Evaluation of soil moisture in CMIP6 multi-model simulations over conterminous China

Aihui Wang[1*], Xianghui Kong[1], Yue Chen[2], and Xin Ma[1]

1. Nansen-Zhu International Research Centre, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing China

2. College of Atmospheric Sciences, Lanzhou University, Lanzhou, Gansu, China

*Corresponding author: Aihui Wang (wangaihui@mail.iap.ac.cn)

Key points:

- Station observation and a pre-derived dataset in China are used to assess near-surface and root-zone soil moisture in CMIP6 simulations.
- The soil moisture simulations are similar for models in the same institutions, most models reproduce the recent drying trend in China.
- The multi model mean performs overall superior to individual model, which is suggested for future applications.

Abstract

The Sixth Phase of the Coupled Model Intercomparison Project (CMIP6) provides the long-term soil moisture (SM) products and this study conducts a comprehensive assessment of SM products of multiple CMIP6 model simulations over conterminous China. Both near-surface (0-10 cm) SM simulations from 40 models and root-zone (0-100 cm) SM from 25 models are compared with a set of station measurements in the growing season (April to September) for 1992-2013 in term of magnitude, spatial and temporal variability, and the long-term trend and interannual variability of near-surface SM for 1961-2014 are further evaluated with an offline land surface modeling dataset. Simulations from most models broadly capture the spatial characteristics of observation and the multi-model mean (MME) well reproduces seasonal variations over majority regions regardless of large-spread across models. Models from the same institution likely manifest similar performances and the land surface scheme plays a dominant role in the SM reproduction. The majority of models well simulate the overall drying trend in China as a whole and the signs of SM trend are highly consistent across models, but the areas with significant wetting/drying trends vary with models. The spatial patterns of SM interannual variability are model-dependent and the result of MME relatively resembles to the reference dataset. In summary, MME is overall superior to the simulation of individual model and may have potential applications in the future research. The heterogeneity SM performances across models reveal the complexity in modeling land surface variables, suggesting the need for improving representations of land surface processes in the coupled models.

Plain Language Summary

Soil moisture influences the exchange of energy, water and carbon within the soil-vegetation-atmosphere system. Climate models are extensively applied to simulate soil moisture data that are required to be assessed prior to applications. In this study, we comprehensively evaluate the soil moisture simulations in the Sixth Phase of the Coupled Model Intercomparison Project (CMIP6) historical experiments, with a set of station observations and a high-quality pre-derived SM product in China as the reference datasets. The results show that the spatial and temporal characteristics of soil moisture simulated by majority models are consistent with observation and the multi-model mean (MME) well reproduces SM seasonal variations in most regions. Resemble performances are found for models from the same institution. For China as a whole during 1961-2014, the drying trend is reproduced by most models. There are high consistencies of trend patterns across country among models although the magnitudes of trend are model-dependent. The interannual variability of soil moisture varies with models in term of spatial patterns, for which MME overall shows better performance than the simulation of individual model. This study provides a basic reference for selecting soil moisture products in CMIP6 for future applications.

1 Introduction

Soil moisture (SM) is one of key elements in the terrestrial hydrology and ecosystem and it determines exchange of energy, water and carbon within the soil-vegetation-atmosphere system. The role of SM in the climate system has been extensively studied (e.g., Seneviratne et al., 2010) and it has been used in various applications including drought reconstruction and identification (e.g., Wang et al., 2011; Wang & Kong, 2021), hydrological forecast (e.g., Koster et al., 2014), short-term climate prediction (e.g., Yeh et al., 1984), climate variability (e.g., Yang et al., 2016), dust emission (e.g., Kim & Choi, 2015) and wildfire behavior prediction (e.g., Krueger et al., 2016). For above applications, the long-term reliable SM datasets with continuously spatial coverage are crucial.

There are several ways to obtain SM dataset and each of them has strengths and weaknesses. Ground-based measurements provide the most accurate SM records, but they are less spatial coverage and temporal continuity (Dorigo et al., 2011; Robock et al., 2000; Wang & Shi, 2019). Remote-sensing based SM datasets may extend longer term with broad coverage, whereas they only represent moisture state within a few centimeters below ground and their accuracies are impacted by many factors such as the retrieval approach and frequent and coverage of satellite measurements (Liu et al., 2011). Numerical models are widely used to produce SM dataset for decades. For instance, land surface model (LSMs) forced by observation-based meteorology datasets can offline simulate the land surface hydrology elements and fluxes over regional and global land areas, but the accuracy of simulations strongly depends on the model physical process, parametrization schemes and forcing dataset (Miao & Wang, 2020; Wang et al., 2016; Xia et al., 2014).

The Six phase of Coupled Model Intercomparison Project (CMIP6) provides the collection of climate and land datasets derived from Earth System Models (ESMs) for historical simulations and future climate projections (Eyring et al., 2016). Products of CMIP6 have been and will support many aspects not only in research communities, but also in the social economic policy-makings in the next few years. For instance, the CMIP6 SM products have been used to project drought in the future under different SSP scenarios (Cook et al., 2020). There is a common consensus that better performances of climate models in certain aspects during historical simulations would also be more reliable in future projection. Therefore, evaluation of CMIP6 model historical simulations with reliable datasets is very important and may provide a reference for model selection in future projection.

In CMIP6 models, SM is one of the prognostic variables in the land surface model (i.e., LSM) that is an essential component of ESM. The performances of SM in ESMs depend on many aspects, primarily including representation of model physical processes, coupling schemes, as well as the external forcing imposed on the CMIP6 historical experiments. Among them, LSM is crucial for SM production and its physical processes and corresponding computation schemes directly determine the accuracy of SM (Wang et al., 2016). Previous evaluations of CMIP simulations mainly focus on the near-surface SM (e.g., SM within top 10 cm soil layer). For example, compared to its predecessor (i.e., CMIP5), Yuan et al. (2021) indicated that surface SM in CMIP6 better captures its trends in term of the whole contiguous United States and the multi model agreements are also enhanced. Sang et al. (2021) also compared the 0-10 cm SM of BCC-CSM simulations and reported overall improvements of SM in CMIP6 against CMIP5. There are two reasons: i) it is one of standard output variables in CMIPs and all endorsed models are required to provide this variable; and ii) the remote sensed SM that represents within 2~5cm soil depth can be directly used to be compared with simulations in CMIPs. However, the root zone SM (roughly denoted as SM in 0-100 cm soil layer) is essentially important for climate prediction and ecosystem production. Near-surface SM has large temporal variability because it is directly affected by atmosphere process, while root-zone SM may have season to intra-seasonal time scales (Nie et al., 2008). Although some studies showed that there is a strong relationship of SM between near-surface and root-zone (Ford et al., 2014; Mahmood & Hubbard, 2007), model performances of both SMs are not necessarily consistent (e.g., Yuan & Quiring, 2017). Berg et al. (2017) also reported the divergent projections of SM in the vertical soil profile under global warming. Therefore, to fully understand the SM from multi model simulations in CMIP6, it necessarily performs evaluation both near-surface and root-zone SM.

In this study, we conduct a comprehensive assessment of SM products from 40 CMIP6 models over the conterminous China in terms of their magnitude, spatial and temporal variation, long-term variabilities. A set of high-quality station-based SM dataset and a well-validated LSM simulated long-term SM product are used as the reference datasets. We compare both near-surface and root-zone

SM at individual sites, regional scale and country scale. Results of this evaluation may provide guidelines for both model output users and modeling groups to understand their SM simulations among others. Section 2 describes SM simulations from CMIP6 models, reference datasets and methods. Sections 3 and 4 present the detailed comparisons against two reference datasets, respectively. Section 5 discusses possible sources of uncertainties in CMIP6 SM simulations and reference datasets, and conclusions of this study are summarized in Section 6.

2. Data and method

2.1 CMIP6 SM datasets

In this study, we evaluate 0-10 cm and 0-100 cm SM, representing near-surface and root-zone SM, respectively. We downloaded the outputs of "mrsos" and "mrsol" from historical (1850–2014) simulations that are the core DECK experiments from the data site of CMIP6 (https://esgf-node.llnl.gov/search/cmip6/). In CMIP6 variable name system, "mrsos" is referred to as "the mass of water in the upper 0-10 cm of the soil layer", and "mrsol" is "the mass of water in each soil layer". At the time of this study, there are 40 CMIP6 models with "mrsos" available (the first ensemble, r1i1p1) and 25 models provide the "mrsol" and soil layer thickness dataset. Information of each model including resolution, institution, and land module and its vertical layer are summarized in Table 1. Of all 40 CMIP6 models, 16 LSMs are adopted and the total soil column depth varies from 1.9 to 35.2 meters and the number of soil levels varies 5 to 20 (Table 1). The root-zone SM is computed as the integrated "mrsol" within 0-100 cm weighted by the soil layer thicknesses in each model. Several ESMs share identical LSMs, in particular for models hosted by the same institution. For instance, the community land model (CLM) with different versions has been used as the land module in 13 CMIP6 models. Among them, eight models (i.e., FGOALS-f3-L, FGOALS-g3, FIO-ESM-2-0, NorCPM1, NorESM2-LM, NorESM2-MM, SAM0-UNICON, TaiESM1) adopt CLM4 as the land module, while two CESM2 models use CLM5 and three CMCC-models use CLM4.5. Differences among those versions of CLM have been well documented (Lawrence et al., 2019). Intercomparisons of SM from different ESMs with the same LSM may exclude the impact of land process parameterizations and facilitate exploring the influences of other model components on SM simulations.

2.2 Reference SM datasets

a. Station observation

We adopted an in-situ measured SM dataset in China as a reference (Wang & Shi, 2019, hereafter referred as to WS2019). This dataset was originally derived from the gravimetric method at 1471 Chinese agrometeorology stations and then after strictly quality-controlled and post-processed, the WS2019 SM dataset contains the monthly volumetric SM records for up to 732 stations in five soil layers (Figure 1). The gravimetric method is regarded as the most accurate approach to measure total moisture in the soil sample (Lekshmi et al., 2014).

