. Monthly climate (except for precipitation) is identical to CRU TS, and NCEP is used only to reconstruct the 6-hourly variability within each month after bi-linearly interpolated to 0.5°. For precipitation the original NCEP values are used for temporal linear interpolation in those CRU grid cells (0.5°) covered by the specific NCEP grid cell (2.5°) in each month. CRU-NCEP dataset is available from 1901 to 2015 at global scale and it is updated every
25 year.

WFDEI

The WFDEI forcing (version 31 Jul 2012) is generated by applying the WATCH Forcing Data methodology (<http://www.eu-watch.org>, Weedon et al. (2014)) to the ERA-Interim reanalysis (Dee et al., 2011) providing 3-hourly data at 0.5° from 1979 to 2009. The ERA-Interim blends GCM modeled variables and a suite of observations by a 4D-Var (4-dimensional variable
30 analysis) data assimilation system (Weedon et al., 2014). All variables are bias-corrected using CRU TS. For precipitation, we use a version that has been bias-corrected by GPCC v5 and v6.

2.3 Soil moisture datasets

International Soil Moisture Network ISMN

ISMN is an international cooperative project providing a global gauged SM database (Dorigo et al., 2011). It is based on in-situ measurements from multiple monitoring regional sub-projects. Here only data from the CHINA sub-project is used (Robock et al., 2000) with in-situ volumetric water content (depth of water column over depth of soil in $\text{m}^3.\text{m}^{-3}$) from 40 stations between 1981 and 1999. SM profiles on 11 vertical layers were collected three times per month and (on 8th, 18th, and 28th of each month). The 11 sampled soil layers are: 0-5 cm, 5-10 cm, and then every 10 cm layers until 1 m. Most stations are located in cropland or grasslands, but information about land use types and soil texture of each site are not provided. Moreover, there is no information about management practices affecting SM, such as irrigation or tillage.

In spite of the long length of this dataset, the data **availability and monitoring period among stations vary widely**. Some stations only recorded SM during the growing season, while others have a full year record. Furthermore, the measurements including the 5 deep layers (below 50 cm) are less than those including the top 6 layers. Only stations with more than 15 years of data were selected, which at least cover the same period (1984-1999). To make sure that there is at least half of the data available in the 15-year time series, stations with less than 270 measurement points in the top 6 layers are removed. This leads to selecting a subset of 20 stations, and given the sparseness of data below 50 cm, only SM in the top 6 layers is used for model evaluation.

In-situ SM from Peking University

The SM was measured over 778 stations of agro-meteorological stations over China by the Chinese Meteorological Administration (Xu, 2014) and collected and harmonized by the research team in Peking University (PKU; Piao et al. (2009)). The dataset provides 10-day SM variation during the growing season (mainly between May and September) from 1991 to 2007. It provides SM profiles in 7 soil layers (0-10 cm, 10-20 cm, 20-30 cm, 30-40 cm, 40-50 cm, 50-70 cm, and 70-100 cm), but the bottom four layers often have missing records. This dataset concerns exclusively croplands but there is no explicit information of soil texture and irrigation. Similar to the ISMN, the monitoring durations among gauging stations are different. 203 stations that cover the period of 1992-2006 are chosen.

ESA CCI SM

The ESA CCI SM is a multi-satellite based product (Liu et al., 2011, 2012; Wagner et al., 2012) **and has been validated both at global scale (Dorigo et al., 2015) and in China (Peng et al., 2015; An et al., 2016)**. The daily SM is retrieved from a suite of microwave sensors spreading the period of 1979-2010 with 0.5° resolution. The representative soil layer depth is approximately 0.5-2 cm. Multi sensors ensure a long term records of SM dynamics, however the **uncertainty** of the data varies with the change of available sensors and corresponding algorithms. Moreover, the remote sensing technique limits its ability to detect SM in frozen soils or under snow cover. Therefore SM data is not available during winter in high latitude regions (e.g., Northern

China). The data availability also varies along the period according to the number of available instruments and the increase of their temporal and spatial resolutions. In China, the fraction of days with available records (Figure 4 of Dorigo et al. (2015)) is lower than 20% from 1979 to 2006. More importantly, large spatial variation of gaps exists as well before 2006 (Fig. B1). The period after the launch of METOP-A-ASCAT (Advanced Scatterometer) at the end of 2006 appears much stable. To provide a reliable validation, we only use the CCI SM data between 2007 and 2009.

GLEAM v3.0A SM

The GLEAM v3.0 is a multiple algorithm, observation-based model reconstructing the components of the land evaporation process, including daily SM, evapotranspiration, and interception at 0.25° resolution (Martens et al., 2017). It has three sub versions. Due to the short duration of version B (2003-2015) and C (2011-2015) only version A, which covers the period 1980-2014, is used here. Radiation and air temperature used in GLEAM 3.0A are from ERA-Interim, and precipitation is from MSWEP (Multi-Source Weighted-Ensemble Precipitation; Beck et al. (2016)).

Both surface and root-zone SM from GLEAM, which has been validated by Martens et al. (2017), are used for comparison. The surface SM in the top 0-10 cm is a combination of simulated SM from the GLEAM soil module, SMOS (the Soil Moisture Ocean Salinity satellite mission; Kerr et al. (2001)), and ESA CCI SM (ESA Climate Change Initiative Soil Moisture; Liu et al. (2011, 2012); Wagner et al. (2012)) through the data assimilation system developed by Martens et al. (2016). The Community Noah land surface model SM fields in GLDAS (Global Land Data Assimilation System; Rodell et al. (2004)) was used to estimate the errors of these SM products. Root-zone SM is derived from the GLEAM soil module based on mass balance. GLEAM provides SM in separate land-cover tiles of bare soil (0-10 cm), low vegetation (0-100 cm), and tall vegetation (0-250 cm). These tiles are based on MODIS Vegetation Continuous Fields (MOD44B; Hansen et al. (2003)).

3 Land surface model, simulation protocol, and model-data comparison metrics

3.1 Land surface model

ORCHIDEE (Organizing Carbon and Hydrology In Dynamic Ecosystems; Krinner et al. (2005)) is a physical-based land surface process model. It is mainly composed by two modules. The SECHIBA (surface-vegetation-atmosphere transfer scheme) module calculates the exchange of water and energy between land and atmosphere with a high time interval (half an hour). While the STOMATE (Saclay Toulouse Orsay Model for the Analysis of Terrestrial Ecosystems) module estimates the carbon cycle at daily time scale. The ORCHIDEE-MICT (aMeliorated Interactions between Carbon and Temperature, SVN version 3952; Guimberteau et al. (2018); Zhu et al. (2015)) is a recent version of ORCHIDEE including new processes as the interactions among frozen soil, snow, plants, and soil carbon pools. It accounts for soil freezing, soil carbon discretization, snow processes, and lateral water flows to improve the simulation of the main biogeochemical cycles in permafrost regions. It has been chosen in this study because China has a large permafrost area especially for the Tibetan Plateau, where originate both Yellow and Yangtze rivers. To simulate the SM dynamics, ORCHIDEE-MICT uses a 11-soil layer scheme, whose depth in-

creases exponentially until 2 m. The respective depths (in meters) of the calculation nodes are the following: 0.0005, 0.002, 0.006, 0.014, 0.03, 0.06, 0.12, 0.25, 0.5, 1.0, and 1.75. Each grid cell can include up to three soil tiles: bare soil, trees, and grass/crops, which are filled by the corresponding plant functional types (PFT) of the 13-PFT scheme of ORCHIDEE-MICT to allow better representation of their specific hydrology. The hydrological budget is calculated separately in each soil tile.

- 5 The amplitude of SM depends on soil texture, which is a part of boundary conditions. Explicit description of the ORCHIDEE-MICT model can be found in Guimberteau et al. (2018). ORCHIDEE will be referred to ORCHIDEE-MICT for brevity in the following text.

There are two main outputs of SM in ORCHIDEE. The total SM (θ_t) indicates the total amount of soil water volume in the top 2 m soil layer in a grid cell. The SM profile (θ_p) records the vertical distribution of soil water content in the 11 soil layers.

- 10 Note that the θ_p in each soil layer is an average value among the three soil tiles. The initial unit of ORCHIDEE SM is $\text{m}^3 \cdot \text{m}^{-3}$.

3.2 Simulation protocol

Four simulations were performed driven by different forcing datasets described in Sect 2.2. In the simulations, CO_2 rise, and land use change are taken into account but without human processes like irrigation. The 13-PFT map is from LUH2 (<http://luh.umd.edu>) and the soil texture map is from Zobler (1986). For the 3-soil texture scheme of Zobler86, the minimum residual and maximum saturated SM are 0.065 and $0.43 \text{ m}^3 \cdot \text{m}^{-3}$, respectively. The model domain covers the main part of China ($[85\text{-}124^\circ\text{E}] \times [20\text{-}44^\circ\text{N}]$). The spatial resolution is as same as the atmospheric forcing (Table 1). The simulation period covers 39 years, from 1971 to 2009, except for the one driven by WFDEI, which is from 1979 to 2009. To make sure that carbon (LAI and biomass) and water cycle variables can reach equilibrium, a 100-year spin-up was performed by repeating 10 times the forcing of the period 1971-1980 (for WFDEI, 50 times the period 1979-1980). Starting from the end of the spin-up, simulations were run from 1981 to 2009. The output driven by PGF forcing was re-gridded at $0.5^\circ \times 0.5^\circ$ to match the resolution of other simulation outputs.

- 25 The temporal resolution of forcing datasets is either 3-hourly or 6-hourly (CRU-NCEP), which is larger than the simulation time step of SECHIBA (30 min). To have a reasonable precipitation intensity and thus a good infiltration of water in the soil, the default precipitation splitting algorithm of ORCHIDEE is applied in our simulations. At the beginning of each forcing time step, if precipitation occurred, the precipitation amount (precipitation rate multiplied by the time interval of specific forcing) will be uniformly distributed to the first half of the forcing time step.

3.3 Model-data comparison methodology and metrics

Comparison protocol

As the soil depths, periods, and spatio-temporal resolutions are different in the four SM datasets (Sect. 2.3), we have to chose corresponding ORCHIDEE outputs for each comparison. To compare with the in-situ data of ISMN and PKU, we first extracted modeled daily SM profile (θ_p) from the nearest grid cell for each station. Then the SM above a certain soil depth was chosen (50 cm for ISMN and 20 cm for PKU). PKU SM is provided in degree of saturation, defined as the volume ratio of actual

- 30

water content to its maximum value when the soil is saturated. As the soil porosity is unknown, the PKU SM dataset cannot be directly compared with simulated SM from ORCHIDEE, which is defined from modeled porosity. To overcome this problem, normalization was applied on both datasets before comparison. The normalized data at each station and in the corresponding grid cell of the model is the ratio of the difference between the original value and its mean (during the observation period) to its standard deviation.

According to the sampled depth of the ESA CCI SM, the daily top 4-layer (2.2 cm) averaged SM from ORCHIDEE is used. Regarding the definition of GLEAM SM (Sect. 2.3), we used the daily top 6-layer (approximately 9.2 cm depth) averaged SM and the total SM of ORCHIDEE to compare with GLEAM surface and root-zone SM, respectively. The period length and soil depth of each comparison are shown in Table 2. **In addition, the timing of all SM datasets is uniformed to the Coordinated Universal Time (UTC).**

Metrics

Pearson correlation coefficient (r) is calculated to estimate the correlation between simulated and observed SM. Daily SM corresponding to the measurement date reported in ISMN was collected to calculate r . As there is no date information from the 10-day PKU dataset, we used the 10-day averaged SM from ORCHIDEE for comparison.

The Root Mean Square Error (RMSE) is applied in order to estimate the temporal differences between simulation and observation. The same data pairs are used for RMSE calculation as the correlation coefficient except for PKU due to the normalization. Note that RMSE is related to the magnitude of SM, which varies significantly in China. To make it comparable in space, relative RMSE is calculated by dividing the mean of the simulated and observed SM.

According to Kobayashi and Salam (2000), the mean squared deviation (MSD), which is $RMSE^2$, can be decomposed into squared bias (SB), squared difference between standard deviation (SDSD), and lack of correlation weighted by the standard deviation (LCS), as:

$$MSD = RMSE^2 = SB + SDSD + LCS. \quad (1)$$

SB is the bias between simulations and observations. It is independent from other two components:

$$SB = (\bar{s} - \bar{m})^2, \quad (2)$$

where \bar{s} and \bar{m} are the mean of simulated and measured values, respectively. The SDSD indicates the mismatch of variation magnitude between simulated and observed variables, defined as:

$$SDSD = (SD_s - SD_m)^2, \quad (3)$$

where SD_s and SD_m are standard deviation of simulations and measurements, respectively. High SDSD implies a failure of the model in simulating the degree of fluctuation across the n measurements. Note that SDSD correlates with LCS, which accounts for SD_s and SD_m as well:

$$LCS = 2SD_s SD_m (1 - r), \quad (4)$$

where r is the Pearson correlation coefficient. The LCS is an indicator of the performance of the model to simulate the pattern of fluctuation of the measurements. The lower the LCS is, the better the model performs.

Finally, to evaluate the characteristic time scale of modeled SM response to hydrological processes, lag- k autocorrelation coefficient (R_k) is computed. The R_k is the correlation coefficient of a time series with itself but with a k time step lag, as:

$$5 \quad R_k = \frac{\sum_{i=1}^{n-k} (x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (5)$$

where n ($n > k$) is the length of the specific time series; \bar{x} is the mean value. For SM time series in a specific grid cell, R_k was computed for different k values. The value of R_k decreases with increasing k and the k -lag time series are considered not auto-correlated if R_k is less than a threshold $1/e$ (Maurer et al., 2001; Rebel et al., 2012). The day number when R_k first drops below a threshold of $1/e$ is called number of lag days (NLD). The NLD difference is used to compare the overall characteristic
10 time scales between datasets. The difference of R_k profiles gives additional information on the autocorrelations for lag. The R_k comparison was implemented between GLEAM and ORCHIDEE because other datasets do not have complete daily records.

The linear trend of SM change in the 29 years is of interest as well. The Mann-Kendall test (Kendall, 1975; Mann, 1945) is applied to test if simulations capture observed trends of SM, with p -value < 0.05 indicating a significant trend.

3.4 Correlation of uncertainties between SM and meteorological factors

15 In our simulations, the difference in atmospheric forcing is the only source of difference in simulated SM. **We look at different climate variables to explain SM differences among simulations.** These variables include monthly precipitation amount (P) and the number of precipitation days in one month (N_p) excluding days with $P < 0.01 \text{ mm.d}^{-1}$. Precipitation days are categorized into 5 classes of 0.01-1, 1-5, 5-10, 10-15, and $> 15 \text{ mm.d}^{-1}$. The number of days with precipitation amount in each class was calculated, denoted by N_p^i with $1 \leq i \leq 5$. **Other meteorological variables are incoming short/long wave radiation (R_s/R_l), air
20 temperature (T_a), air humidity (Q_a), and wind speed (W).** Regarding SM, both total SM (θ_t) and SM in each soil layer (θ_p^i , i is the index of soil layer) were correlated with these variables. To estimate the difference of a variable x among the four simulations, the averaged MSD (D_x) is computed as:

$$D_x = \frac{\frac{1}{n} \sum_{i \neq j} \sum_{t=1}^n (x_{t,i} - x_{t,j})^2}{\binom{n}{2}}, \quad (6)$$

where $N = 4$ is the number of simulations; i and j ($1 \leq i, j \leq N$) are indexes of the four simulations; $\binom{n}{2}$ is the binomial
25 coefficient; n is the length of the time series; t is the time step. Note that we use the absolute value of D_x not relative D_x (D_x over averaged value of x in the specific grid cell) for the analysis because the relative D_x cannot reflect the linkage of uncertainty between inputs and outputs. Detailed explanation is shown in Supplement C.

4 Results

4.1 SM evaluation against multiple datasets

Comparison with ISMN and PKU in-situ data

In most cases, the correlations between modeled and measured SM at ISMN stations (see Sect. 2.3) are significantly positive (Fig. 2). High correlations ($r > 0.6$) are found over the Loess Plateau in the semi-arid zone, where is the region of rainfed agriculture and SM is less affected by anthropogenic processes. In the North China Plain water is limited as well, whereas irrigation is widely applied for agriculture leading to low r (below 0.5). To further compare the simulated and measured SM, three ISMN stations (marked by squares in Fig. 2(a)) are chosen to represent for different wet conditions and model-data comparisons are shown in Figure 3. Xifeng locates in the semi-arid zone (MAP = 556 mm.yr⁻¹), where θ_t is low (0.2 m³.m⁻³ on average) with a large inter-annual variation. The variability of simulated θ_t is consistent with observations ($0.73 < r < 0.87$; when CRU-NCEP is excluded) due to lower human impacts on rainfed agriculture in this region (Li et al., 2014). Xinxian is located in the North China Plain with similar MAP (580 mm.yr⁻¹) to Xifeng, but in a traditional irrigation region (Wang et al., 2016). θ_t at Xinxian is underestimated, possibly because irrigation is not included in our simulations. Thus the model cannot capture the seasonal variations of θ_t , given r values ranging between 0.11 and 0.21. Xuzhou is in the North China Plain as well but with a higher MAP (847 mm.yr⁻¹). The fluctuation of simulated and observed θ_t are coherent, leading to r from 0.55 to 0.64. However the magnitude of θ_t is systematically underestimated as well (Fig. 3).

The correlation coefficients of θ_t between simulations and PKU dataset are shown in Figure 2 as circles. Modeled θ_t has a better performance in the Loess Plateau and the North China Plain than other regions, suggesting that ORCHIDEE is able to capture the variations of SM in semi-arid and temperate zones. In comparison to ISMN, r between ORCHIDEE and PKU θ_t is lower. This may be caused by the shallower depth of the PKU data (20 cm) with stronger influence from fast infiltration and transpiration processes than in the ISMN records (1 m). Moreover, the PKU dataset only records θ_t during the growing season, leading to lower r in absence of full seasonal variations. Negative correlations are found in several sites located along river networks. The negative r ($-0.4 < r < -0.2$) coincides with the coupling of wetness anomaly and irrigation: when droughts occur (reflected by low simulated SM), more water will be withdrawn from the river and irrigated on the crop lands (reflected by high observed SM). Thus the negative r found in Fig. 2 reveals that SM dynamics cannot be well understood without considering anthropogenic activities.

According to the r shown in Figure 2, we find that GSWP3 and WFDEI provide better simulated SM than the other two. The main difference is found in the North China Plain, where the r values of GSWP3 and WFDEI are higher. It indicates that simulation can be improved by selecting suitable atmospheric forcing. Nevertheless, the r is still limited by the lack of measurement information (soil texture, irrigation flag, land cover, etc), and anthropogenic processes in ORCHIDEE. The disagreements between simulated and measured SM are caused by the spatial scale as well. The spatial resolutions of forcings (0.5° \approx 55 km) are too coarse to represent the specific climatic conditions of gauging stations. On the other hand, the com-

parison cannot provide a comprehensive validation in the YZRB where few measurements locate in. Thus remote sensing and hybrid SM datasets are required to evaluate the simulations.

Comparison with ESA CCI SM data

Figure 4 (left panel) shows r between CCI and ORCHIDEE θ_s from 2007 to 2009. High r is found in both the North China Plain and the southern China. In southern China, SM is less disturbed by anthropogenic factors due to its wet condition. Thus SM has small variation driven by climate, which can be well simulated by the model. On the other hand, human activities strongly affect SM in the North China Plain, whereas their impacts on r are neutralized by the large annual variation due to seasonality. Weak correlations only exist in the transition zone from the south to the north along the Yangtze river network, where compounds both human disturbances and small annual variation.

Figure B2 and 4 (right panel) show the relative RMSE and the MSD decomposition of θ_s between CCI and ORCHIDEE. Low relative RMSE (< 0.3) is found in the YZRB, but in the YLRB the value is higher (> 0.4). The mean source of MSD is LCS (phase mismatch). It implies that the magnitude of the simulated θ_s is reasonable, but the timing of the fluctuations differs between ORCHIDEE and CCI. The coincidence of magnitude is reflected from the relative difference (Fig. B3), the absolute value of which is less than 0.1 in 76% grid cells excluding the CRU-NCEP case. Large LCS might be due to human activities and the discretization of CCI SM time series. Irrigation in the northern China may significantly affect the fluctuation of θ_s , which leads to underestimation of simulated SM and contributes to the LCS. Simultaneously, due to the incomplete records of CCI SM (Fig. B1) the seasonal variation of SM cannot be fully taken into account in the comparison. The r is consequently declined and the LCS increases (Eq. 4).

The availability and uncertainty of CCI SM vary with space and time (Sect. 2.3 and Fig. B1). To provide reliable estimation, we performed the analysis exclusively in the period 2007-2009. In fact, there is few difference if the comparison covered the whole period 1981-2009. The patterns of r and MSD decompositions (Fig. B4) are similar to that of comparison 2007-2009. The r of 1981-2009 is lower with no doubt, because longer period contains more errors due to the fragmentary records of CCI SM data.

Comparison with GLEAM v3.0A data

The left panel of Figure 5 shows correlation coefficients between GLEAM surface SM (θ_s) and corresponding modeled SM in the surface layer (0-10 cm). Simulated θ_s is significantly correlated with GLEAM (median $r = 0.54$). In the Sichuan Basin, r is lower than its surroundings. According to the spatially averaged r of θ_s , GSWP3 (0.55) and WFDEI (0.66) lead to better performances with ORCHIDEE than PGF (0.43) and CRU-NCEP (0.51). Note that both WFDEI and GLEAM v3.0A used ERA-Interim reanalysis to reconstruct the time series of precipitation, which can explain the higher r when ORCHIDEE is forced by WFDEI. The correlation coefficients of simulated and GLEAM root-zone SM (θ_r , Fig. B5) have the similar patterns as the θ_s , but higher values (median $r = 0.57$) due to the lower variability of θ_r , which smoothes out misfits related to differences in individual rainfall events between ORCHIDEE and GLEAM for θ_s . Compared to CCI θ_s , the r between ORCHIDEE and GLEAM θ_s is much higher. It is probably due to the shallower depth of the CCI θ_s , which is more sensitive to surface processes

and forcing data errors. Moreover, CCI θ_s is a purely satellite product while GLEAM θ_s (v3.0A) is a combination of modeled, in-situ and satellite SM. The latter one is not totally independent of the forcing datasets and therefore more comparable to our simulations.

Figure B6 and the right panel of 5 show the relative RMSE and the MSD decomposition of θ_s between GLEAM and ORCHIDEE. Low relative RMSE (< 0.3) covers most regions except for the North China Plain (> 0.5), where the MSD is dominated by the squared bias values (SB, Fig. 5(b)). This is clearly shown in the relative difference (Fig B7) between GLEAM and ORCHIDEE where simulated θ_s is approximately 30% lower than in GLEAM. Southern China has lower relative RMSE (< 0.2), and MSD is dominated by SB as well. Different from the North China Plain, SB in southern China may be due to the mismatch of land cover and soil parameterization between ORCHIDEE and GLEAM. For instance, the saturated SM in South China is $0.36 \text{ m}^3 \cdot \text{m}^{-3}$ while the maximum SM in GLEAM is $0.45 \text{ m}^3 \cdot \text{m}^{-3}$. A high contribution of LCS to MSD is found in Qinghai-Tibetan Plateau, the upper part of the YZRB and the YLRB, suggesting a mismatch of the phase of SM variability. The MSD is dominated by SDSD in northwestern China ($P < 200 \text{ mm} \cdot \text{yr}^{-1}$) suggesting different magnitudes of SM fluctuations. Nevertheless, the relative RMSE in Qinghai-Tibetan Plateau and northwestern China is as low as in southern China ($< 20\%$). Overall, ORCHIDEE is able to give a reasonable estimation of θ_s in regions where irrigation is not widespread.

Figure 6(a)-(e) shows NLD of ORCHIDEE and GLEAM θ_s computed based on the k -lag autocorrelation coefficient R_k . High NLD implies that θ_s has a longer memory in response to rainfall inputs. However, the spatial distribution of NLD depends not only on rainfall frequency and intensity but also on evapotranspiration and runoff losses after SM recharge by rainfall. The NLD patterns of GLEAM and ORCHIDEE θ_s are similar, which is encouraging in terms of how ORCHIDEE simulates the processes controlling the decrease of SM after each rainfall. Both southern and southeastern China have higher NLD, like in GLEAM. Lower NLD (≈ 20 days) prevail around 30°N in eastern China, whilst the North China Plain has NLD values of 40 days. The main difference of NLD between GLEAM and ORCHIDEE is in Inner Mongolia and over the Loess Plateau, where the ORCHIDEE NLD has values of 20 days, against 40 days in GLEAM. R_k of spatially averaged θ_s in three regions is shown in Figure 6(f)-(h). Overall, R_k of ORCHIDEE θ_s is consistent with that of GLEAM. The GLEAM R_k is close to the ORCHIDEE R_k in the YZRB with difference less than 6 days. In the YLRB, GLEAM R_k is larger than ORCHIDEE R_k , suggesting that modeled θ_s has a faster response to rainfall input. Such bias can be explained by higher simulated evapotranspiration in the YLRB compared to GLEAM (Fig. B8), suggesting that the decline of ORCHIDEE θ_s is faster after rainfall events than in GLEAM and lead to a lower R_k .

The trend of ORCHIDEE θ_s (Fig. B9) is less significant than that of GLEAM θ_s (Fig. B9). In northwestern China, increasing θ_s is found in simulations ($< 0.2 \times 10^{-3} \text{ m}^3 \cdot \text{m}^{-3} \cdot \text{yr}^{-1}$) and GLEAM ($0.2-0.4 \times 10^{-3} \text{ m}^3 \cdot \text{m}^{-3} \cdot \text{yr}^{-1}$). The trend may be due to increasing P (Fig. B10). GLEAM θ_s decreased dramatically in eastern China ($[103-122^\circ\text{E}] \times [20-35^\circ\text{N}]$) while the trends of ORCHIDEE θ_s are not homogeneous in this region. In addition, all forcing datasets show an increasing P in the North China Plain, which leads to slight increase of simulated θ_s . But GLEAM shows decreasing θ_s in most area of the North China Plain. The mismatch of θ_s and P trends suggests that the change of precipitation amount is not the only driver of the trend of SM.