Thus, those measurements are widely used as the benchmark for other datasets. For example, WS2019 SM has been used as a reference to assess the climate model and LSM simulations (Miao & Wang, 2020; Sang et al., 2021; Wang & Kong, 2021). In this study, WS2019 SM at 0-10cm (732 stations) and 0-100 cm (117 stations) during growing seasons (April to September) for 1992-2013 are used to evaluate CMIP6 simulations.

b. Offline VIC model simulations

A common deficiency of SM station measurements, particularly based on the gravimetric method, is the lack of continuity in both space and time. Most stations in WS2019 locate in the east part of China and few of them are in the western China, especially the northwest and the Tibetan Plateau (Figure 1). Moreover, in-situ measurements in cold months are usually not available or contain too many missing values due to soil frozen and snow cover. To evaluate the long-term variability of CMIP6 SM over the whole China domain, it is necessary to have a reliable long-term SM reference dataset. Using the advanced Variable Infiltration Model (VIC, Liang et al., 1994) model and a pure observation-based atmospheric forcing dataset, Miao & Wang (2020) conducted an offline modeling experiment at 0.25°×0.25° horizontal resolution during1961-2017 and then produced a set of land hydrology and fluxes variables (referred as to VIC-CN05). Compared with various sources of datasets, the VIC-CN05 well reproduced land surface hydrology and fluxes in term of their magnitude and spatiotemporal variability. Compared with WS2019 and other independent datasets, the VIC-CN05 SM displayed high performance in SM spatiotemporal characteristics. The VIC-CN05 contains three layers of SM datasets with the first layer given as 0-10 cm below ground, and the second and the third layers vary between 50 cm to 250 cm across China. In this study, the 0-10 cm VIC-CN05 SM for 1961-2014 is adopted to evaluate CMIP6 SM in terms of their annual trend and interannual variability across the whole China mainland.

2.3 Method

To facilitate comparisons, all monthly CMIP6 SM values are firstly remapped to 0.25°×0.25° using the near-neighbor interpolation method to match the CN05-VIC resolution. In order to compare model simulations with WS2019, the near-neighbor interpolation method is employed again to interpolate each model simulation to the station location. Because the station measurements are strongly influenced by local land surface conditions, regional average may reduce the noises of point observation and highlight the distinct characteristics of SM. For both observation and model simulations, the regional mean is computed as the average of values at all available stations with each subregion (Figure 1a). The comparisons are conducted at both point sites and subregions for the growing season during 1992-2013.

In the following texts, SM from both station observation and CN05-VIC are denoted as "OBS" for convenience. Results from individual model and multi-model ensemble mean (MME), derived from 40 models for 0-10 cm and 25

models for 0-100 cm SM, are compared with OBS. To quantify the performance of simulated SM, we also compute and compare both temporal and pattern correlation coefficients between simulations and OBS, and the long-term linear trend and interannual variability of SM. The linear trend is computed as the least-square regression method and the interannual variability is represented by the standard deviation (STD) of annual data.

3. Results

3.1 Spatial variation at point-to-point sites

In this section, we evaluate CMIP6 multi-model simulation with the in-situ observation at each station and each subregion. To perform the meaningful comparison and also consider the observation data availability (Wang & Shi, 2019), the annual mean SM are computed from April to September in each year firstly, and then they are averaged for the period of 1992-2013. Figure 1 shows the spatial distribution of SM from MME and OBS. Similar plots from individual CMIP6 model are shown in Figure S1 and Figure S2. In both soil layers, MME broadly captures the spatial pattern of OBS, gradually wetting from north to south. As shown, station-to-station variations in MME are smaller than that in OBS, partially due to MME offsetting the variabilities across models (Figure S1). Because station measured SM values are derived from the soil sample at the point site, it represents a small-scale soil hydrological condition. The magnitude of SM is strongly affected by the soil characteristic, vegetation cover and other local factors, which are very heterogeneous over large areas (Entin et al., 2000; Wu et al., 2014). Intercomparisons of simulations from individual model, the spatial patterns of mean SM show apparently disparity across models (Figure S1). For example, the KACE-1-0-G simulates very extreme dry soil in East China where the climate is relatively humid, while some other models, e.g., NorESM2 and CESM2, produce relatively wetting soil. Because soil variables in near ground are directly interacted with overlying atmosphere process, the near-surface SM is to large extent reflection of near surface climate state, in particular, precipitation. The large biases of surface SM somehow imply the poor simulations of precipitation in the models (e.g., Dirmeyer, et al., 2006; Li, et al., 2020; Wang & Zeng, 2011).

We also adopt boxplots to display the SM distribution across all available stations from individual model, MME and OBS (Figure 2). It can be seen that SM for 10th, 25th, mean,75th, and 90th spatial percentiles across all available stations are also indicated in the figure. For near-surface soil layer, the mean SM from MME (0.26 cm^3/cm^3) are slightly larger than OBS (0.24 cm^3/cm^3), while 10^{th}-90^{th} spatial ranges of SM from MME (0.19 ~0.30 cm^3/cm^3) are slightly smaller than OBS (0.16~0.32 cm^3/cm^3). The mean SM from individual model varies between 0.16 cm^3/cm^3 (KACE-1-0-G) and 0.32 cm^3/cm^3 (NorESM2-LM). For the mean root-zone SM, the simulation of MME (0.26 cm^3/cm^3) is very closer to OBS (0.27 cm^3/cm^3). Above results indicate that MMEs in both soil layers well reproduce the mean SM for China as whole. Meanwhile, we also find that ESMs from the same institution show similar performance. For example,

SM from all five EC-earth ESMs do not show large differences in terms of mean and spread, and three GISS ESMs and two CESM ESMs respectively simulate nearly identical SM. Both CESM2 and CESM2-WACCM use the same modeling framework, but their atmospheric models have different vertical representations, with the latter one spanning much high altitude from the Earth's surface to the thermosphere (Marsh et al., 2013). This suggests that the atmospheric vertical structures do not affect simulations of land surface hydrology process because they have less effects on precipitation simulations (Marsh et al., 2013). For those CMIP6 models sharing the same LSMs but hosted by different institutions, their simulated SM show relatively small differences. Whereas there are large differences for ESMs from different institutions. The above results indicate that the model horizontal resolutions or slightly different atmosphere schemes are likely have little effect on the mean SM magnitude and spatial spread.

To further quantify spatial variability of mean SM across stations from each model simulation and MME, we compute the pattern-correlation (CORR) and normalized spatial standard derivation (NorSTD) against OBS. The Taylor diagram is adopted to display CORR and NorSTD (Figure 3). As seen clearly, simulations from all models in both soil layers show positive CORR, indicating overall reproduction skills of SM spatial patterns in CMIP6 models. For near-surface SM, the CORR varies between 0.10 (NorCPM1) and 0.51(EC-Eath3-Veg and EC-Earth-CC) across 40 models, and it is 0.46 for MME that is larger than simulations of most models. For root-zone SM, CORRs are overall larger than those from the near-surface layer, varying between 0.22 (NorESM2-LM and MIROC6) and 0.52 (EC-Eath3-Veg and EC-Earth-CC) across 25 models and being 0.49 in MME. The NorSTDs of SM across all stations represent the spatial variability of model simulations in comparison with OBS. It denotes the perfect matching of simulation and OBS when NorSTD is equal to 1, The NorSTD varies between 0.54 cm^3/cm^3 (ACCESS-ESM1-5) to 1.31 cm^3/cm^3 (CAS-ESM-0) for the near-surface SM, and between 0.44 cm^3/cm^3 (ACCESS-ESM1-5) to 1.04 cm^3/cm^3 (MPI_ESM1-1-HR) for root-zone SM. Especially, NorSTD in three models (i.e., GFDL-CM4 and KACE-1-0-G for near-surface SM, EC-Eath3-Veg for root-zone SM) are very close to one, indicating high skill in reproduction of SM spatial variation. NorSTDs of MME for both soil layers are less than one (0.66 for near-surface SM and 0.58 for root-zone SM), indicating the underestimation of spatial gradient of SM across stations in China. Consistent with the results in Figure 2, similar magnitudes of CORR and NorSTD are also found in models from the same institution. For instance, regarding to results from five EC-Earth3 ESMs for 0-10 cm and 0-100 cm SM, the CORR is 0.48~0.50 and 0.48~0.52, while NorSTD is 1.21~1.28 cm^3/cm^3 and 0.63~0.64 cm^3/cm^3 respectively. This result further emphasizes the strong dependence of SM simulations on the LSM implemented in the climate models.

3.2 Regional comparison

Figures 4 and 5 receptively show the monthly climatology of near-surface and root-zone SM from OBS, individual model and the MME averaged in six subre-

gions for the period of 1992-2013. As seen clearly, the SM seasonal variations of MME and OBS are well matched in five of six subregions (except for HY) although there are large spreads across models. In XJ, specifically, all models produce apparently drier soil compared to OBS, indicating a systematical underestimation of SM. This underestimation is also evident for root-zone SM (Figure 5). In NW and CE, MMEs are wetter than OBS in both layers. In YH, the OBS SM slightly increases from April to September, but MME shows oppositely variations. Because the climate of this region is strongly impacted by the East Asian Monsoon System as well as other larger scale circulation teleconnections (Xiao et al., 2015), climate prediction and simulation are extremely difficult (Si & Ding, 2011). For example, Wang & Kong (2021) assessed the SM products from the Weather Forecast Model (WRF) experiments driven by two reanalysis products and also found a negative correlation against OBS in this region.