4.2 Comparison of the four forcing datasets

To find the most realistic forcing dataset for SM performance given the ORCHIDEE model, several metrics were calculated and shown in Figure 7. Radar charts show the correlation coefficients (r) and RMSE of simulated SM in comparison to different datasets. Histograms show MSD and its three components. The median of specific metrics is listed in Table 3. GSWP3 has the best performance in estimating the magnitude of SM (lowest MSD) while WFDEI shows the best score in simulating SM variation (highest r). PGF provides as good estimation as GSWP3 in the YZRB, but performs more poorly in capturing SM variation in the YLRB, which is also reflected from the components of MSD. The largest MSD is found in CRU-NCEP in most of comparisons, which is mainly contributed by SB. The SDS and LCS of CRU-NCEP are also larger than others but the differences are not as significant as SB. In addition, we performed the comparison over the full period (1981-2009). Corresponding metrics are shown in Table A1. The values vary slightly, but they do not change our conclusions.

Thus we conclude that both GSWP3 and WFDEI are suitable to simulate SM dynamics in China with ORCHIDEE. The best choice can be made based on the main focus of specific research. For estimating magnitude of SM, GSWP3 is preferable; for investigating SM variation, WFDEI is the best choice. Note that this study only provides the evaluation of SM, but other hydrological components should be compared with observations to confirm the superiority of GSWP3 and WFDEI.

4.3 Source of SM difference among simulations

By investigating the D of meteorological variables and simulated SM among the four simulations (D_x for variable x ; Eq. 6), two questions are addressed: (1) How is D of simulated SM and forcing variables spatially distributed? (2) Can spatial patterns of D of SM be explained by that of meteorological variables? Note that the relative value of D , D over the magnitude of specific variable in each grid cell, is not suitable for the analysis (detailed explanation is in Supplement C).

Figure B11 shows maps of D of θ_t and meteorological variables. As the unit of D depends on specific variables, it can only be used to compare spatial distributions, not values. High D_{θ_t} is found in southwest of China ($[92-104^\circ\text{E}] \times [28-35^\circ\text{N}]$). However, similar patterns do not exist in the D_P (Fig. B11(b)), suggesting that the difference of simulated θ_t is not caused by the difference of precipitation amount of forcing data. Similarly, in southwestern China, no high D is found in meteorological variables except for number of precipitation days (N_p) and air humidity (Q_a), although the patterns of D_{N_p} and D_{Q_a} overlap with D_{θ_t} but extend to zones with low D_{θ_t} as well (Fig. B11(c) and (g)).

To look for clearer links between input and SM D , we decompose N_p and θ_t by scales of P and soil layers respectively (Sect. 3.4). The r of D between simulated SM and meteorological variables are shown in Figure 8. D_{Q_a} , D_{N_p} , and $D_{N_p^2}$ are highly correlated with D_{θ_t} , implying that the difference of simulated θ_t can be explained by the differences of Q_a and N_p among the four forcing datasets. The r between D_{θ_t} and D_P is less than 0.3. All in all, the results suggest that the uncertainty of precipitation frequency and intensity is more important than that of precipitation amount in influencing SM differences among the simulations.

5 Discussion

5.1 Performance of the model to simulate SM

Due to the spatio-temporal complexity of SM and its vertical profile, four datasets were selected to drive the simulations and modeled SM at different depths was validated against multiple datasets. The results showed that ORCHIDEE SM coincides well with CCI (median $r = 0.48$; median RMSE = 0.06) and GLEAM SM (median $r = 0.55$; median RMSE = 0.07) in comparison to other model studies (Lai et al., 2016).

Higher r were systematically found in southern China, the Loess Plateau, and the North China Plain; lower r were found in northwestern China, western Tibetan Plateau, eastern Sichuan basin and downstream of the YZRB. SM is underestimated significantly in the Loess Plateau and the North China Plain with modeled values being 20% and 30% less than in CCI and GLEAM, respectively (Fig. B7 and B3). It is not only due to model parameterization but also due to irrigation activities in those agricultural regions (Fig. 1), which are not considered in the simulations.

Because the in-situ SM measurements were collected only for croplands and grasslands (Piao et al., 2009; Robock et al., 2000), implying potential disturbances from human activities, r was low in the comparison to ISMN and PKU datasets (median $r = 0.37$, Fig. 2). For instance, drought occurred in northern China during 1987-1988 (Yang et al., 2012), which is reflected in the variation of measured SM at Xifeng and Xinxian (Fig. 3(a)-(b)). ORCHIDEE successfully reproduced the drought induced SM decline at the two stations. But SM measured at Xinxian was maintained at a high level. A possible explanation is that the soil at Xinxian was irrigated. Consequently SM at Xinxian did not vary with precipitation leading to a low r (< 0.23). Another possible reason leading to the mismatch between simulations and in-situ measurements is scale effects. Local measurements can be an ideal choice for model validation only if the atmospheric forcing was provided at the same scale due to the spatial variability of precipitation and of landscape. Otherwise remote sensing products derived from multiple observations averaged or aggregated at daily time step are probably more comparable to model simulations obtained with meteorological reanalysis than local in-situ measurements.

In the comparison to CCI and GLEAM SM, low r did not occur in the northern China, such as the Loess Plateau and the North China Plain, but was found in the climatic transition zone between southern and northern China (Fig. 4, B5, and 5). Irrigation may strongly influence SM dynamics in northern China, and in turn reduce r . However, such effect to r is not significant because of the large seasonality of SM in this region. Instead of r , the impacts of irrigation are mainly reflected from the RMSE and relative difference (Fig. B2, B3, B6, and B7). Thus for a region with both irrigation and strong seasonality, bias and RMSE are recommended to trace the footprint of irrigation rather than correlation coefficients. In the climatic transition zone (e.g., Sichuan Basin, mid- and down-stream of the YZRB), climatic seasonality is not as large as the northern China. Meanwhile irrigation is still needed for agriculture, which consequently results in low r between simulated and observed SM.

From the results we conclude that ORCHIDEE provides a satisfactory simulation of SM dynamics in China, except in areas subject to irrigation. This calls for inclusion of irrigation and realistic crop phenology (Wang et al., 2017) as a priority for future application of this model for SM and river discharge dynamics.

5.2 Linkage of discrepancies between meteorological factors and SM through ORCHIDEE

In Section 4.3, we showed that the spatial differences of simulated SM among the four forcing datasets were highly correlated with forcing differences in N_p and Q_a . This suggests that the **uncertainty of precipitation frequency is more critical than that of precipitation amount in determining variation of SM patterns**, as pointed out by other studies, especially in arid and semi-arid regions (Baudena and Provenzale, 2008; Cissé et al., 2016; Piao et al., 2009). To precise the result, we studied the correlation coefficients between the spatial averaged D of SM in different soil layers and of N_p categorized by classes of precipitation intensity (Fig. 8). The result showed that differences in small rainfall events N_p^i with $1 < P < 5 \text{ mm.d}^{-1}$ are more important than other precipitation classes in explaining SM differences due to atmospheric forcing datasets.

Differences in Q_a were also shown to explain a large fraction of the simulated SM differences across different forcings. Q_a determines vapor pressure deficit, which in turn controls transpiration (Farquhar and Sharkey, 1982) and evaporation (Monteith, 1965), suggesting a strong control by atmospheric dryness of the differences in SM found among the four forcing datasets. Both of Q_a and N_p have positive impacts on SM, which enhances the correlation in Figure 8.

Estimating impacts of meteorological factors on SM dynamics is difficult. First of all, the importance of a meteorological variable on SM may vary with climate regimes. For instance, the importance of precipitation and radiation on SM changes from water to energy limited regions. Secondly, impacts of meteorological variables can be nonlinear through interactions with local ecosystem (Seneviratne et al., 2010), suggesting that even with same meteorological variable the simulated SM can be totally different (e.g., with different soil texture or land cover types). Moreover, SM can be strongly coupled with atmosphere (Koster, 2004; Taylor et al., 2012), implying that meteorological factors can be influenced by SM as well (such as cloudiness, precipitation, air humidity, etc), which is not included in this study. However, the logic of our importance analysis is simple. If the model inputs (forcing data) were the same, the outputs (SM) should be the same. In other words, the differences of outputs can only be caused by the difference of inputs in our simulation results. It does not matter whether the quality of atmospheric forcing is good. On the contrary, the more differences exists among these forcing datasets, the better our analysis is. To keep the analysis simple, we did not investigate temporal correlations in each pixel but focused on spatial patterns of D at continental scale. Therefore, our results provided a general estimation of the importance of meteorological variable uncertainties to SM simulation through ORCHIDEE.

Indeed this approach is not able to demonstrate explicit links between meteorological variables and SM. We underlined the impacts of N_p and Q_a uncertainties, but it does not mean that other factors are unimportant. For instance, assuming that a variable can strongly influence simulated SM, if there was no much difference of the variable among forcing datasets, its importance cannot be detected in this work. Moreover, only one model was used in this study. Although ORCHIDEE performed very well in SM simulation, the lack of unknown mechanisms may weaken the linkage between SM and specific atmospheric variables. In one word, our analysis only focused on the inputs and outputs of the model and tried to diagnose the relationship between their differences.

6 Conclusions

Simulations in China were performed in ORCHIDEE-MICT driven by different forcing datasets: GSWP3, PGF, CRU-NCEP, and WFDEI. Simulated soil moisture was compared to several datasets to evaluate the ability of ORCHIDEE-MICT in reproducing soil moisture dynamics in China. Results showed that ORCHIDEE soil moisture coincided well with other datasets in wet areas and in non-irrigated areas. It suggested that the ORCHIDEE-MICT was suitable for further hydrological studies in China. However, the abnormal variation of observed SM in North China Plain implied potential impacts of irrigation, which was recommended to be considered in further simulations. Moreover, results showed that bias was mainly from model parameterization and atmospheric forcing. Thus parameterizations in ORCHIDEE-MICT should be calibrated, and atmospheric forcing should be carefully selected to meet the situation of China.

Several criteria were chosen and compared among the four simulations in China, YZRB, and YLRB. Results showed that GSWP3 and WFDEI, which had the best performances in correlation coefficients and RMSE respectively, were ideal choices for hydrological study in China. However, higher MSD in the Yellow River basin reflected the complicated climate condition in northern China, which might be significantly influenced by human activities as well. Finally, we used the differences of simulated soil moisture and meteorological variables to simply investigate the linkage between them. Results showed that the differences of simulated soil moisture were mainly explained by the differences of air humidity and precipitation frequency among the four atmospheric forcing. However, this coarse analysis cannot give explicit explanations about related mechanisms. Further study is needed to discover the interactions between soil water and climate through tracing the surface hydrological cycles and energy balances.

Code and data availability. The SVN version of ORCHIDEE-MICT used in this study is 3952, which is available at <https://forge.ipsl.jussieu.fr/orchidee/wiki/DevelopmentActivities/ORCHIDEE-MICT-IMBALANCE-P>. The ORCHIDEE code and scripts of analysis are available by contacting the correspond author. The GSWP3, PGF, WFDEI, ISMN, GLEAM and ESA CCI datasets are freely available online; for the CRU-NCEP and PKU datasets, please contact the corresponding author and Prof. Shilong Piao (slpiao@pku.edu.cn) respectively.

Members of the China-Trend-Stream project are (alphabetically): Philippe Ciais¹, Patrice Dumas², Xiaoming Feng³, Matthieu Guimberteau^{1,4}, Laurent Li⁵, Catherine Ottlé¹, Shushi Peng⁶, Shilong Piao⁶, Jan Polcher⁵, Pengfei Shi⁷, Shuai Wang³, Xuhui Wang^{1,5,6}, Yi Xi⁶, Hui Yang⁶, Tao Yang⁷, Zun Yin¹, Xuanze Zhang⁶, Feng Zhou⁶, and Xudong Zhou⁷.

¹Laboratoire des Sciences du Climat et de l'Environnement, CNRS-CEA-UVSQ, Gif-sur-Yvette 91191, France

²Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Avenue Agropolis, 34398 Montpellier Cedex 5, France

³State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China

⁴UMR 7619 METIS, Sorbonne Universités, UPMC, CNRS, EPHE, 4 place Jussieu, Paris 75005, France

⁵Laboratoire de Météorologie Dynamique, UPMC/CNRS, IPSL, Paris 75005, France

⁶Sino-French Institute for Earth System Science, College of Urban and Environmental Sciences, Peking University, Beijing 100871, China

⁷State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Center for Global Change and Water Cycle, Hohai University, Nanjing 210098, China

5 *Author contributions.* P. Ciais, C. Otlé and Z. Yin designed research. Z. Yin performed research, analyzed data and wrote the draft; all authors contributed to interpreting results, discussing findings and improving the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This study was supported by the National Natural Science Foundation of China (grant number 41561134016) and by the CHINA-TREND-STREAM French national project (ANR Grant No. ANR-15-CE01-00L1-0L). M. Guimberteau and P. Ciais acknowledge
10 support from the European Research Council Synergy grant ERC-2013-SyG-610028 IMBALANCE-P. **H. Kim was supported by Japan Society for the Promotion of Science KAKENHI (16H06291).** We thank Brecht Martens and Suxia Liu for helpful discussion about GLEAM and ISMN datasets. **We gratefully acknowledge two anonymous referees and the editor for their helpful comments and efforts.**

References

- Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P. P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., Arkin, P., and Nelkin, E.: The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–Present), *Journal of Hydrometeorology*, 4, 1147–1167, [https://doi.org/10.1175/1525-7541\(2003\)004<1147:TVGPCP>2.0.CO;2](https://doi.org/10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2), 2003.
- 5 An, R., Zhang, L., Wang, Z., Quaye-Ballard, J. A., You, J. J., Shen, X. J., Gao, W., Huang, L. J., Zhao, Y. H., and Ke, Z. Y.: Validation of the ESA CCI soil moisture product in China, *International Journal of Applied Earth Observation and Geoinformation*, 48, 28–36, <https://doi.org/10.1016/j.jag.2015.09.009>, <http://linkinghub.elsevier.com/retrieve/pii/S0303243415300325>, 2016.
- An, W. M., Li, Z. S., Wang, S., Wu, X., Lu, Y. H., Liu, G. H., and Fu, B. J.: Exploring the effects of the “Grain for Green” program on the differences in soil water in the semi-arid Loess Plateau of China, *Ecological Engineering*, 107, 144–151, <https://doi.org/10.1016/j.ecoleng.2017.07.017>, <http://linkinghub.elsevier.com/retrieve/pii/S0925857417304147>, 2017.
- 10 Ayalew, T. B., Krajewski, W. F., Mantilla, R., and Small, S. J.: Exploring the effects of hillslope-channel link dynamics and excess rainfall properties on the scaling structure of peak-discharge, *Advances in Water Resources*, 64, 9–20, <https://doi.org/10.1016/j.advwatres.2013.11.010>, 2014.
- 15 Basheer, M. and Elagib, N. A.: Sensitivity of Water-Energy Nexus to dam operation: A Water-Energy Productivity concept, *Science of the Total Environment*, 616–617, 918–926, <https://doi.org/10.1016/j.scitotenv.2017.10.228>, <http://linkinghub.elsevier.com/retrieve/pii/S0048969717329467>, 2018.
- Baudena, M. and Provenzale, A.: Rainfall intermittency and vegetation feedbacks in drylands, *Hydrology and Earth System Sciences*, 12, 679–689, <https://doi.org/10.5194/hess-12-679-2008>, 2008.
- 20 Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., and de Roo, A.: MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data, *Hydrology and Earth System Sciences Discussions*, pp. 1–38, <https://doi.org/10.5194/hess-2016-236>, 2016.
- Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., and Ziese, M.: A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present, *Earth System Science Data*, 5, 71–99, <https://doi.org/10.5194/essd-5-71-2013>, 2013.
- 25 Bonan, G. B.: Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests, *Science*, 320, 1444–1449, <https://doi.org/10.1126/science.1155121>, 2008.
- Bouwer, L. M., Bubeck, P., Wagtendonk, A. J., and Aerts, J. C. J. H.: Inundation scenarios for flood damage evaluation in polder areas, *Natural Hazards and Earth System Science*, 9, 1995–2007, <https://doi.org/10.5194/nhess-9-1995-2009>, <http://www.nat-hazards-earth-syst-sci.net/9/1995/2009/>, 2009.
- 30 Calvet, J. C., Rivalland, V., Picon-Cochard, C., and Guehl, J. M.: Modelling forest transpiration and CO₂ fluxes—response to soil moisture stress, *Agricultural and Forest Meteorology*, 124, 143–156, <https://doi.org/10.1016/j.agrformet.2004.01.007>, 2004.
- Cissé, S., Eymard, L., Otlé, C., Ndione, J., Gaye, A., and Pinsard, F.: Rainfall Intra-Seasonal Variability and Vegetation Growth in the Ferlo Basin (Senegal), *Remote Sensing*, 8, 66, <https://doi.org/10.3390/rs8010066>, 2016.
- 35 Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J.,

- Maugeri, M., Mok, H. Y., Nordli, Ø., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D., and Worley, S. J.: The Twentieth Century Reanalysis Project, *Quarterly Journal of the Royal Meteorological Society*, 137, 1–28, <https://doi.org/10.1002/qj.776>, 2011.
- Dai, A. G., Trenberth, K. E., and Qian, T. T.: A Global Dataset of Palmer Drought Severity Index for 1870–2002: Relationship with Soil Moisture and Effects of Surface Warming, *Journal of Hydrometeorology*, 5, 1117–1130, <https://doi.org/10.1175/JHM-386.1>, 2004.
- 5 Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597, <https://doi.org/10.1002/qj.828>, 2011.
- 10 Deng, K., Yang, S. Y., Lian, E. G., Li, C., Yang, C. F., and Wei, H. L.: Three Gorges Dam alters the Changjiang (Yangtze) river water cycle in the dry seasons: Evidence from H-O isotopes, *Science of The Total Environment*, 562, 89–97, <https://doi.org/10.1016/j.scitotenv.2016.03.213>, <http://linkinghub.elsevier.com/retrieve/pii/S0048969716306611>, 2016.
- Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., van Oevelen, P., 15 Robock, A., and Jackson, T.: The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements, *Hydrology and Earth System Sciences*, 15, 1675–1698, <https://doi.org/10.5194/hess-15-1675-2011>, 2011.
- Dorigo, W. A., Gruber, A., De Jeu, R. A. M., Wagner, W., Stacke, T., Loew, A., Albergel, C., Brocca, L., Chung, D., Parinussa, R. M., and Kidd, R.: Evaluation of the ESA CCI soil moisture product using ground-based observations, *Remote Sensing of Environment*, 162, 380–395, <https://doi.org/10.1016/j.rse.2014.07.023>, 2015.
- 20 Draper, C. S., Reichle, R. H., De Lannoy, G. J. M., and Liu, Q.: Assimilation of passive and active microwave soil moisture retrievals, *Geophysical Research Letters*, 39, n/a–n/a, <https://doi.org/10.1029/2011GL050655>, <http://doi.wiley.com/10.1029/2011GL050655>, 2012.
- Farquhar, G. D. and Sharkey, T. D.: Stomatal Conductance and Photosynthesis, *Annual Review of Plant Physiology*, 33, 317–345, <https://doi.org/10.1146/annurev.pp.33.060182.001533>, 1982.
- Feng, X. M., Fu, B. J., Piao, S. L., Wang, S., Ciais, P., Zeng, Z. Z., Lü, Y. H., Zeng, Y., Li, Y., Jiang, X. H., and Wu, B. F.: 25 Revegetation in China’s Loess Plateau is approaching sustainable water resource limits, *Nature Climate Change*, 6, 1019–1022, <https://doi.org/10.1038/nclimate3092>, <http://www.nature.com/doi/10.1038/nclimate3092>, 2016.
- Getirana, A. C. V., Dutra, E., Guimberteau, M., Kam, J., Li, H. Y., Decharme, B., Zhang, Z. Q., Ducharne, A., Boone, A., Balsamo, G., Rodell, M., Toure, A. M., Xue, Y. K., Peters-Lidard, C. D., Kumar, S. V., Arsenault, K., Drapeau, G., Ruby Leung, L., Ronchail, J., and Sheffield, J.: Water Balance in the Amazon Basin from a Land Surface Model Ensemble, *Journal of Hydrometeorology*, 15, 2586–2614, <https://doi.org/10.1175/JHM-D-14-0068.1>, <http://journals.ametsoc.org/doi/abs/10.1175/JHM-D-14-0068.1>, 2014.
- 30 Grillakis, M. G., Koutroulis, A. G., Komma, J., Tsanis, I. K., Wagner, W., and Blöschl, G.: Initial soil moisture effects on flash flood generation? A comparison between basins of contrasting hydro-climatic conditions, *Journal of Hydrology*, 541, 206–217, <https://doi.org/10.1016/j.jhydrol.2016.03.007>, 2016.
- Guimberteau, M., Laval, K., Perrier, A., and Polcher, J.: Global effect of irrigation and its impact on the onset of the Indian summer monsoon, *Climate Dynamics*, 39, 1329–1348, <https://doi.org/10.1007/s00382-011-1252-5>, <http://link.springer.com/10.1007/s00382-011-1252-5>, 2012.
- 35

- Guimberteau, M., Ducharne, A., Ciais, P., Boisier, J. P., Peng, S., De Weirdt, M., and Verbeeck, H.: Testing conceptual and physically based soil hydrology schemes against observations for the Amazon Basin, *Geoscientific Model Development*, 7, 1115–1136, <https://doi.org/10.5194/gmd-7-1115-2014>, <http://www.geosci-model-dev.net/7/1115/2014/>, 2014.
- Guimberteau, M., Ciais, P., Pablo Boisier, J., Paula Dutra Aguiar, A., Biemans, H., De Deurwaerder, H., Galbraith, D., Kruijt, B., Langerwisch, F., Poveda, G., Rammig, A., Andres Rodriguez, D., Tejada, G., Thonicke, K., Von Randow, C., Randow, R., Zhang, K., and Verbeeck, H.: Impacts of future deforestation and climate change on the hydrology of the Amazon Basin: A multi-model analysis with a new set of land-cover change scenarios, *Hydrology and Earth System Sciences*, 21, 1455–1475, <https://doi.org/10.5194/hess-21-1455-2017>, <https://www.hydrol-earth-syst-sci.net/21/1455/2017/>, 2017.
- Guimberteau, M., Zhu, D., Maignan, F., Huang, Y., Yue, C., Dantec-Nédélec, S., Otlé, C., Jornet-Puig, A., Bastos, A., Laurent, P., Goll, D., Bowring, S., Chang, J. F., Guenet, B., Tifafi, M., Peng, S. S., Krinner, G., Ducharne, A., Wang, F. X., Wang, T., Wang, X. H., Wang, Y. L., Yin, Z., Lauerwald, R., Joetzer, E., Qiu, C. J., Kim, H., and Ciais, P.: ORCHIDEE-MICT (v8.4.1), a land surface model for the high latitudes: model description and validation, *Geoscientific Model Development*, 11, 121–163, <https://doi.org/10.5194/gmd-11-121-2018>, <https://www.geosci-model-dev.net/11/121/2018/>, 2018.
- Hansen, M. C., DeFries, R. S., Townshend, J. R. G., Carroll, M., Dimiceli, C., and Sohlberg, R. A.: Global Percent Tree Cover at a Spatial Resolution of 500 Meters: First Results of the MODIS Vegetation Continuous Fields Algorithm, *Earth Interactions*, 7, 1–15, [https://doi.org/10.1175/1087-3562\(2003\)007<0001:GPTCAA>2.0.CO;2](https://doi.org/10.1175/1087-3562(2003)007<0001:GPTCAA>2.0.CO;2), 2003.
- Harris, I., Jones, P. D., Osborn, T. J., and Lister, D. H.: Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset, *International Journal of Climatology*, 34, 623–642, <https://doi.org/10.1002/joc.3711>, 2014.
- Hirschi, M., Seneviratne, S. I., Alexandrov, V., Boberg, F., Boroneant, C., Christensen, O. B., Formayer, H., Orlowsky, B., and Stepanek, P.: Observational evidence for soil-moisture impact on hot extremes in southeastern Europe, *Nature Geoscience*, 4, 17–21, <https://doi.org/10.1038/ngeo1032>, 2011.
- Hirschi, M., Mueller, B., Dorigo, W., and Seneviratne, S. I.: Using remotely sensed soil moisture for land-atmosphere coupling diagnostics: The role of surface vs. root-zone soil moisture variability, *Remote Sensing of Environment*, 154, 246–252, <https://doi.org/10.1016/j.rse.2014.08.030>, <http://linkinghub.elsevier.com/retrieve/pii/S003442571400337X>, 2014.
- Hong, S. Y. and Chang, E. C.: Spectral nudging sensitivity experiments in a regional climate model, *Asia-Pacific Journal of Atmospheric Sciences*, 48, 345–355, <https://doi.org/10.1007/s13143-012-0033-3>, 2012.
- Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G. J., Hong, Y., Bowman, K. P., and Stocker, E. F.: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales, *Journal of Hydrometeorology*, 8, 38–55, <https://doi.org/10.1175/JHM560.1>, 2007.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-Year Reanalysis Project, *Bulletin of the American Meteorological Society*, 77, 437–471, [https://doi.org/10.1175/1520-0477\(1996\)077<0437:TNYRP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2), 1996.
- Kendall, M. G.: Rank correlation methods, Charles Griffin, London, 4th edn., 1975.
- Kerr, Y. H., Waldteufel, P., Wigneron, J. P., Martinuzzi, J., Font, J., and Berger, M.: Soil moisture retrieval from space: the Soil Moisture and Ocean Salinity (SMOS) mission, *IEEE Transactions on Geoscience and Remote Sensing*, 39, 1729–1735, <https://doi.org/10.1109/36.942551>, 2001.