To examine the SM temporal variability of model simulations, the temporal correlation (CORR) between individual model and OBS, between MME and OBS are computed from the available monthly data in each subregion (Figure 6). For near-surface SM, simulations of majority models and MME are positively correlated with OBS, and CORRs are significant at the 95% level in five subregions except for YH, where CORR is negative or very small positive among models. It is surprising that the CORRs in XJ are significant in all models and also highest among all subregions. For root-zone SM, CORRs are slightly smaller than those in near-surface layer and only simulations from a few models and MME in NW and CE are significantly correlated with OBS. In YH, similar as near-surface layer, most models produce negative CORRs. The possible reason of that has been discussed preciously. Therefore, it is suggested that caution should be paid when using CMIP6 SM in YH subregion. Regarding the SM modelling deficiency in this region, it is needed to have in-depth research.

4. Evaluation of long-term variability of soil moisture

In this section, we evaluate the long-term variability of near-surface SM from individual CMIP6 model, MME and OBS for the period of 1961-2014. The long-term linear trend for annual SM and the interannual standard deviation (STD) of detrended SM at each grid cell are computed and compared. Noted that the SM unit of plots in this section use millimeter (i.e., mm) for visual purpose and it can be easily converted to the volumetric unit (i.e., cm^3/cm^3).

4.1 Annual trends

Figures 7 a) and b) respectively show the maps of annual SM trend from OBS and MME. Similar results from individual model are in Figure S4. Significant wetting trends of OBS appear in the northwest, Three-rivers source region, northeast China, but there are offset by significant drying trends in the eastern Gansu along the middle reach Yellow River basin, eastern Tibetan Plateau and southwest China. In contrast, the significant wetting trends of MME occur in the western of 105°E and lower reaches of Yellow River basin, but a large portion of the rest of land area displays significant downtrend, patterns

of which are roughly consistent with OBS except for the Xinjiang-Tibet border along Kunlun Mountain, where trends show opposite signs. Consequently, the country-mean SM shows a slightly increasing (wetting) trend from both OBS (0.54×10^2 mm/year) and MME (0.26×10^2 mm/year). Intercomparison of SM trend simulations from individual model, the most distinct biases against OBS are that several models (e.g., NorESM2-LM, NorESM2-MM, CESM2, CEM2-WACCM, and five EC-Earth3 models) produce strongly wetting trend (>0.04 mm/year) in the western Tibetan Plateau that induces a wetting center in MME. This disparity might be partially resulted from poor representations of land surface processes in the complex topography regions and the near-surface climate simulations also contains lager uncertainties in those regions (Wang et al., 2018).

In order to explore the inter-consistency of SM trend sign across models, we also compute the percentage of models with the same trend sign (negative or positive) as the OBS in each grid cell (Figure 7c). The average consistency across conterminous China is near 56% and areas with relatively high consistency appear in Northwest, the lower reach of Yellow River basin, and Yunan province, where more than 80% (i.e., at least 32 models) of the models show consistent SM trend sign. In contrast, the lowest consistency locates at the south of lower reach Yangtze River basin where less than 30% of the models agree on the sign of SM trend. One purpose of the CMIP6 historical experiments is to reproduce the reliable variables that may support to exploring the long-term variability in climate, hydrology and other related research (Eyring et al., 2016). High consistency denotes the robustness and the reliability in representation of long-term SM trend within the CMIP6 multi-model ensemble. Similar analysis may also be applied to assess other modeled variables.

Figure 7d) shows the aggregation of land areas with positive or negative SM trends that significantly exceed the 95% confidence level. The significant wetting areas are much smaller than that of drying from OBS, indicating overall expansion of dryland in China (Cheng et al., 2015; Wang et al., 2011; Wang & Kong, 2021). The MME captures this feature, although either significantly wetting or drying areas are more extensively compared to those in OBS. There are nearly 70% of land areas that show significant SM trends in MME, which are much larger than that from individual models as well as OBS (25%). For individual model, majority of models produce more drying areas than wetting areas, although the magnitude of above metric varies with models. This may imply certain ability of ESMs on reproducing SM trend pattern in China. It should be noted that models with the same LSM do not necessarily produce similar SM trend patterns.

4.2 Interannual variability

Figure 8 shows the spatial distribution of interannual variability of near-surface SM from OBS and MME, represented by the interannual STD of detrended annual mean SM for 1961-2014. The results from each individual model are listed in Figure S5. The STD varies spatially and it is lowest (less than 0.4 mm) in

northwest regions, where MME is consistent with OBS (Figure 8a). A noticeable mismatch between MME and OBS appears in northeast China and the Tibetan Plateau, which is resulted from the diverse performance of several models (e.g., NorESM2_MM, NorESM2_MM, CESM2, CENS2_WACCM, CanESM5 and KACE-1-0-G). Over the Tibetan Plateau, high STD in MME may be also induced by the uncertainties of parameterization schemes of LSM used in models. To further quantify the spatial patterns of interannual variability, we also calculate the centered pattern correlations (CORR) of STD from each individual model and MME against OBS (Figure 8c). CORRs from all models and MME are positive, indicating overall consistency of simulations and OBS in terms of spatial pattern of STD. It is noticeable that correlations derived from five EC-Earth models are relatively larger, especially CORR of EC-Earth3-Veg (0.51) is even larger than that of MME (0.48). Similar to Figure 7d, models with the same LSMs do not produce coherently patterns of STD. The long-term variability of near surface SM mainly depends on the atmospheric forcing dataset but is less affected by land surface scheme (Cheng et al., 2015; Wang et al., 2016). The magnitude of STD denotes the year-to-year variability of near-surface SM, which mainly reflects the annual variation of near surface climate, in particular precipitation (Liu et al., 2017). This needs to explore the model simulated near surface climate variables, which are beyond the scope of current work.

5 Discussions

5.1 Uncertainties of soil moisture in the CMIP6 model simulations

As a prognostic variable in the model, SM is usually computed through solving land surface water balance equations in the climate models. The performance of SM simulations in climate models depends on many factors that may be roughly divided into two categories: near-surface atmosphere variables and parametrization schemes in LSM. Previous studies based on offline LSM experiments have indicated that the simulated SM magnitude has largely depended on the accuracy of input surface meteorology variables, in particular precipitation, while its temporal variability may be also affected by the parameterization schemes in the LSM (Wang & Yang, 2018; Decker &Zeng, 2009; Dirmeyer et al., 2006). Variations of near-surface moisture from monthly to season scale are generally followed by the near land surface climate, and it highly correlates with precipitation that arrives at the soil surface (Koster et al., 2009). Evaluation of near surface climate simulations from CMIP6 multi-model ensemble in China have been conducted and the results show overall improvements compared to its predecessors (i.e., CIMP5) resulted from the improvements of physical process representation and enhanced spatial resolution (e.g., Chen et al., 2020). Compared to the in-situ observations, the climate products from CMIP6 models still contain larger biases, which in turn may transfer to the land surface hydrological process.

The representations of land surface schemes are another important source of uncertainty in the simulated SM. Different LSMs, even driven by the same atmospheric forcing dataset, may also produce larger uncertainties in the land

surface variables (e.g., Dirmeyer et al., 2006; Wang et al., 2011; Xia et al., 2014). The comparison of station observation and SM simulations in current study shows the dependence on the LSMs (Figures 2 and 3). As 40 CMIP6 models employed 16 LSMs (Table 1 and Figure 2), the vertical structures of soil column and soil water transfer schemes are model-depended. The water transfer scheme regulates the water movement between adjacent soil layers and determines the soil water retention time in each layer. The vertical soil layer division within soil column also represents the vertical resolution of LSM. For instance, the total soil depth is about 2.9 meters in ACCESS and BCC models, but they are subdivided into 6 and 10 soil layers, respectively. The different vertical resolution may produce different variable SM profiles that are potential causes of simulated SM disparities among models. Furthermore, soil textures (i.e., sand, clay and other matters in the soil), which determine the soil hydrological parameters (e.g., soil hydraulic conductivity and soil matric potential), can directly impact the water transfer in the soil (Wu et al., 2014). In CMIP6 historical experiments, soil texture datasets are specified by each model group and might be another potential source of uncertainty in simulated SM. Furthermore, although all modeling groups in the CMIP6 historical experiments were required to use the "land use harmonization" version 2 (LUH2) as the land cover input (Hurtt et al., 2011), the detail settings might not be strictly the same for different models (Lawrence et al., 2016) and they may be also different from the real land surface conditions. All above settings in CMIP6 models might be sources of uncertainties in SM simulations.

5.2 Uncertainties in the reference datasets and the comparison methods

We used a set of volumetric SM records at multi stations to compare the CMIP6 multi-model simulations in terms of magnitude, spatial pattern and temporary variability at point sites. We then use a pre-produced offline hydrology model simulation product to compare the long-time variability of SM in contiguous China. The measured SM displays larger spatial variability because the soil water holding capacity and soil texture are heterogeneous from station to station (Wang & Shi, 2019). Although the point-based measurement may provide the most accurate SM in quantity, its spatial representation is much smaller than a grid cell area (i.e., 500~10000 km^2) in climate models (Entin et al., 2000; Famiglietti et al., 2008). Scale-mismatching induced uncertainties are always challengeable for model-data comparison. Comparisons conducted with regional average reduces observation noises across stations, but characteristics of regional SM strongly depend on the density and distribution of stations within each region (Xia et al., 2014; Yuan & Quiring, 2017). Upscaling sparse SM measurements from the point site to relatively large scale may enhance the robustness of the comparison (Crow et al., 2012), but this process requires additional information in the observation networks that do not exist currently in Chinese in-situ SM observation. Moreover, because the measured SM used in this study are from agrotechnological stations in China, there are inevitable that SM records at some stations are influenced by irrigation (Wang & Shi, 2019). It should be

pointed that current comparisons are only conducted during the growing seasons due to limited observation, and the performance of SM in CMIP6 models might be different in cold months, in which soil water may experience phase change owing to the freezing-thawing transition induced by soil temperature change (e.g., Guo et al., 2018).