- Kim, H.: Global Soil Wetness Project Phase 3 Atmospheric Boundary Conditions (Experiment 1) [Data set], Data Integration and Analysis System (DIAS), <https://doi.org/10.20783/DIAS.501>, 2017.
- Kobayashi, K. and Salam, M. U.: Comparing Simulated and Measured Values Using Mean Squared Deviation and its Components, *Agronomy Journal*, 92, 345, <https://doi.org/10.1007/s100870050043>, 2000.
- 5 Koster, R. D.: Regions of Strong Coupling Between Soil Moisture and Precipitation, *Science*, 305, 1138–1140, <https://doi.org/10.1126/science.1100217>, 2004.
- Kottek, M., Grieser, J., Beck, C., Rudolf, B., and Rubel, F.: World Map of the Köppen-Geiger climate classification updated, *Meteorologische Zeitschrift*, 15, 259–263, <https://doi.org/10.1127/0941-2948/2006/0130>, 2006.
- Krinner, G., Viovy, N., de Noblet-Ducoudre, N., Ogée, J., Polcher, J., Friedlingstein, P., Ciais, P., Sitch, S., and Prentice, I. C.: A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, *Global Biogeochemical Cycles*, 19, 1–33, <https://doi.org/10.1029/2003GB002199>, 2005.
- 10 Lai, X., Wen, J., Cen, S. X., Huang, X., Tian, H., and Shi, X. K.: Spatial and Temporal Soil Moisture Variations over China from Simulations and Observations, *Advances in Meteorology*, 2016, 1–14, <https://doi.org/10.1155/2016/4587687>, 2016.
- Li, L. C., Zhang, L. P., Xia, J., Gippel, C. J., Wang, R. C., and Zeng, S. D.: Implications of Modelled Climate and Land Cover Changes on Runoff in the Middle Route of the South to North Water Transfer Project in China, *Water Resources Management*, 29, 2563–2579, <https://doi.org/10.1007/s11269-015-0957-3>, <http://link.springer.com/10.1007/s11269-015-0957-3>, 2015.
- 15 Li, L. L., Zhang, R. Z., Luo, Z. Z., Liang, W. L., Xie, J. H., Cai, L. Q., and Bellotti, B.: Evolution of soil and water conservation in rain-fed areas of China, *International Soil and Water Conservation Research*, 2, 78–90, [https://doi.org/10.1016/S2095-6339\(15\)30015-0](https://doi.org/10.1016/S2095-6339(15)30015-0), 2014.
- Liu, B., Xu, M., and Henderson, M.: Where have all the showers gone? Regional declines in light precipitation events in China, 1960-2000, *International Journal of Climatology*, 31, 1177–1191, <https://doi.org/10.1002/joc.2144>, 2011.
- 20 Liu, S. X., Mo, X. G., Li, H. B., Peng, G. B., and Robock, A.: Spatial Variation of Soil Moisture in China: Geostatistical Characterization., *Journal of the Meteorological Society of Japan*, 79, 555–574, <https://doi.org/10.2151/jmsj.79.555>, 2001.
- Liu, W. F., Wei, X. H., Fan, H. B., Guo, X. M., Liu, Y. Q., Zhang, M. F., and Li, Q.: Response of flow regimes to deforestation and reforestation in a rain-dominated large watershed of subtropical China, *Hydrological Processes*, 29, 5003–5015, <https://doi.org/10.1002/hyp.10459>, <http://doi.wiley.com/10.1002/hyp.10459>, 2015a.
- 25 Liu, Y. L., Pan, Z. H., Zhuang, Q. L., Miralles, D. G., Teuling, A. J., Zhang, T. L., An, P. L., Dong, Z. Q., Zhang, J. T., He, D., Wang, L. W., Pan, X. B., Bai, W., and Niyogi, D.: Agriculture intensifies soil moisture decline in Northern China, *Scientific Reports*, 5, 11 261, <https://doi.org/10.1038/srep11261>, <http://www.nature.com/articles/srep11261>, 2015b.
- Liu, Y. W., Wang, W., and Hu, Y. M.: Investigating the impact of surface soil moisture assimilation on state and parameter estimation in SWAT model based on the ensemble Kalman filter in upper Huai River basin, *Journal of Hydrology and Hydromechanics*, 65, 123–133, <https://doi.org/10.1515/johh-2017-0011>, <http://content.sciendo.com/view/journals/johh/65/2/article-p123.xml>, 2017.
- 30 Liu, Y. Y., Dorigo, W. A., Parinussa, R. M., de Jeu, R. A. M., Wagner, W., McCabe, M. F., Evans, J. P., and van Dijk, A. I. J. M.: Trend-preserving blending of passive and active microwave soil moisture retrievals, *Remote Sensing of Environment*, 123, 280–297, <https://doi.org/10.1016/j.rse.2012.03.014>, 2012.
- 35 Mann, H. B.: Nonparametric Tests Against Trend, *Econometrica*, 13, 245, <https://doi.org/10.2307/1907187>, 1945.
- Martens, B., Miralles, D., Lievens, H., Fernández-Prieto, D., and Verhoest, N. E. C.: Improving terrestrial evaporation estimates over continental Australia through assimilation of SMOS soil moisture, *International Journal of Applied Earth Observation and Geoinformation*, 48, 146–162, <https://doi.org/10.1016/j.jag.2015.09.012>, 2016.

- Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, *Geoscientific Model Development*, 10, 1903–1925, <https://doi.org/10.5194/gmd-10-1903-2017>, 2017.
- Maurer, E. P., O'Donnell, G. M., Lettenmaier, D. P., and Roads, J. O.: Evaluation of NCEP/NCAR reanalysis water and energy budgets using macroscale hydrologic model simulations, in: *Land Surface Hydrology, Meteorology, and Climate: Observations and Modeling*, pp. 137–158, Wiley Online Library, <https://doi.org/10.1029/WS003p0137>, 2001.
- Monteith, J. L.: Evaporation and environment, <https://doi.org/10.1613/jair.301>, 1965.
- Njoku, E. G., Jackson, T. J., Lakshmi, V., Chan, T. K., and Nghiem, S. V.: Soil moisture retrieval from AMSR-E, *IEEE Transactions on Geoscience and Remote Sensing*, 41, 215–229, <https://doi.org/10.1109/TGRS.2002.808243>, 2003.
- 10 Peng, J., Niesel, J., Loew, A., Zhang, S. Q., and Wang, J.: Evaluation of Satellite and Reanalysis Soil Moisture Products over Southwest China Using Ground-Based Measurements, *Remote Sensing*, 7, 15 729–15 747, <https://doi.org/10.3390/rs71115729>, <http://www.mdpi.com/2072-4292/7/11/15729>, 2015.
- Peng, S. S., Piao, S., Zeng, Z., Ciais, P., Zhou, L., Li, L. Z. X., Myneni, R. B., Yin, Y., and Zeng, H.: Afforestation in China cools local land surface temperature, *Proceedings of the National Academy of Sciences*, 111, 2915–2919, <https://doi.org/10.1073/pnas.1315126111>,
15 <http://www.pnas.org/cgi/doi/10.1073/pnas.1315126111>, 2014.
- Piao, S. L., Yin, L., Wang, X. H., Ciais, P., Peng, S. S., Shen, Z. H., and Seneviratne, S. I.: Summer soil moisture regulated by precipitation frequency in China, *Environmental Research Letters*, 4, 044 012, <https://doi.org/10.1088/1748-9326/4/4/044012>, 2009.
- Piao, S. L., Ciais, P., Huang, Y., Shen, Z. H., Peng, S. S., Li, J. S., Zhou, L. P., Liu, H. Y., Ma, Y. C., Ding, Y. H., Friedlingstein, P., Liu, C. Z., Tan, K., Yu, Y. Q., Zhang, T. Y., and Fang, J. Y.: The impacts of climate change on water resources and agriculture in China, *Nature*, 467,
20 43–51, <https://doi.org/10.1038/nature09364>, 2010.
- Pierdicca, N., Fascetti, F., Pulvirenti, L., Crapolicchio, R., and Muñoz-Sabater, J.: Analysis of ASCAT, SMOS, in-situ and land model soil moisture as a regionalized variable over Europe and North Africa, *Remote Sensing of Environment*, 170, 280–289, <https://doi.org/10.1016/j.rse.2015.09.005>, 2015.
- Polcher, J., Piles, M., Gelati, E., Barella-Ortiz, A., and Tello, M.: Comparing surface-soil moisture from the SMOS mission and the ORCHIDEE land-surface model over the Iberian Peninsula, *Remote Sensing of Environment*, 174, 69–81, <https://doi.org/10.1016/j.rse.2015.12.004>, <http://linkinghub.elsevier.com/retrieve/pii/S003442571530225X>, 2016.
- Puma, M. J. and Cook, B. I.: Effects of irrigation on global climate during the 20th century, *Journal of Geophysical Research*, 115, D16 120, <https://doi.org/10.1029/2010JD014122>, <http://doi.wiley.com/10.1029/2010JD014122>, 2010.
- Rebel, K. T., de Jeu, R. A. M., Ciais, P., Viovy, N., Piao, S. L., Kiely, G., and Dolman, A. J.: A global analysis of soil moisture derived from
30 satellite observations and a land surface model, *Hydrology and Earth System Sciences*, 16, 833–847, <https://doi.org/10.5194/hess-16-833-2012>, 2012.
- Reichle, R. H., Koster, R. D., Liu, P., Mahanama, S. P. P., Njoku, E. G., and Owe, M.: Comparison and assimilation of global soil moisture retrievals from the Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) and the Scanning Multichannel Microwave Radiometer (SMMR), *Journal of Geophysical Research Atmospheres*, 112, D09 108, <https://doi.org/10.1029/2006JD008033>,
35 <http://doi.wiley.com/10.1029/2006JD008033>, 2007.
- Robock, A., Vinnikov, K. Y., Srinivasan, G., Entin, J. K., Hollinger, S. E., Speranskaya, N. A., Liu, S. X., and Namkhai, A.: The Global Soil Moisture Data Bank, *Bulletin of the American Meteorological Society*, 81, 1281–1299, [https://doi.org/10.1175/1520-0477\(2000\)081<1281:TGSMDB>2.3.CO;2](https://doi.org/10.1175/1520-0477(2000)081<1281:TGSMDB>2.3.CO;2), 2000.

- Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin*, J. K., Walker, J. P., Lohmann, D., and Toll, D.: The Global Land Data Assimilation System, *Bulletin of the American Meteorological Society*, 85, 381–394, <https://doi.org/10.1175/BAMS-85-3-381>, 2004.
- 5 Rogers, S., Barnett, J., Webber, M., Finlayson, B., and Wang, M.: Governmentality and the conduct of water: China’s South–North Water Transfer Project, *Transactions of the Institute of British Geographers*, 41, 429–441, <https://doi.org/10.1111/tran.12141>, <http://doi.wiley.com/10.1111/tran.12141>, 2016.
- Schmied, H. M., Adam, L., Eisner, S., Fink, G., Flörke, M., Kim, H., Oki, T., Portmann, F. T., Reinecke, R., Riedel, C., Song, Q., Zhang, J., and Döll, P.: Impact of climate forcing uncertainty and human water use on global and continental water balance components, *Proceedings of the International Association of Hydrological Sciences*, 374, 53–62, <https://doi.org/10.5194/piahs-374-53-2016>, <http://www.proc-iahs.net/374/53/2016/>, 2016.
- 10 Schymanski, S. J., Roderick, M. L., Sivapalan, M., Hutley, L. B., and Beringer, J.: A canopy-scale test of the optimal water-use hypothesis, *Plant, Cell and Environment*, 31, 97–111, <https://doi.org/10.1111/j.1365-3040.2007.01740.x>, 2008.
- Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling, A. J.: Investigating soil moisture–climate interactions in a changing climate: A review, *Earth-Science Reviews*, 99, 125–161, <https://doi.org/10.1016/j.earscirev.2010.02.004>, 2010.
- 15 Seneviratne, S. I., Wilhelm, M., Stanelle, T., van den Hurk, B. J. J. M., Hagemann, S., Berg, A., Cheruy, F., Higgins, M. E., Meier, A., Brovkin, V., Claussen, M., Ducharne, A., Dufresne, J. L., Findell, K. L., Ghattas, J., Lawrence, D. M., Malyshev, S., Rummukainen, M., and Smith, B.: Impact of soil moisture-climate feedbacks on CMIP5 projections: First results from the GLACE-CMIP5 experiment, *Geophysical Research Letters*, 40, 5212–5217, <https://doi.org/10.1002/grl.50956>, 2013.
- 20 Sheffield, J., Goteti, G., and Wood, E. F.: Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling, *Journal of Climate*, 19, 3088–3111, <https://doi.org/10.1175/JCLI3790.1>, 2006.
- Stephens, G. L., Wild, M., Stackhouse, P. W., L’Ecuyer, T., Kato, S., and Henderson, D. S.: The Global Character of the Flux of Downward Longwave Radiation, *Journal of Climate*, 25, 2329–2340, <https://doi.org/10.1175/JCLI-D-11-00262.1>, 2012.
- Su, Z. B., Yacob, A., Wen, J., Roerink, G., He, Y. B., Gao, B. H., Boogaard, H., and van Diepen, C.: Assessing relative soil moisture with remote sensing data: theory, experimental validation, and application to drought monitoring over the North China Plain, *Physics and Chemistry of the Earth, Parts A/B/C*, 28, 89–101, [https://doi.org/10.1016/S1474-7065\(03\)00010-X](https://doi.org/10.1016/S1474-7065(03)00010-X), 2003.
- 25 Tangdamrongsub, N., Han, S. C., Decker, M., Yeo, I. Y., and Kim, H.: On the use of the GRACE normal equation of inter-satellite tracking data for estimation of soil moisture and groundwater in Australia, *Hydrology and Earth System Sciences*, 22, 1811–1829, <https://doi.org/10.5194/hess-22-1811-2018>, <https://www.hydrol-earth-syst-sci.net/22/1811/2018/>, 2018.
- 30 Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An Overview of CMIP5 and the Experiment Design, *Bulletin of the American Meteorological Society*, 93, 485–498, <https://doi.org/10.1175/BAMS-D-11-00094.1>, 2012.
- Teuling, A. J., Seneviratne, S. I., Stöckli, R., Reichstein, M., Moors, E., Ciais, P., Luysaert, S., van den Hurk, B. J. J. M., Ammann, C., Bernhofer, C., Dellwik, E., Gianelle, D., Gielen, B., Grünwald, T., Klumpp, K., Montagnani, L., Moureaux, C., Sottocornola, M., and Wohlfahrt, G.: Contrasting response of European forest and grassland energy exchange to heatwaves, *Nature Geoscience*, 3, 722–727, <https://doi.org/10.1038/ngeo950>, 2010.
- 35 Van Den Hurk, B., Kim, H., Krinner, G., Seneviratne, S. I., Derksen, C., Oki, T., Douville, H., Colin, J., Ducharne, A., Cheruy, F., Viovy, N., Puma, M. J., Wada, Y., Li, W. P., Jia, B. H., Alessandri, A., Lawrence, D. M., Weedon, G. P., Ellis, R., Hagemann, S., Mao, J. F., Flanner, M. G., Zampieri, M., Matera, S., Law, R. M., and Sheffield, J.: LS3MIP (v1.0) contribution to CMIP6: The Land Surface, Snow

- and Soil moisture Model Intercomparison Project - Aims, setup and expected outcome, *Geoscientific Model Development*, 9, 2809–2832, <https://doi.org/10.5194/gmd-9-2809-2016>, <http://www.geosci-model-dev.net/9/2809/2016/>, 2016.
- Wada, Y., de Graaf, I. E. M., and van Beek, L. P. H.: High-resolution modeling of human and climate impacts on global water resources, *Journal of Advances in Modeling Earth Systems*, 8, 735–763, <https://doi.org/10.1002/2015MS000618>, <http://doi.wiley.com/10.1002/2015MS000618>, 2016.
- Wada, Y., Bierkens, M. F. P., de Roo, A., Dirmeyer, P. A., Famiglietti, J. S., Hanasaki, N., Konar, M., Liu, J. G., Müller Schmied, H., Oki, T., Pokhrel, Y., Sivapalan, M., Troy, T. J., van Dijk, A. I. J. M., van Emmerik, T., Van Huijgevoort, M. H. J., Van Lanen, H. A. J., Vörösmarty, C. J., Wanders, N., and Wheatler, H.: Human–water interface in hydrological modelling: current status and future directions, *Hydrology and Earth System Sciences*, 21, 4169–4193, <https://doi.org/10.5194/hess-21-4169-2017>, <https://www.hydrol-earth-syst-sci.net/21/4169/2017/>, 2017.
- Wagner, W., Dorigo, W., de Jeu, R., Fernandez, D., Benveniste, J., Haas, E., and Ertl, M.: Fusion of Active and Passive Microwave Observations To Create an Essential Climate Variable Data Record on Soil Moisture, *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, I-7, 315–321, <https://doi.org/10.5194/isprsannals-I-7-315-2012>, 2012.
- Wanders, N., Karssenber, D., de Roo, A., de Jong, S. M., and Bierkens, M. F. P.: The suitability of remotely sensed soil moisture for improving operational flood forecasting, *Hydrology and Earth System Sciences*, 18, 2343–2357, <https://doi.org/10.5194/hess-18-2343-2014>, 2014.
- Wang, S. S., Mo, X. G., Liu, S. X., Lin, Z. H., and Hu, S.: Validation and trend analysis of ECV soil moisture data on cropland in North China Plain during 1981–2010, *International Journal of Applied Earth Observation and Geoinformation*, 48, 110–121, <https://doi.org/10.1016/j.jag.2015.10.010>, 2016.
- Wang, X. H., Ciais, P., Li, L., Ruget, F., Vuichard, N., Viovy, N., Zhou, F., Chang, J. F., Wu, X. C., Zhao, H. F., and Piao, S. L.: Management outweighs climate change on affecting length of rice growing period for early rice and single rice in China during 1991–2012, *Agricultural and Forest Meteorology*, 233, 1–11, <https://doi.org/10.1016/j.agrformet.2016.10.016>, 2017.
- Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., and Viterbo, P.: The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA-Interim reanalysis data, *Water Resources Research*, 50, 7505–7514, <https://doi.org/10.1002/2014WR015638>, 2014.
- Wei, X. H., Li, Q., Zhang, M. F., Giles-Hansen, K., Liu, W. F., Fan, H. B., Wang, Y., Zhou, G. Y., Piao, S. L., and Liu, S. R.: Vegetation cover—another dominant factor in determining global water resources in forested regions, *Global Change Biology*, 24, 786–795, <https://doi.org/10.1111/gcb.13983>, <http://doi.wiley.com/10.1111/gcb.13983>, 2018.
- Williams, M. R., King, K. W., and Fausey, N. R.: Drainage water management effects on tile discharge and water quality, *Agricultural Water Management*, 148, 43–51, <https://doi.org/10.1016/j.agwat.2014.09.017>, 2015.
- Wu, Y. L., Peng, S., Ciais, P., Guimberteau, M., Piao, S. L., Polcher, J., and Zhou, F.: Estimating water withdrawals and its impacts on water budget of an eutrophic Lake Dianchi, China, *Journal of Hydrology*, 2018.
- Xia, Y., Sheffield, J., Ek, M. B., Dong, J., Chaney, N., Wei, H., Meng, J., and Wood, E. F.: Evaluation of multi-model simulated soil moisture in NLDAS-2, *Journal of Hydrology*, 512, 107–125, <https://doi.org/10.1016/j.jhydrol.2014.02.027>, 2014.
- Xu, W. N. e. a.: Agro-meteorological observation standard (eds Xu, W. N. et al.), Tech. rep., China Meteorological Administration, 2014.
- Yan, R., Gao, J., and Li, L.: Modeling the hydrological effects of climate and land use/cover changes in Chinese lowland polder using an improved WALRUS model, *Hydrology Research*, 47, 84–101, <https://doi.org/10.2166/nh.2016.204>, <http://hr.iwaponline.com/cgi/doi/10.2166/nh.2016.204>, 2016.

- Yang, H., Piao, S. L., Zeng, Z. Z., Ciais, P., Yin, Y., Friedlingstein, P., Sitch, S., Ahlström, A., Guimberteau, M., Huntingford, C., Levis, S., Levy, P. E., Huang, M. T., Li, Y., Li, X. R., Lomas, M. R., Peylin, P., Poulter, B., Viovy, N., Zaehle, S., Zeng, N., Zhao, F., and Wang, L.: Multicriteria evaluation of discharge simulation in Dynamic Global Vegetation Models, *Journal of Geophysical Research: Atmospheres*, 120, 7488–7505, <https://doi.org/10.1002/2015JD023129>, <http://doi.wiley.com/10.1002/2015JD023129>, 2015.
- 5 Yang, J., Gong, D. Y., Wang, W. S., Hu, M., and Mao, R.: Extreme drought event of 2009/2010 over southwestern China, *Meteorology and Atmospheric Physics*, 115, 173–184, <https://doi.org/10.1007/s00703-011-0172-6>, 2012.
- Ye, J. S., Li, W. H., Li, L. F., and Zhang, F.: "North drying and south wetting" summer precipitation trend over China and its potential linkage with aerosol loading, *Atmospheric Research*, 125-126, 12–19, <https://doi.org/10.1016/j.atmosres.2013.01.007>, 2013.
- Yin, Z., Dekker, S. C., van den Hurk, B. J. J. M., and Dijkstra, H. A.: Bimodality of woody cover and biomass across the precipitation gradient in West Africa, *Earth System Dynamics*, 5, 257–270, <https://doi.org/10.5194/esd-5-257-2014>, 2014.
- 10 Yoshimura, K. and Kanamitsu, M.: Dynamical Global Downscaling of Global Reanalysis, *Monthly Weather Review*, 136, 2983–2998, <https://doi.org/10.1175/2008MWR2281.1>, 2008.
- Yoshimura, K. and Kanamitsu, M.: Incremental Correction for the Dynamical Downscaling of Ensemble Mean Atmospheric Fields, *Monthly Weather Review*, 141, 3087–3101, <https://doi.org/10.1175/MWR-D-12-00271.1>, 2013.
- 15 Zhai, P. M., Zhang, X. B., Wan, H., and Pan, X. H.: Trends in Total Precipitation and Frequency of Daily Precipitation Extremes over China, *Journal of Climate*, 18, 1096–1108, <https://doi.org/10.1175/JCLI-3318.1>, 2005.
- Zhao, F., Veldkamp, T. I. E., Frieler, K., Schewe, J., Ostberg, S., Willner, S., Schauburger, B., Gosling, S. N., Schmied, H. M., Portmann, F. T., Leng, G. Y., Huang, M. Y., Liu, X. C., Tang, Q. H., Hanasaki, N., Biemans, H., Gerten, D., Satoh, Y., Pokhrel, Y., Stacke, T., Ciais, P., Chang, J. F., Ducharne, A., Guimberteau, M., Wada, Y., Kim, H., and Yamazaki, D.: The critical role of the routing scheme in simulating peak river discharge in global hydrological models, *Environmental Research Letters*, 12, 075 003, <https://doi.org/10.1088/1748-9326/aa7250>, 2017.
- 20 Zhou, X. D., Polcher, J., Yang, T., Hirabayashi, Y., and Nguyen-Quang, T.: Understanding the water cycle over the upper Tarim basin: retrospect the estimated discharge bias to atmospheric variables and model structure, *Hydrology and Earth System Sciences Discussions*, pp. 1–38, <https://doi.org/10.5194/hess-2018-88>, <https://www.hydrol-earth-syst-sci-discuss.net/hess-2018-88/>, 2018.
- 25 Zhu, D., Peng, S. S., Ciais, P., Viovy, N., Druel, A., Kageyama, M., Krinner, G., Peylin, P., Ottlé, C., Piao, S. L., Poulter, B., Schepaschenko, D., and Shvidenko, A.: Improving the dynamics of Northern Hemisphere high-latitude vegetation in the ORCHIDEE ecosystem model, *Geoscientific Model Development*, 8, 2263–2283, <https://doi.org/10.5194/gmd-8-2263-2015>, 2015.
- Zhu, Z. C., Piao, S. L., Myneni, R. B., Huang, M. T., Zeng, Z. Z., Canadell, J. G., Ciais, P., Sitch, S., Friedlingstein, P., Arneeth, A., Cao, C. X., Cheng, L., Kato, E., Koven, C., Li, Y., Lian, X., Liu, Y. W., Liu, R., Mao, J. F., Pan, Y. Z., Peng, S. S., Peñuelas, J., Poulter, B., Pugh, T. A. M., Stocker, B. D., Viovy, N., Wang, X. H., Wang, Y. P., Xiao, Z. Q., Yang, H., Zaehle, S., and Zeng, N.: Greening of the Earth and its drivers, *Nature Climate Change*, 6, 791–795, <https://doi.org/10.1038/nclimate3004>, 2016.
- 30 Zobler, L.: A world soil file for global climate modeling, Tech. rep., NASA Tech, 1986.

Tables

Table 1. General information of the climate forcing datasets. ‘Reanalysis’ and ‘Observations’ are corresponding datasets used in producing the atmospheric forcing. Detailed description can be found in Sect. 2.2.

Dataset	Resolution		Duration	Reanalysis	Observations
	Spatial	Temporal			
GSWP3	0.5°	3-hourly	1901–2010	20CR	GPCC, CRU TS, SRB
PGF	1°	3-hourly	1901–2012	NCEP-NCAR	CRU TS, GPCP, TRMM, SRB
CRU-NCEP	0.5°	6-hourly	1901–2015	NCEP	CRU TS
WFDEI	0.5°	3-hourly	1979–2009	ERA-Interim	CRU TS, GPCC

Table 2. Summary of the SM datasets for validation. ‘M+RS+RA’ indicates that the dataset is a model output driven by both remote sensing and reanalysis data. More details can be found in Sect. 2.3.