Offline LSM modeling products, driven by the observation-based atmospheric dataset and parameters provide fully land surface hydrological variables that are commonly used as a reference dataset to validate the coupled model simulations. Much of research have been proved the great reliability of offline LSMs modelling in reproduction of land surface hydrological and energy variables (e.g., Miao & Wang, 2020; Wang et al., 2016; Xia et al., 2014). The offline LSM simulation is indeed a passive response of forced historical climate states provided by prescribed atmosphere forcing dataset. In other word, they do not account for the concurrent land-atmosphere interaction, which has been proved to play an important role in the climate variability and prediction within the climate system (Seneviratne et al., 2010; Van den Hurk et al., 2011). Strictly, the offline LSM simulation is a surrogate reference that may contain biases from both atmospheric forcing dataset and land parametrization schemes.

6 Summary

This study assesses the SM simulations in CMIP6 historical experiment over mainland of China. Both SM in 0-10 cm (40 models) and 0-100 cm (25 models) soil layers are compared with a set of in-situ station observation (OBS) in term of its magnitude, spatial and temporal variabilities. Then, the long-term trend and inter-annual variability of near-surface SM are evaluated with a prevalidated offline LSM modeling product. Comparisons of simulated SM with OBS reveal that a large range of variability and inconsistencies across CMIP6 models. In general, simulations of majority CMIP6 models can broadly capture spatial characteristics of mean SM. Consistent with OBS, MME shows increasing SM from northwest to southeast. The pattern correlations between CMIP6 simulations and OBS are all positive and the correlation of MME, compared with that of individual model, is relatively high. The spatial variability of simulated SM varies with models and the majority of them show slightly less variability than OBS. Results indicate that climate models from the same institution produce a similar performance in terms of SM magnitude and spatial patterns, which reveals the importance of LSM in the coupled CMIP6 models. The seasonal SM variations are well reproduced by MME in most of the subregions although a large multi-model spread exists. Especially, the near-surface SM in all models and MME show significant correlation against OBS in arid region, whereas most models underestimate the magnitude of SM, possibly resulting from the systematic biases in the simulated near surface meteorology variables, in particular, precipitation (e.g., Wang et al., 2016; Wang & Zeng, 2011). This also illustrates the need to improve atmosphere model in order to provide high quality near-surface atmospheric forcing dataset for LSMs.

For the long-term variability of near-surface SM during 1961-2014, MME shows

an overall drying trend in China as whole, consistent with OBS although the spatial feature varies with regions. Signs of trend in 40 models are highly consistent with the OBS across China mainland, and their spatial patterns with significant wetting or drying trend are model-depended. Moreover, the MME broadly captures the spatial features of interannual variability although there are diverse across models. In summary, the MME seems better than individual model in most cases and may provide a reasonable SM product for future applications. The heterogeneity performances among models reveal the complexity in modeling land surface processes within a coupled model. In order to reproduce reliable SM simulations, there needs to synthetically improve various components in the models, especially the representation of land-atmosphere exchange in water and energy and soil water transfer scheme in LSMs.

Acknowledgments

The work is jointly supported by the National Natural Science Fund for Distinguished Young Scholars (Grant No. 41925021), and the National Natural Science Foundation of China (Grant No. 41875106). We thank the support of the National Key Scientific and Technological Infrastructure project "Earth System Science Numerical Simulator Facility (EarthLab)". We also acknowledge the World Climate Research Programme (WCRP) through its Working Group on Coupled Modelling providing their model outputs. The in-situ station measurements of soil moisture in China are available from the lead author via email. The offline VIC modeling product in China can be obtained at ScienceDB website: http://www.doi.org/10.11922/sciencedb.00364.

References:

Berg, A., Sheffield, J., & Milly, P. C. D. (2017), Divergent surface and total soil moisture projections under global warming. *Geophysical Research Letters*, 44, 236–244. https://doi.org/10.1002/2016GL071921

Chen, H., Sun, J., Lin, W., & Xu, H. (2020). Comparison of CMIP6 and CMIP5 models in simulating climate extremes. *Science Bulletin*, *65*(17), 1415–1418. https://doi.org/10.1016/j.scib.2020.05.015

Cheng, S., Guan, X., Huang, J., Ji, F., & Guo, R. (2015). Long-term trend and variability of soil moisture over East Asia. *Journal of Geophysical Research: Atmospheres*, 120, 8658– 8670. doi:10.1002/2015JD023206

Cook, B. I., Mankin, J. S., Marvel, K.,Williams, A. P., Smerdon, J. E., & Anchukaitis, K. J. (2020). Twenty-first century drought projections in the CMIP6 forcing scenarios. *Earth's Future*, 8, e2019EF001461. https://doi.org/10.1029/2019EF001461

Crow, W. T., Berg, A. A., Cosh, M. H., Loew, A., Mohanty, B. P., Panciera, R., et al. (2012). Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products. *Reviews of Geophysics*, 50, RG2002. http://dx.doi.org/10.1029/2011RG000372

Decker, M. & Zeng, X.B. (2009). Impact of modified Richards equation on global soil moisture simulation in the community land model (CLM3.5). *Journal of Advances in Modeling Earth Systems, 1*(Art. 5), 22. https://doi.org/10.3894/James.2009.1.5

Dirmeyer, P. A., Gao, X., Zhao, M., Guo, Z., Oki, T., & Hanasaki, N. (2006). GSWP-2: Multimodel analysis and implications for our perception of the land surface. *Bulletin of the American Meteorological Society*, *87*(10), 1381–1398. https://doi.org/10.1175/BAMS-87-10-1381

Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., et al. (2011). The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements. *Hydrology and Earth System Sciences*, *15*(5), 1675–1698. https://doi.org/10.5194/hess-15-1675-2011

Entin, J. K., Robock, A., Vinnikov, K. Y., Hollinger, S. E., Liu, S., & Namkhai, A. (2000). Temporal and spatial scales of observed soil moisture variations in the extratropics. *Journal of Geophysical Research: Atmospheres*, *105*(D9), 11865–11877. https://doi.org/10.1029/2000JD900051

Eyring, V., Bony, S. Meehl, G. A. C. Senior, A. Stevens, B. Stouffer, R. J. & Taylor, K E. (2016), Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9, 1937−1958, https://doi.org/10.5194/gmd-9-1937-2016

Famiglietti, J. S., Ryu, D., Berg, A. A., Rodell, M., & Jackson, T. J. (2008). Field observations of soil moisture variability across scales. *Water Resources Research*, *44*(1). https://doi.org/10.1029/2006WR005804

Ford, T.W., Harris, E., & Quiring, S. M. (2014). Estimating root zone soil moisture using near-surface observations from SMOS. *Hydrology and Earth System Sciences,* 18, 139–154, https://doi.org/10.5194/hess-18-139-2014

Guo, D., Wang, A., Li, D., & Hua, W. (2018). Simulation of changes in the near-surface soil freeze/thaw cycle using CLM4.5 with four atmospheric forcing data sets. *Journal of Geophysical Research: Atmospheres,* 123. https://doi.org/10.1002/2017JD028097

Hurtt, G. C., Chini, L. P., Frolking, S., Betts, R. A., Feddema, J., Fischer, G., et al. (2011). Harmonization of land-use scenarios for the period 1500–2100: 600 years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands. *Climatic change*, *109*(1), 117–161. https://doi.org/10.1007/s10584-011-0153-2

Kim, H., & Choi, M. (2015). Impact of soil moisture on dust outbreaks in East Asia: Using satellite and assimilation data. *Geophysical Research Letters*, *42*(8), 2789–2796. https://doi.org/10.1002/2015GL063325

Koster, R. D., Guo, Z., Yang, R., Dirmeyer, P. A., Mitchell, K., & Puma, M. J. (2009). On the nature of soil moisture in land surface models. *Journal of Climate*, *22*(16), 4322–4335. https://doi.org/10.1175/2009JCLI2832.1

Koster, R. D., Walker, G. K., Mahanama, S. P., & Reichle, R. H. (2014). Soil moisture initialization error and subgrid variability of precipitation in seasonal streamflow forecasting. *Journal of Hydrometeorology*, *15*(1), 69–88. https://doi.org/10.1175/JHM-D-13-050.1

Krueger, E. S., Ochsner, T. E., Carlson, J. D., Engle, D. M., Twidwell, D., & Fuhlendorf, S. D. (2016). Concurrent and antecedent soil moisture relate positively or negatively to probability of large wildfires depending on season. *International Journal of Wildland Fire*, *25*(6), 657–668. https://doi.org/10.1071/WF15104

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., et al. (2019). The Community Land Model version 5: Description of new features, benchmarking, and impact of forcing uncertainty. *Journal of Advance Modeling Earth System*, 11, 4245–4287. https://doi.org/10.1029/2018MS001583

Lekshmi S.U., Singh, D. N., & Baghini, M. S. (2014). A critical review of soil moisture measurement. *Measurement*, 54, 92–105. https://doi.org/10.1016/j.measurement.2014.04.007

Li, M., Wu, P., & Ma, Z. (2020). A comprehensive evaluation of soil moisture and soil temperature from third-generation atmospheric and land reanalysis data sets. *International Journal of Climatology*, 40, 5744-5766. https://doi.org/10.1002/joc.6549

Liang, X., Lettenmaier, D.P., Wood, E.F., Burges, S.J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Meteorological Research*, *99* (D7), 14415–14428. https://doi.org/10.1029/94jd00483

Liu, D., Yu, Z., & Mishra, A. (2017). Evaluation of soil moisture-precipitation feedback at different time scales over Asia. *International Journal of Climatology, 37*(9), 3619-3629. https://doi.org/10.1002/joc.4943