Dataset	Type	Unit	Resolution	Duration	Contents	Corresponding ORCHIDEE soil layer
				Analysis period	Analysis depth	
ISMN	in-situ	m ³ .m ⁻³	station, 10-day	1981-1999	11 layers; 0-100 cm	1-9 layers (0-75 cm)
				1984-1999	0-50 cm	
PKU	in-situ	% of porosity	station, 10-day	1991-2007	7 layers; 0-100 cm	1-8 layers (0-37 cm)
				1992-2006	0-30 cm	
ESA CCI	RS	m ³ .m ⁻³	0.25°, daily	1979-2010	Top layer, depth ≈ 0.5-2 cm	1-4 layers (0-2 cm)
				2007-2009		
GLEAM surface	M+RS +RA	m ³ .m ⁻³	0.25°, daily	1980-2014	0-10 cm	1-6 layers (0-9 cm)
				1981-2009		
GLEAM root-zone	M+RS +RA	m ³ .m ⁻³	0.25°, daily	1980-2014	Mixture of bare soil (1-10 cm), low vegetation (0-100 cm) and high vegetation (0-250 cm)	all layers (0-200 cm)
				1981-2009		

Table 3. Median of metrics in specific comparisons. The subscripts of correlation coefficients indicate the quantile of stations (samples) with significant correlation (p -value < 0.05).

Dataset	Simulations	Correlation			RMSE (m ³ .m ⁻³)		
ISMN	GSWP3	0.52 _{0.85}			0.07		
	PGF	0.46 _{0.90}			0.07		
	CRU-NCEP	0.36 _{0.95}			0.10		
	WFDEI	0.55 _{0.95}			0.08		
PKU	GSWP3	0.38 _{0.91}			NA		
	PGF	0.31 _{0.85}			NA		
	CRU-NCEP	0.31 _{0.86}			NA		
	WFDEI	0.45 _{0.93}			NA		
		China	Yangtze	Yellow	China	Yangtze	Yellow
ESA CCI	GSWP3	0.47 _{0.93}	0.42 _{0.94}	0.58 _{0.99}	0.06	0.06	0.06
	PGF	0.26 _{0.83}	0.32 _{0.91}	0.28 _{0.96}	0.06	0.07	0.07
	CRU-NCEP	0.51 _{0.94}	0.50 _{0.94}	0.54 _{0.99}	0.06	0.07	0.07
	WFDEI	0.61 _{0.97}	0.60 _{0.96}	0.68 ₁	0.05	0.05	0.06
GLEAM surface SM	GSWP3	0.54 ₁	0.60 ₁	0.52 ₁	0.07	0.07	0.10
	PGF	0.42 ₁	0.51 ₁	0.35 ₁	0.08	0.08	0.10
	CRU-NCEP	0.49 _{0.99}	0.61 ₁	0.49 ₁	0.10	0.10	0.12
	WFDEI	0.68 _{0.99}	0.77 ₁	0.63 ₁	0.08	0.09	0.10
GLEAM root-zone SM	GSWP3	0.60 _{0.98}	0.67 _{0.99}	0.60 _{0.99}	0.05	0.04	0.08
	PGF	0.57 _{0.98}	0.69 ₁	0.57 _{0.99}	0.06	0.04	0.09
	CRU-NCEP	0.40 _{0.96}	0.48 _{0.97}	0.37 _{0.97}	0.08	0.08	0.11
	WFDEI	0.63 _{0.98}	0.74 ₁	0.59 ₁	0.06	0.04	0.10

Figures

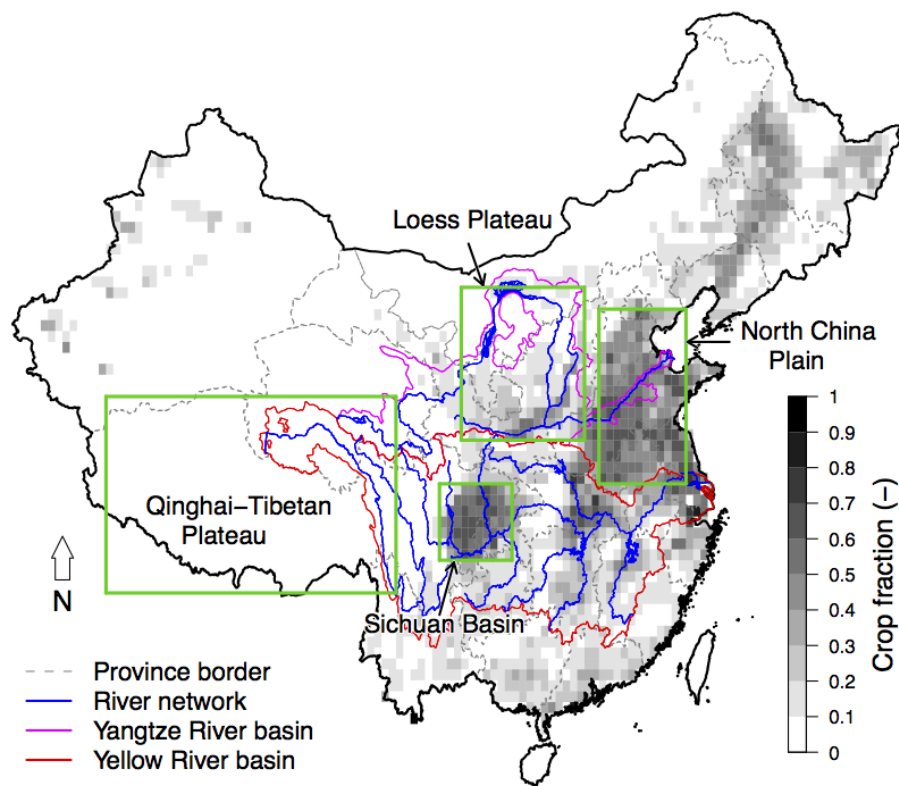


Figure 1. Map of China. The grey background is cropland fraction. Green rectangulars show four important regions mentioned in this paper.

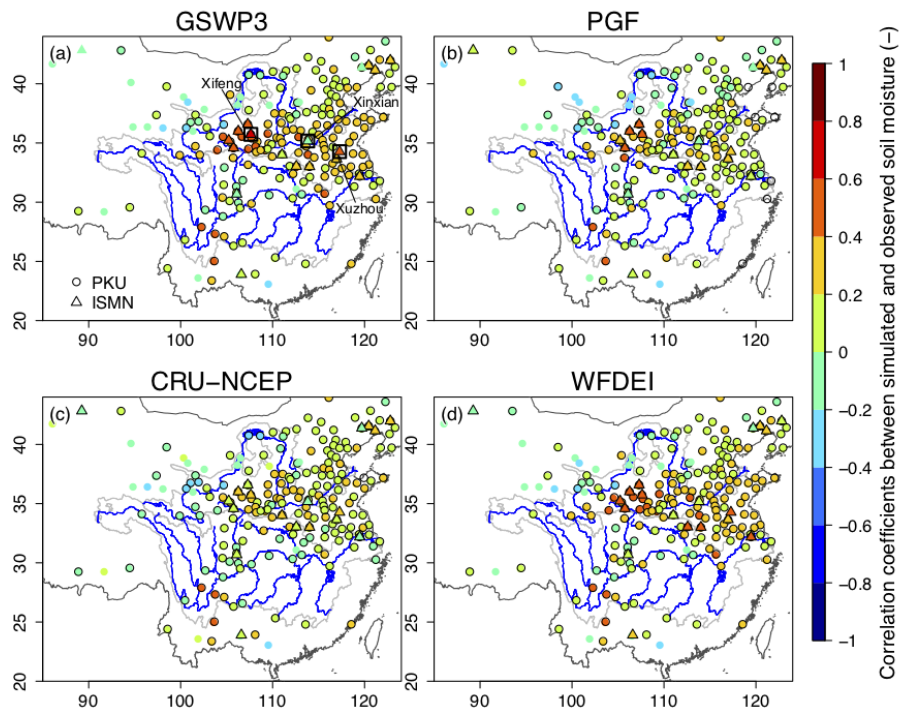


Figure 2. Pearson correlation coefficients of modeled and measured SM at each gauging station from ISMN (triangles) and PKU (circles). Symbols with dark border indicate significant correlations ($p < 0.05$). The locations of three ISMN stations shown in Figure 3 are marked by black squares in panel (a).

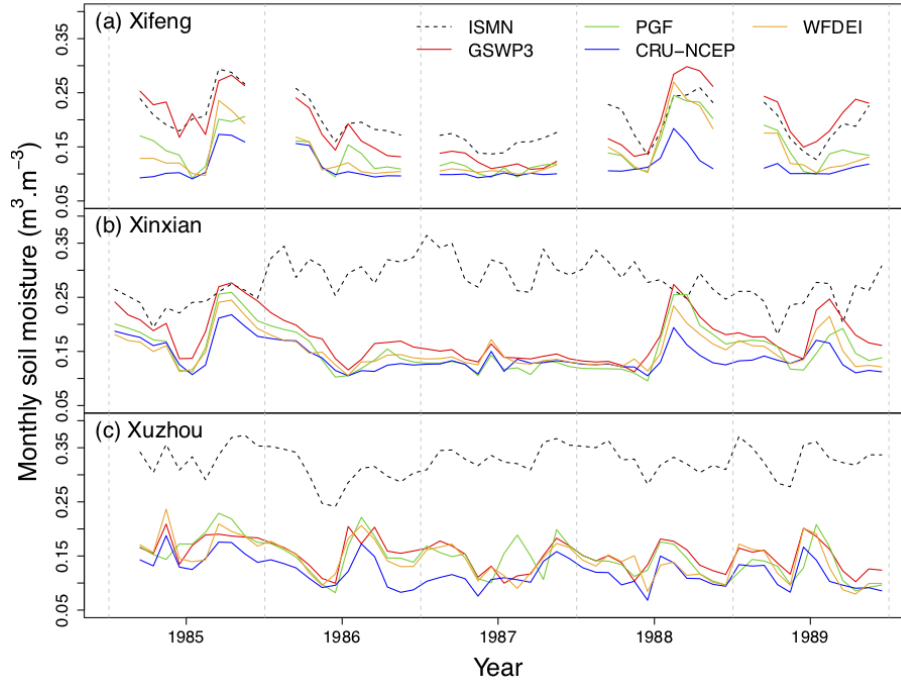


Figure 3. Time series of 10-day SM from ORCHIDEE and ISMN at three stations. The station locations are shown in Fig. 2(a). The mean annual precipitation at Xifeng, Xinxian, and Xuzhou (according to GSWP3) are 556, 580, and 847 $\text{mm} \cdot \text{yr}^{-1}$, respectively. Dark dashed lines indicate ISMN SM. Red, green, blue, and orange lines indicate simulated SM based on GSWP3, PGF, CRU-NCEP, and WFDEI, respectively.

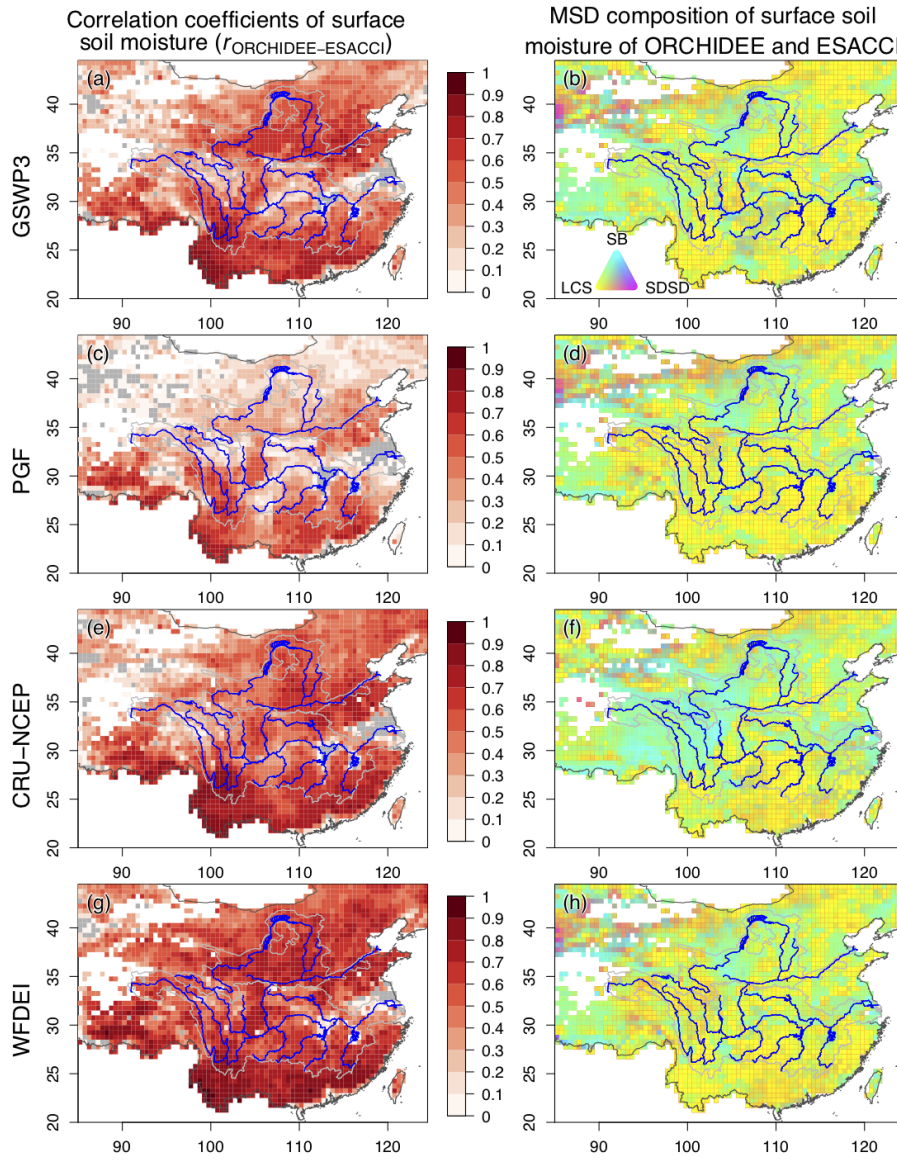


Figure 4. Left panel: Correlation coefficients of the ESA CCI SM and the corresponding ORCHIDEE SM. Gray pixels indicate non and negative correlation. Right panel: decomposition of the MSD between the daily ESA CCI SM and the corresponding ORCHIDEE SM (Eq. 1). Cyan, magenta, and yellow indicate the fractions of SB, SDSD, and LCS respectively.

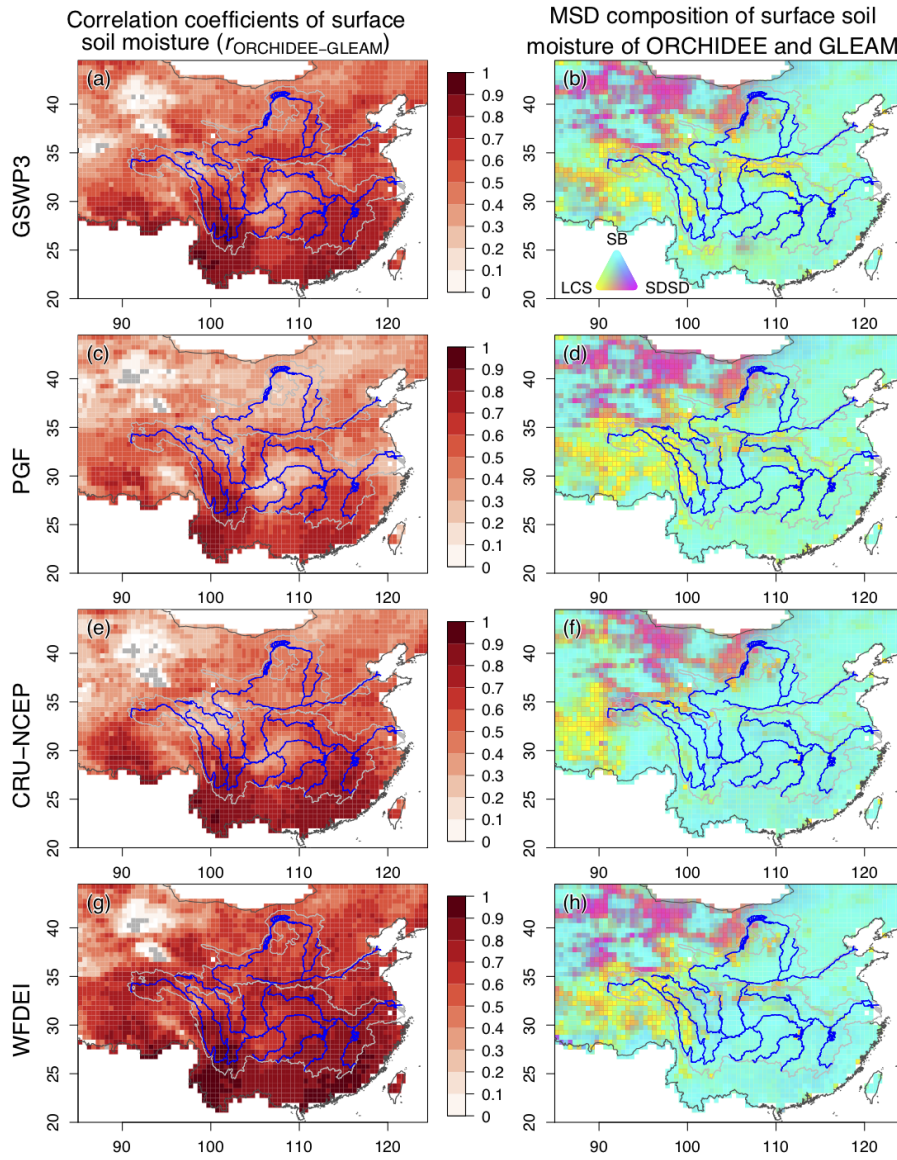


Figure 5. Left panel: Pearson correlation coefficients of the GLEAM surface SM and the corresponding ORCHIDEE SM. Gray indicates non and negative correlation. Right panel: decomposition of the MSD between the daily GLEAM surface SM and the corresponding ORCHIDEE SM (Eq. 1). Cyan, magenta, and yellow indicate the fractions of SB, SDSD, and LCS respectively.

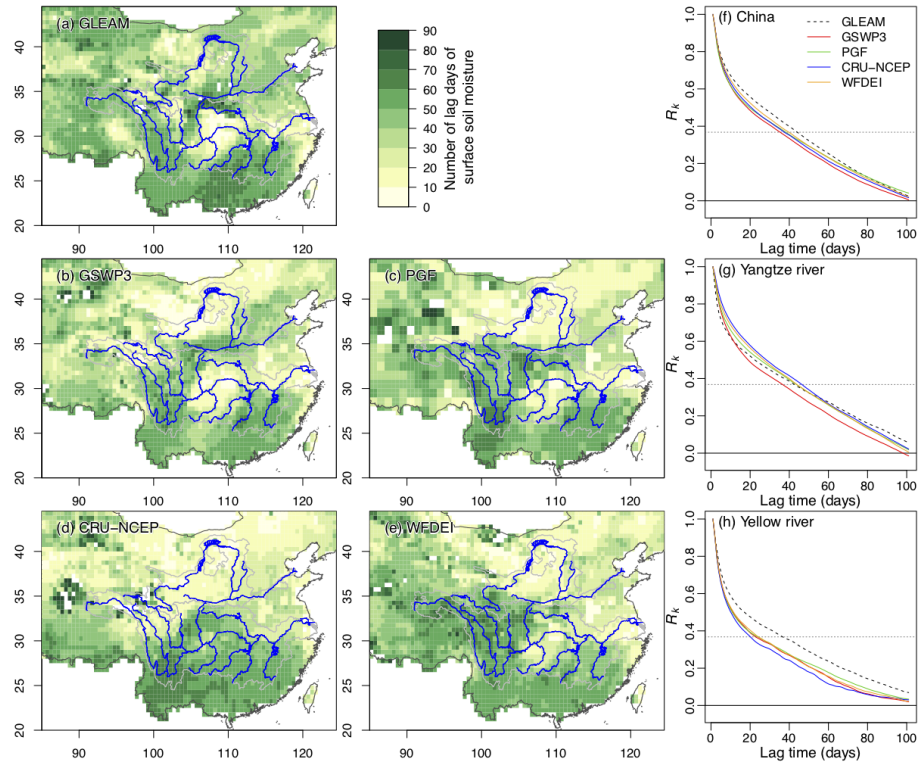


Figure 6. (a): Number of lag days (NLD) of GLEAM surface SM. (b)-(e): Difference of NLD between GLEAM and ORCHIDEE surface SM. (f)-(h): Autocorrelation coefficient R_k of spatial averaged surface SM as a function of NLD. The dashed line ($y = 1/e$) is the threshold of significant correlation.

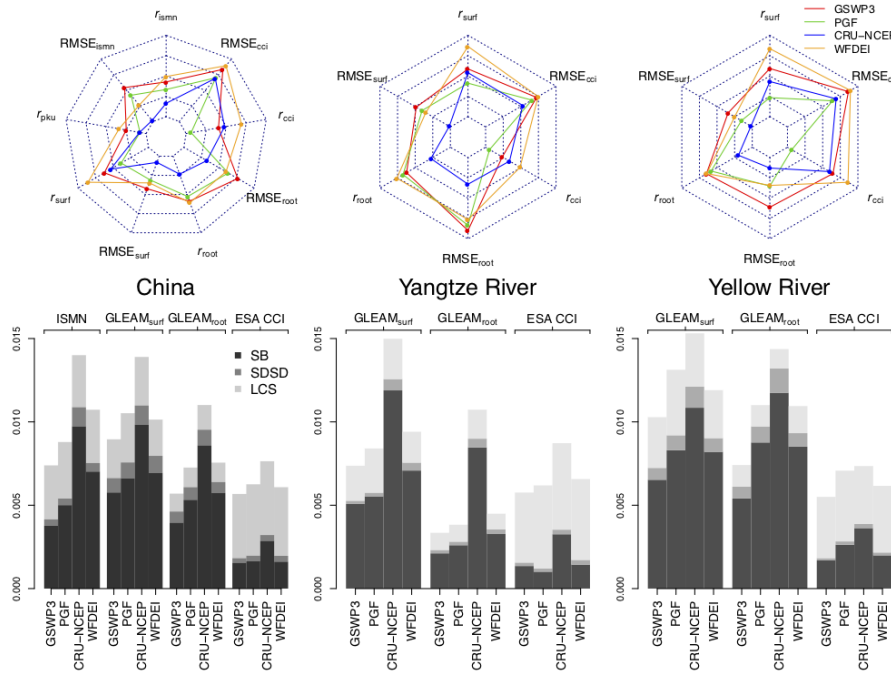


Figure 7. Evaluation of the forcing datasets for simulating SM dynamics in China, YZRB, and YLRB. Top panel: Radar charts of criteria of the four forcing datasets. Center implies bad criteria. Red, green, blue, and orange lines indicate GSWP3, PGF, CRU-NCEP, and WFDEI, respectively. ‘surf’ and ‘root’ indicates surface and root-zone SM of GLEAM 3.0A. Bottom panel: Composition of MSD from each comparison. x -axis indicates the drivers of specific simulations; top labels indicate the data set used in the specific comparison.

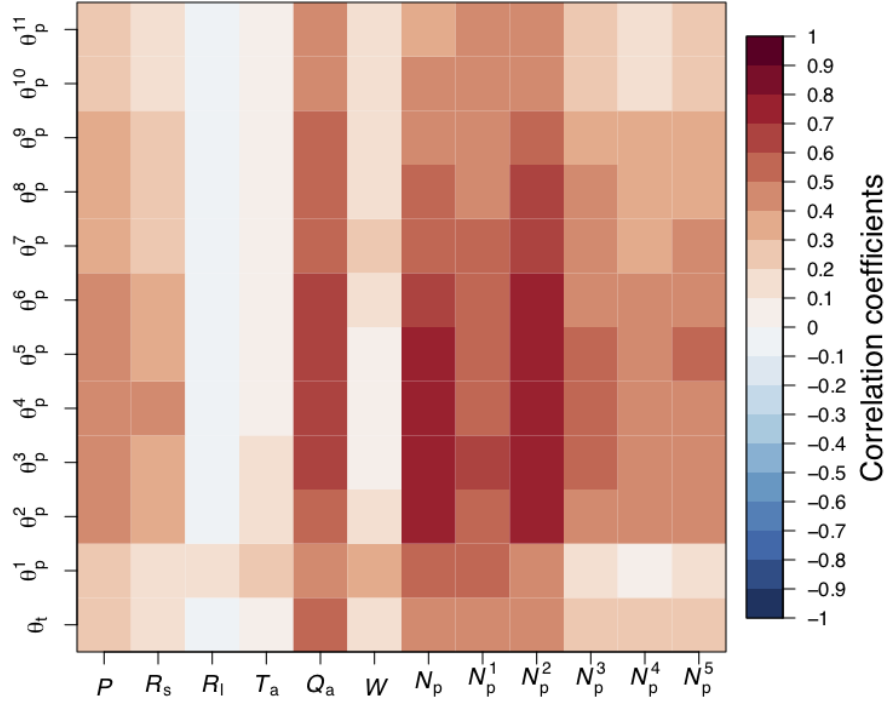


Figure 8. Matrix of correlation coefficients between the D of meteorological variables and the D of simulated SM. D is the averaged MSD defined by Eq. 6. θ_t indicates total SM. θ_p^i indicates SM in i th layer. P indicates annual precipitation. R_s and R_l indicate short and long wave incoming radiation, respectively. T_a indicates air temperature. Q_a indicates air humidity. W indicates wind speed. N_p indicates the number of days with precipitation no less than 0.01 mm.d^{-1} . N_{p^i} indicates the number of days with a specific precipitation range (Sect. 3.4).