Liu,Y., Parinussa, R., Dorigo, W., De Jeu, R., Wagner, W.,van Dijk, A., McCabe, M.,& Evans, J. (2011). Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. *Hydrology and Earth System Sciences*, 15, 425–436. https://doi.org/10.5194/hess-15-425-2011

Mahmood, R. & Hubbard, K. G. (2007). Relationship between soil moisture of near surface and multiple depths of the root zone under heterogeneous land uses and varying hydroclimatic conditions. *Hydrological Processes*, 21, 3449–3462. https://doi.org/10.1002/hyp.6578

Marsh, D. R., Mills, M. J., Kinnison, D. E., Lamarque, J. F., Calvo, N., & Polvani, L. M. (2013). Climate change from 1850 to 2005 simulated in CESM1 (WACCM). *Journal of climate*, *26*(19), 7372–7391. https://doi.org/10.1175/JCLI-D-12-00558.1

Miao Y., & Wang, A. (2020). A daily 0.25° × 0.25° hydrologically based land surface flux dataset for conterminous China, 1961–2017. *Journal of Hydrology*, 590, 125413, https://doi.org/10.1016/j.jhydrol.2020.125413

Nie, S., Luo, Y., & Zhu, J. (2008). Trends and scales of observed soil moisture variations in China. *Advances in Atmospheric Sciences*, *25*(1), 43–58. https://doi.org/10.1007/s00376-008-0043-3

Robock, A., Vinnikov, K. Y., Srinivasan, G., Entin, J. K., Hollinger, S. E., Speranskaya, N. A., et al. (2000). The global soil moisture data bank. *Bulletin of the American Meteorological Society*, *81*(6), 1281–1300. https://doi.org/10.1175/1520-0477(2000)081%3c1281:TGSMDB%3e2.3.CO;2

Sang, Y., Ren, H. L., Shi, X., Xu, X., & Chen, H. (2021). Improvement of soil moisture simulation in Eurasia by the Beijing Climate Center Climate System Model from CMIP5 to CMIP6. *Advances in Atmospheric Sciences*, *38*(2), 237–252. https://doi.org/10.1007/s00376-020-0167-7

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., et al. (2010). Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Reviews*, *99*(3-4), 125–161. https://doi.org/10.1016/j.earscirev.2010.02.004

Si, D., & Ding, Y. (2012). The tropospheric biennial oscillation in the East Asian monsoon region and its influence on the precipitation in China and large-scale atmospheric circulation in East Asia. *International journal of climatology*, *32*(11), 1697–1716. https://doi.org/10.1002/joc.238

Van den Hurk, B., Doblas-Reyes, F., Balsamo, G., Koster, R. D., Seneviratne, S. I., & Camargo, H. (2012). Soil moisture effects on seasonal temperature and precipitation forecast scores in Europe. *Climate Dynamics*, *38*(1), 349–362. https://doi.org/10.1007/s00382-010-0956-2

Wang, A., Lettenmaier, D. P. & Sheffield, J. (2011). Soil moisture drought in China, 1950-2006. *Journal of Climate*, 24, 3257-3270. https://doi.org/10.1175/2011JCLI3733.1

Wang, A. & Kong, X. (2021). Regional climate model simulation of soil moisture and its application in drought reconstruction across China from 1911 to 2010. *International Journal of Climatology*, *41* (Suppl. 1), E1028– E1044. https://doi.org/10.1002/joc.6748

Wang, A., & Shi, X. (2019). A multilayer soil moisture dataset based on the gravimetric method in China and its characteristics. *Journal of Hydrometeorology, 20*(8), 1721-1736. https://doi.org/10.1175/JHM-D-19-0035.1

Wang, A., & Zeng, X. (2011). Sensitivities of terrestrial water cycle simulations to the variations of precipitation and air temperature in China. *Journal of Geophysical Research: Atmospheres*, *116*(D2). https://doi.org/10.1029/2010JD014659

Wang, A., Zeng, X., & Guo, D. (2016). Estimates of global surface hydrology and heat fluxes from the Community Land Model (CLM4. 5) with four atmospheric forcing datasets. *Journal of Hydrometeorology*, *17*(9), 2493–2510. https://doi.org/10.1175/JHM-D-16-0041.1

Wang, X., Pang, G., & Yang, M. (2018). Precipitation over the Tibetan Plateau during recent decades: a review based on observations and simulations. *International Journal of Climatology*, *38*(3), 1116–1131. https://doi.org/10.1002/joc.5246

Wu, L., Wang A., & Sheng, Y. (2014) Impact of soil texture on the land surface Process in China (in Chinese). *Climatic and Environmental Research*, *19*(05), 559–571. https://doi.org/10.3878/j.issn.1006-9585

Xia, Y., Sheffield, J., Ek, M. B., Dong, J., Chaney, N., Wei, H., et al. (2014). 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

Xiao, M., Zhang, Q., & Singh, V. P. (2015). Influences of ENSO, NAO, IOD and PDO on seasonal precipitation regimes in the Yangtze River basin, China. *International Journal of Climatology*, *35*(12), 3556–3567. https://doi.org/10.1002/joc.4228

Yang, K., Wang, C., & Bao, H. (2016). Contribution of soil moisture variability to summer precipitation in the Northern Hemisphere. *Journal of Geophysical Research: Atmospheres*, 121, 108–124. https://doi.org/10.1002/2016JD025644

Yeh, T. C., Wetherald, R. T., & Manabe, S. (1984). The effect of soil moisture on the short-term climate and hydrology change—A numerical experiment. *Monthly Weather Review, 112*(3), 474–490. https://doi.org/10.1175/1520-0493(1984)112<0474:TEOSMO>2.0.CO;2

Yuan, S., & Quiring, S. M. (2017). Evaluation of soil moisture in CMIP5 simulations over the contiguous United States using in situ and satellite observations. *Hydrology and Earth System Sciences*, *21*(4), 2203–2218. https://doi.org/10.5194/hess-21-2203-2017

Yuan, S., Quiring, S. M., & Leasor, Z. T. (2021). Historical changes in surface soil moisture over the contiguous United States: an assessment of CMIP6. *Geophysical Research Letters*, *48*(1), e2020GL089991. https://doi.org/10.1029/2020GL089991

Figure 1. The mean soil moisture averaged over the growing season (April–September) for the period of 1992-2013. The soil moisture for 0-10cm (a, b) and 0-100 cm (c,d) derived from WS2019 (a, c) and (c,d) CMIP6 multi-model ensemble mean (MME). There are 732 stations for 0-10cm (a,b) and 177 stations for 0-100 cm soil moisture (c,d), respectively. The boxes indicate the regional range with the two-letter abbreviations provided for each region (XJ: 76°-90°E, 38-48°N; NW: 99.5°-110°E, 32.5-41°N, CE: 110°-117.5°E, 34.5-42°N, NE: 118-130°E, 39.5-50.5°N; YH: 110-120°E 29.5-34.5°N, and SW: 102-107°E, 23-31.5°N).

Figure 2. Boxplots of soil moisture for a) 0-10 cm and b) 0-100 cm soil depths from individual CMIP6 model, MME and OBS, derived firstly from the average soil moisture of growing seasons (April to September) from 1992 to 2013 at all available stations (locations are indicated in Figure 1). Boxes indicate the 75th (bottom) and 25th (top) spatial percentiles, center lines are median value across all stations, and dashed outlines are the 90th (bottom) and 10th (top) spatial percentiles. Models with the same LSM are shaded with same color and name of LSM is also indicated at the top of the figure.

Figure 3. Taylor diagrams of (a) near-surface (0-10 cm) and (b) root-zone (0-100cm) soil moisture statistics derived from the individual CMIP6 model simulation, MME and OBS across all available stations. Statistics are computed from the growing seasonal mean soil moisture from 1992 to 2013. Makers with the same color denote CMIP6 models with the same LSM.

Figure 4. Seasonal cycles of regional mean surface (0-10cm) soil moisture for station observation (WS2019, red curve), individual CMIP6 model simulation (thin black curve), and MME (bold black curve) from the average soil moisture of growing seasons (April to September) from 1992 to 2013. Regional divisions are indicated in Figure 1a.

Figure 5. Similar as Figure 4, but for 0-100 cm soil moisture.

Figure 6. Temporal correlation coefficients of regional mean a) 0-10cm and b) 0-100cm soil moisture computed from simulations of individual CMIP6 model simulation and MME against OBS for the period of 1992 to 2013. The model simulations are firstly interpolated at the station locations using near-neighbor interpolation scheme and the regional mean is then computed. Regional divisions are indicated in Figure 1a. The black dots indicate the correlation coefficient passing 95% significant level test.

Figure 7. Linear trend of annual 0-10 cm soil moisture for the period of 1961-2014 from (a) Observation (VIC-CN05) and (b) MME simulation of 25 CMIP6 models. Results from each individual model are indicated in Figure S3. Trend passing 95% significant level test are hatched with dots. c) The percentage number of models with their soil moisture trend-sign consistent with that from observation at each grid cell. The value showing in the figure are the average over conterminous land area in China. (d) The percentage of land area showing soil moisture with significant uptrends (blue) and downtrends (red).

Figure 8. Interannual standard deviation (STD) of 0-10cm soil moisture from (a) Observation (VIC-CN05) and (b) MME simulation of 25 CMIP6 models for the period of 1961-2014. Results from each individual model are indicated in Figure S4. c) shows the centered pattern correlation of interannual STD between individual model simulation and observation, and between MME and observation.

Table 1 List of information of 40 CMIP6 models used in this study, include model names, spatial resolutions, land surface model names, total

soil depth, number of soil vertical layers, and the DOI ID of each model (https://doi.org/10.22033/ESGF/CMIP6.ID).

Model name	Spatial Resolution (~lon°×lat°)	Land surface model (LSM)	Total soil depth (meter)	Number of soil vertical layers	DOI ID
CESM2	×0.94	CLM5			
CESM2-WACCM	×0.94	CLM5			
CMCC-CM2-HR4	×0.94	CLM4.5			
CMCC-CM2-SR5	×0.94	CLM4.5			
CMCC-ESM2	×0.94	CLM4.5			
FGOALS-f3-L	×1.0	CLM4			
FGOALS-g3	×2.0	CLM4			
FIO-ESM-2-0	×0.94	CLM4			
NorCPM1	×1.875	CLM4			
NorESM2-LM	×1.875	CLM4			
NorESM2-MM	×0.94	CLM4			
SAM0-UNICON	×0.94	CLM4			
TaiESM1	×0.94	CLM4			
EC-Earth3-AerChem	×0.70	HTESSEL			
EC-Earth3-CC	×0.70	HTESSEL			
EC-Earth3	×0.70	HTESSEL			
EC-Earth3-Veg	×0.70	HTESSEL			
EC-Earth3-Veg-LR	×1.125	HTESSEL			
AWI-ESM-1-1-LR	×1.875	JSBACH			

Model name	Spatial Resolution (~lon°×lat°)	Land surface model (LSM)	Total soil depth (meter)	Number of soil vertical layers	DOI ID
MPI-ESM-1-2-HAM	×1.875	JSBACH			
MPI-ESM1-2-HR	×0.94	JSBACH			
MPI-ESM1-2-LR	×1.875	JSBACH			
E3SM-1-0	×1.0	ELM1.0			
E3SM-1-1-ECA	×1.0	ELM1.1			
E3SM-1-1	×1.0	ELM1.1			
GISS-E2-1-G-CC	×2.0	GISS-LSM			
GISS-E2-1-G	×2.0	GISS-LSM			
GISS-E2-1-H	×1.25	GISS-LSM			
ACCESS-CM2	×1.25	CABLE2.5			4271
ACESS-ESM1-5	×1.25	CABLE2.4			
BCC-CSM2-MR	×1.125	BCC_AVIM2			
BCC-ESM1	×2.8	BCC_AVIM2			
GFDL-CM4	×1.0	GFDL-LM4.0.1			
GFDL-ESM4	×1.0	GFDL-LM4.1			
CanESM5	×2.8	CLASS3.6			
CAS-ESM2-0	×1.4	CoLM			
IPSL-CM6A-LR	×1.25	ORCHIDEE 2.0			
KACE-1-0-G	×1.25	JULES			
MIROC6	×1.41	MATSIRO6.0			
MRI-ESM2-0	×1.0	HAL			

Note: * models with 0-100 cm soil moisture are used in the analysis.

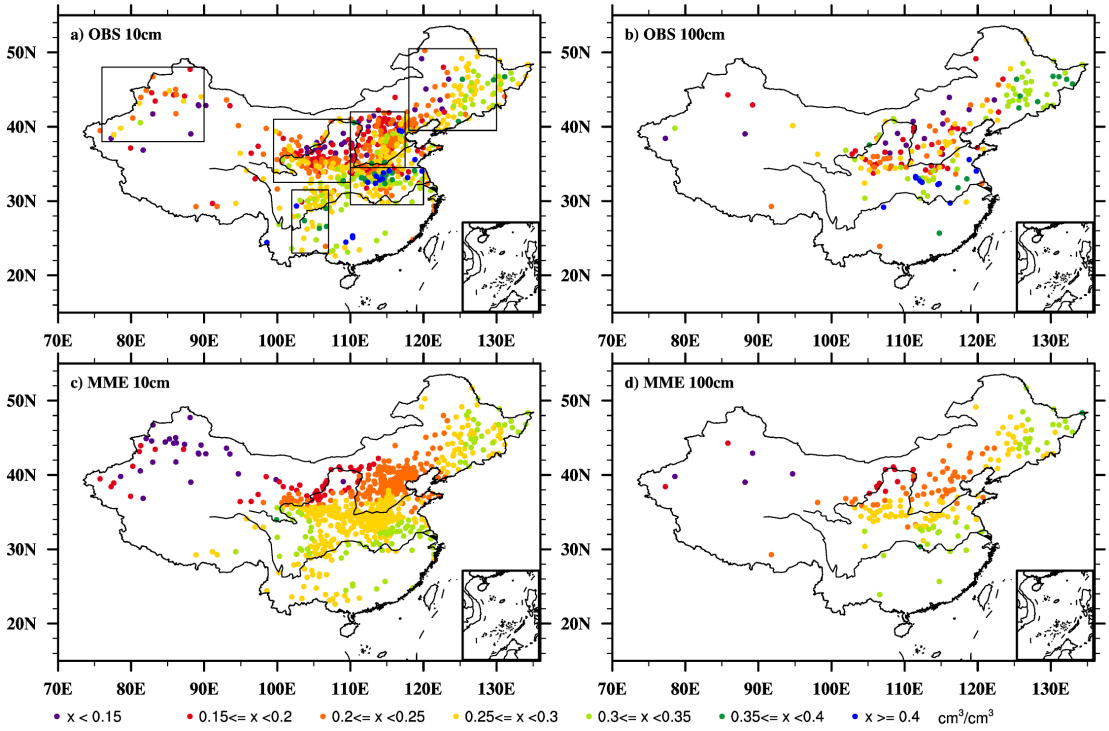

Figure 1. The mean soil moisture averaged over the growing season (April–September) for the period of 1992-2013. The soil moisture for 0-10cm (a, b) and 0-100 cm (c,d) derived from WS2019 (a, c) and (c,d) CMIP6 multi-model ensemble mean (MME). There are 732 stations for 0-10cm (a,b) and 177 stations for 0-100 cm soil moisture (c,d), respectively. The boxes indicate the regional range with the two-letter abbreviations provided for each region (XJ: 76°-90°E, 38-48°N; NW: 99.5°-110°E, 32.5-41°N, CE: 110°-117.5°E, 34.5-42°N, NE: 118-130°E, 39.5-50.5°N; YH: 110-120°E 29.5-34.5°N, and SW: 102-107°E, 23-31.5°N).

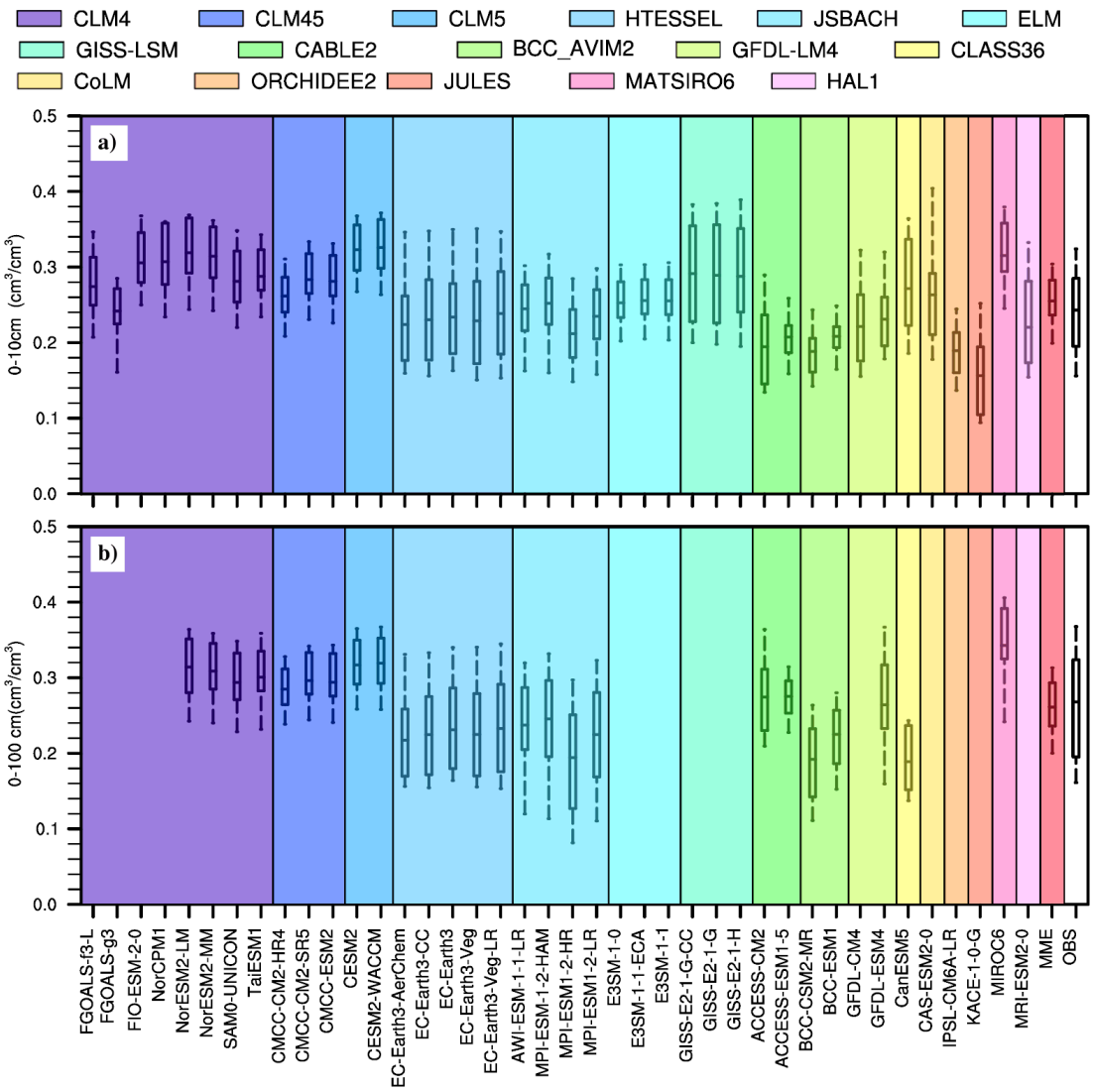

Figure 2. Boxplots of soil moisture for a) 0-10 cm and b) 0-100 cm soil depths from individual CMIP6 model, MME and OBS, derived firstly from the average soil moisture of growing seasons (April to September) from 1992 to 2013 at all available stations (locations are indicated in Figure 1). Boxes indicate the 75^{th} (bottom) and 25^{th} (top) spatial percentiles, center lines are median value across all stations, and dashed outlines are the 90^{th} (bottom) and 10^{th} (top) spatial percentiles. Models with the same LSM are shaded with same color and name of LSM is also indicated at the top of the figure.

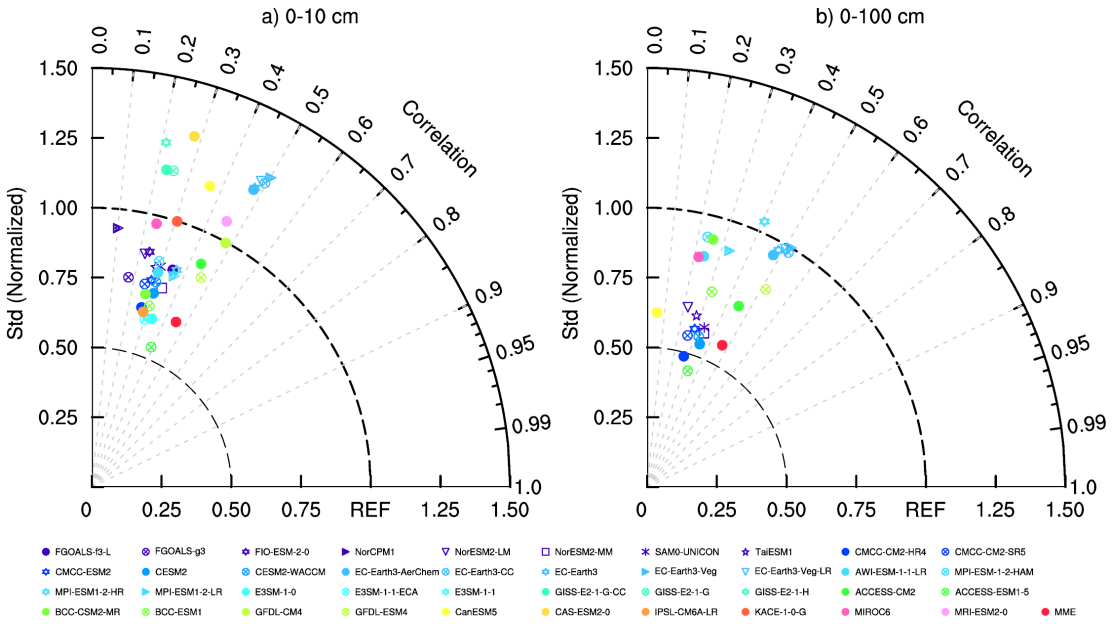

Figure 3. Taylor diagrams of (a) near-surface (0-10 cm) and (b) root-zone (0-100cm) soil moisture statistics derived from the individual CMIP6 model simulation, MME and OBS across all available stations. Statistics are computed from the growing seasonal mean soil moisture from 1992 to 2013. Makers with the same color denote CMIP6 models with the same LSM.

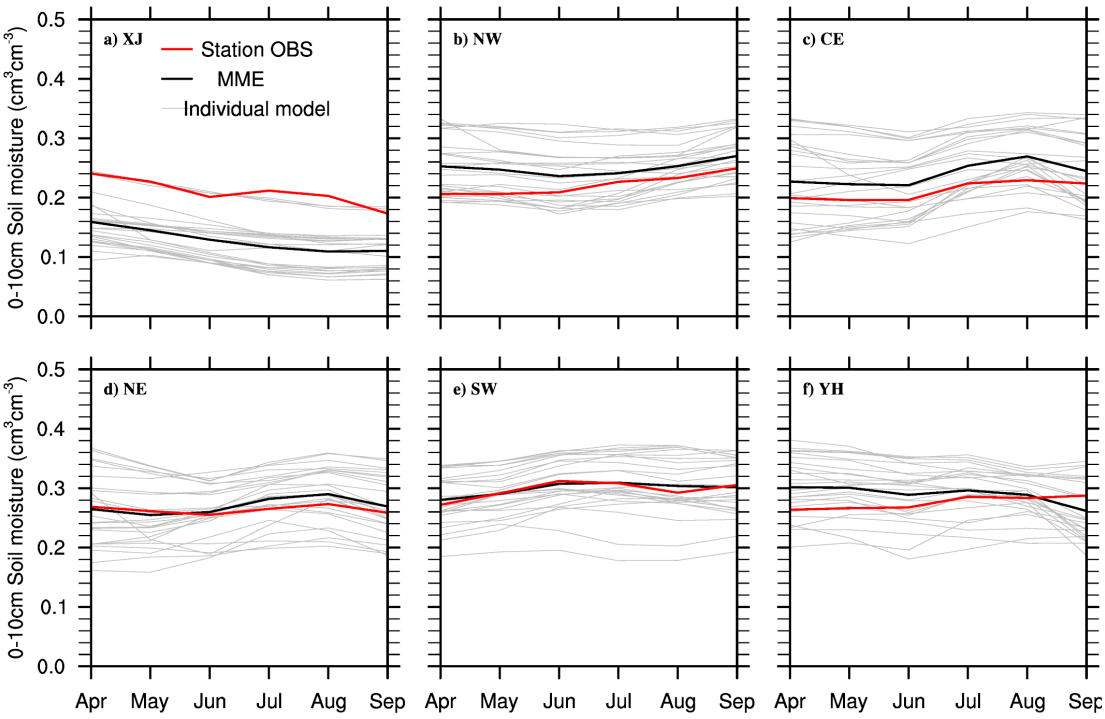

Figure 4. Seasonal cycles of regional mean surface (0-10cm) soil moisture for station observation (WS2019, red curve), individual CMIP6 model simulation (thin black curve), and MME (bold black curve) from the average soil moisture of growing seasons (April to September) from 1992 to 2013. Regional divisions are indicated in Figure 1a.

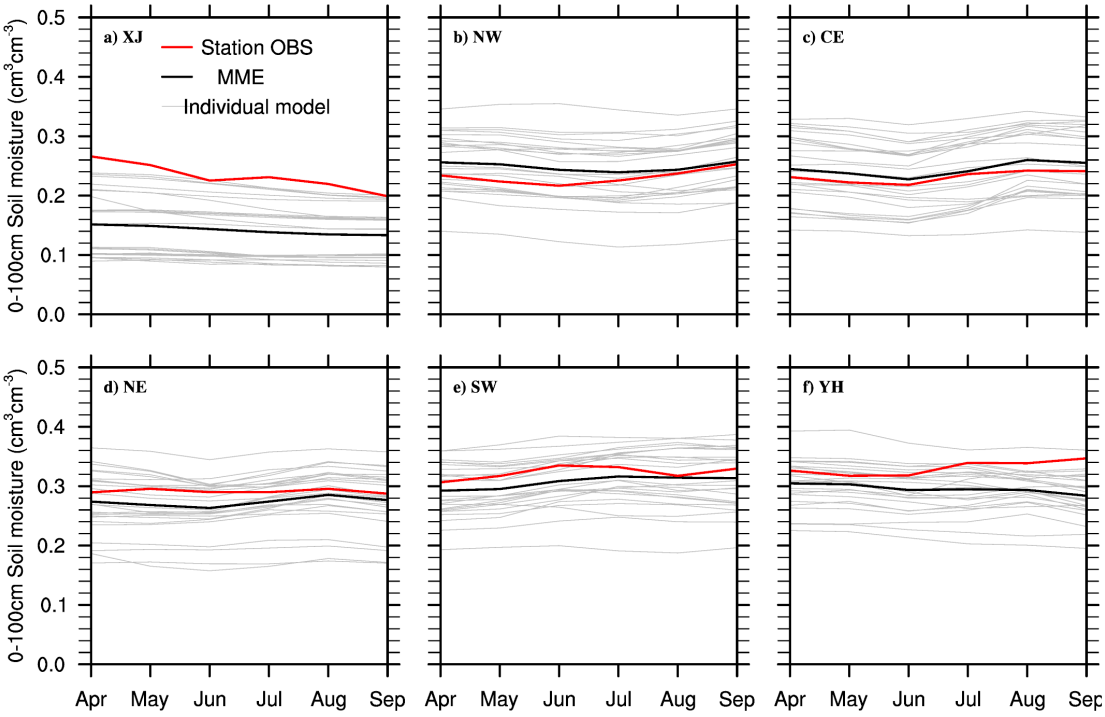

Figure 5. Similar as Figure 4, but for 0-100 cm soil moisture.

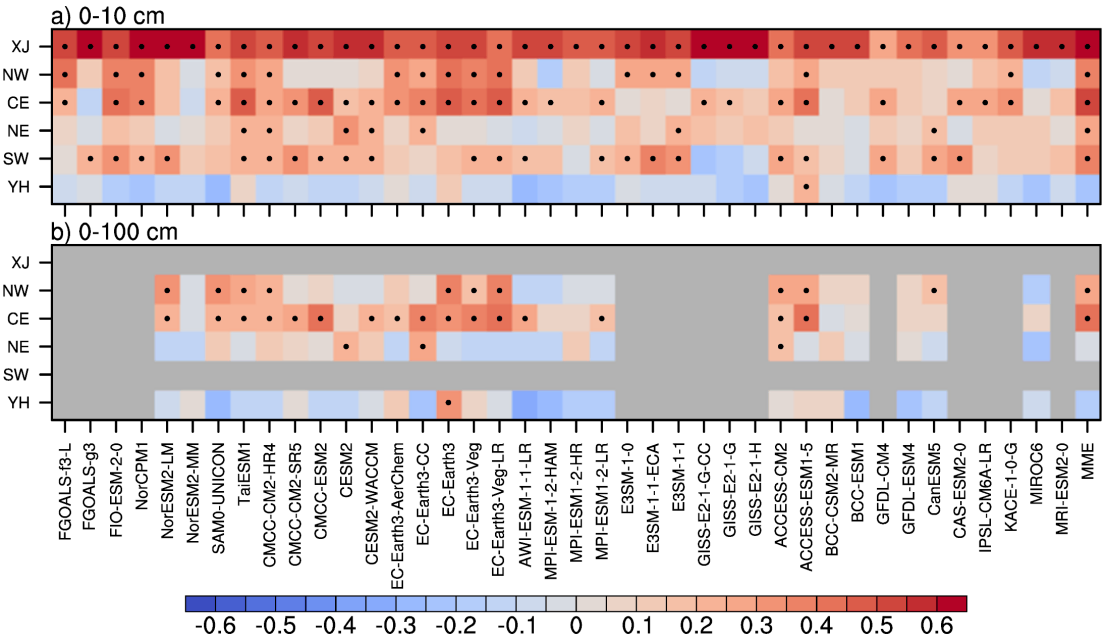

Figure 6. Temporal correlation coefficients of regional mean a) 0-10cm and b) 0-100 cm soil moisture computed from simulations of individual CMIP6 model simulation and MME against OBS for the period of 1992 to 2013. The model simulations are firstly interpolated at the station locations using near-neighbor interpolation scheme and the regional mean is then computed. Regional divisions are indicated in Figure 1a. The black dots indicate the correlation coefficient passing 95% significant level test.

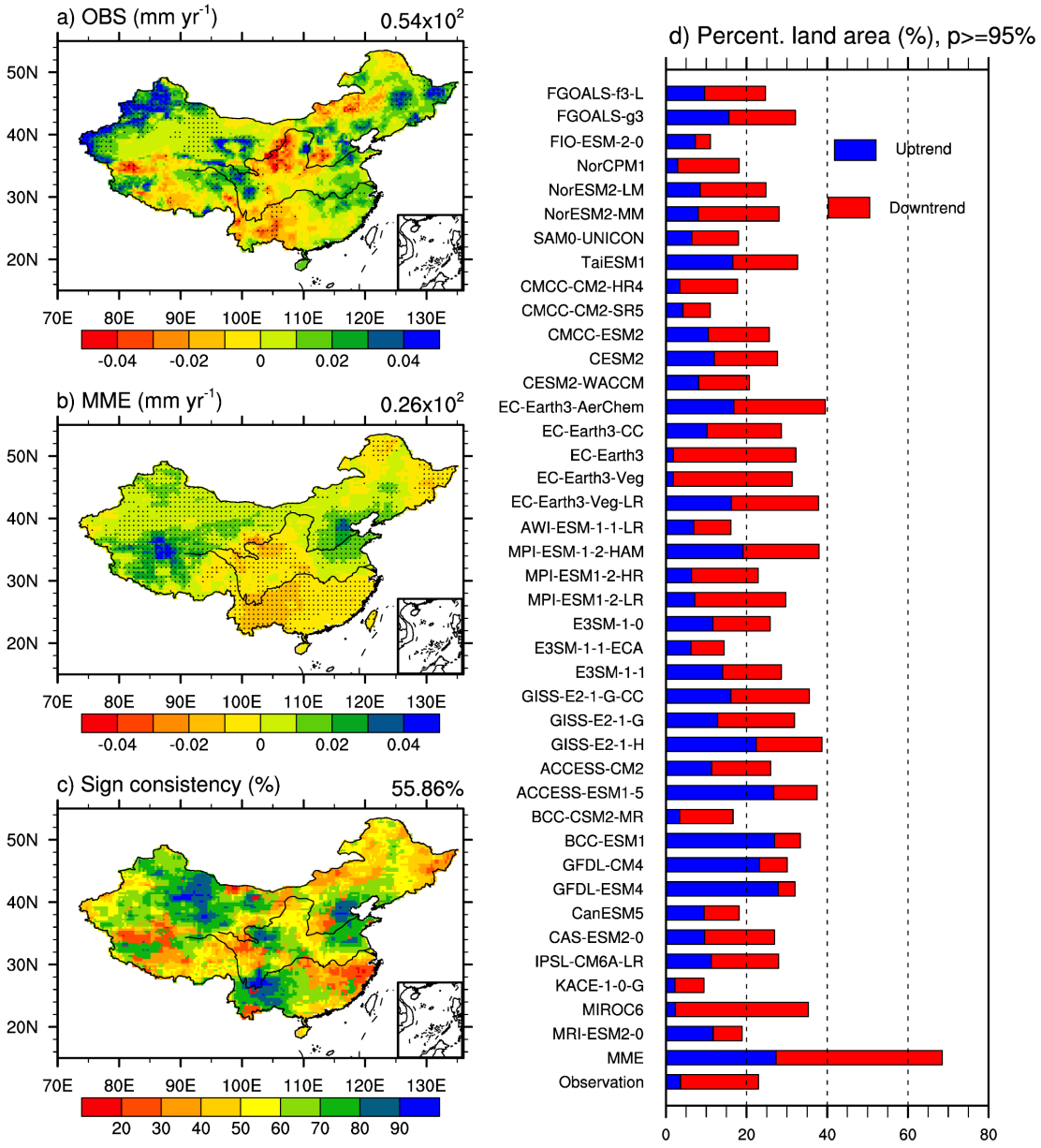

Figure 7. Linear trend of annual 0-10 cm soil moisture for the period of 1961-2014 from (a) Observation (VIC-CN05) and (b) MME simulation of 25 CMIP6 models. Results from each individual model are indicated in Figure S3. Trend passing 95% significant level test are hatched with dots. c) The percentage number of models with their soil moisture trend-sign consistent with that from observation at each grid cell. The value showing in the figure are the average

over conterminous land area in China. (d) The percentage of land area showing soil moisture with significant uptrends (blue) and downtrends (red).

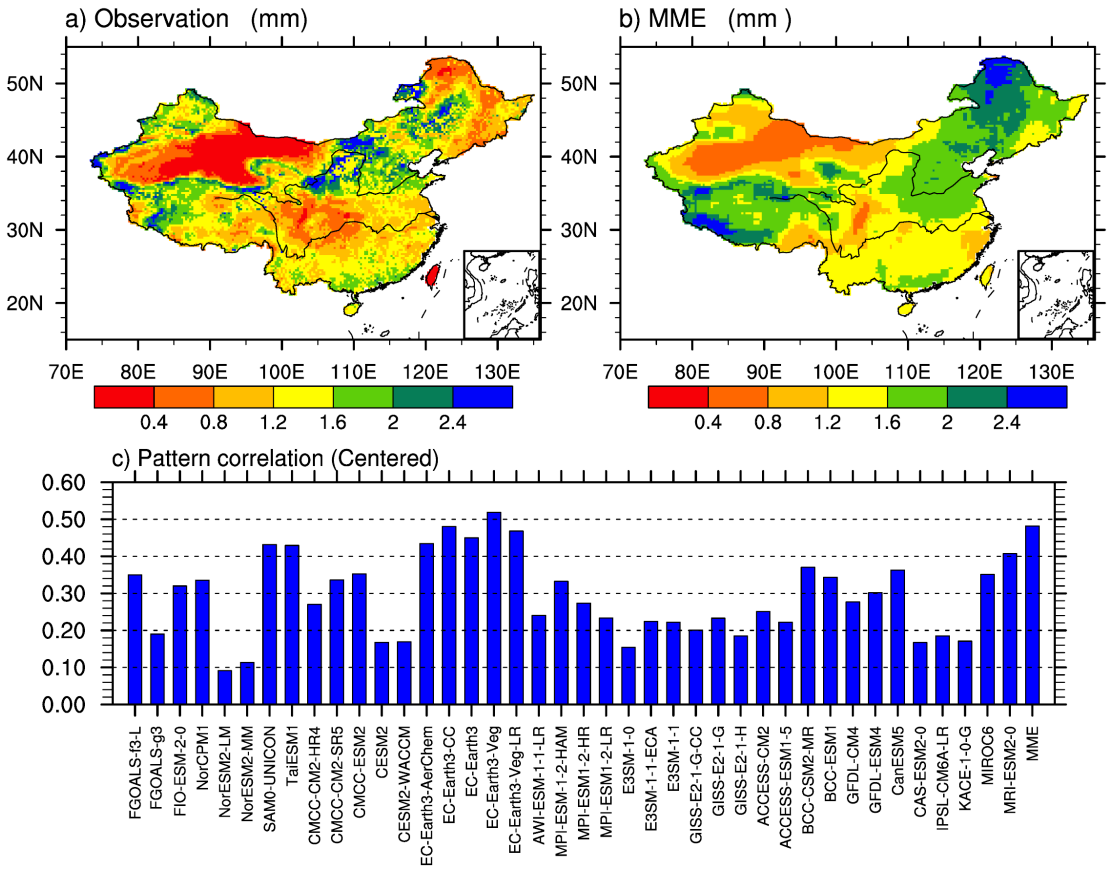

Figure 8. Interannual standard deviation (STD) of 0-10cm soil moisture from (a) Observation (VIC-CN05) and (b) MME simulation of 25 CMIP6 models for the period of 1961-2014. Results from each individual model are indicated in Figure S4. c) shows the centered pattern correlation of interannual STD between individual model simulation and observation, and between MME and observation.