

Reply to Referee #1 for “Evaluation of ORCHIDEE-MICT simulated soil moisture over China and impacts of different atmospheric forcing data” on HESSD

Z. Yin on behalf of all co-authors

1 General comments

1.1 **“The comparison against CCI soil moisture only for 2007 until 2009 is a odd choice. CCI is a unique soil moisture dataset in being based on observations and covering a long time period. This makes it different to other available long-term soil moisture datasets based on model output and other observation based datasets which are usually much shorter. Therefore it should be taken as what it was designed for without cherry-picking the best period. Also, these long time periods will be likely much more interesting for most readers as a limited amount of specific years. Also, the comparison is not too meaningful if the other datasets experiments are not compared for the same time period.”**

A: The primary aim of the comparison using ESA CCI soil moisture (SM) is to assess our model outputs. The availability of ESA CCI SM varies a lot due to changes in sensors. Figure R1 shows the fraction of days with valid observations in different periods from Dorigo et al. (2015). It is clear that the fraction is extremely low in China (less than 0.2) until 2006. If we zoom in China and check CCI SM time series at some grid cells (Fig. R2), the availability varies not only temporally but also spatially. Once again, as our aim is to evaluate ORCHIDEE simulated soil moisture (spatial and temporal patterns), we have to select a period which presents the less gaps in the time series in order to be able to compare our evaluation metrics and their spatial variations. This explains why we choose the more recent period in the dataset, covering our simulations. In the manuscript, we modified as (Page 7, Line 1-5): **“The data availability also varies along the period according to the number of instruments available 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 before 2006 (Fig. A1). ... To provide a reliable validation, we only use the CCI data between 2007-2009.”**

However, we agree that it is interesting for readers to see a long time period comparison. Thus we will provide the comparison between ESA CCI and ORCHIDEE SM from 1981

to 2009 in online supplementary, as Table A1 and Figure B4. Comparing to the same analysis based on period 2007-2009 (Fig. 4 and Table 3 in the revised manuscript), there is no significant change of spatial patterns. Although the values of r and RMSE change slightly, they do not influence our conclusion that GSWP3 and WFDEI performance better than the other two. We added in the manuscript (Page 12, Line 19-23): “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.”, and (Page 14, Line 9-10): “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.”

1.2 “Please make the choice of GLEAM clearer. It uses a lot of observations but it essentially is also model output. So you are comparing your model output to another model (which uses a different precipitation forcing?) Possibly give a little more literature on other soil moisture datasets, why specifically GLEAM, e.g. long time period?”

A: True. GLEAM SM is a model output, but it is corrected by numerous satellite and in-situ measurements through data assimilation. The comparison to GLEAM SM is to provide an assessment of SM dynamics at longer time period. In the introduction, we will explain the aim of using different SM datasets, as (Page 3, Line 22-30): “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 Peking University; Piao et al. (2009); Xu (2014)) are used to evaluate temporal validation of ORCHIDEE 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 soil moisture 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 of the radiative transfer at the soil-vegetation interface calibrated with in-situ data.”

And at Page 3, Line 31 – Page 4, Line 2: “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 data, so that

evaluation of our model against root-zone SM from GLEAM is consistent with evaluation against surface SM from CCI.”

1.3 “Describe why you specifically chose those four forcing datasets. Are they being frequently updated? Also usable for global studies?, etc.”

A: These four datasets are widely used in large scale hydrological studies, which are suitable for global simulation for next step. In the introduction, we added (Page 3, Line 14-21): “Four global atmospheric forcing datasets are chosen to drive the simulations in China, including 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), 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, 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.”

1.4 “The motivation of carrying out the study specifically over China is in my view lacking a little. Also, are there no locally optimised forcing datasets available? Why run a land surface model specifically over China using global input data? Again, just make the motivation of the study a bit clearer. Why was this specific model used for the experiments, does it have any advantages specifically for China (this is actually mentioned in the model section but might be also helpful in the introduction with a little more detail)?”

A: China is selected as the study area before global application, because (1) It covers multiple climate zones, which can help us understand different mechanisms under different climate regimes; (2) It has almost all types of anthropogenic impacts: irrigation, deforestation, afforestation, dam operations, polders, inter basin water transfer, etc, which is an ideal example to investigate climate-water-human interactions in the next step. In the introduction, we added (Page 2, Line 4-8): “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.”

Yes, there is a high resolution Chinese forcing dataset but not regularly updated. Moreover, using global forcing can help us to fast extend some of the metrics developed in this study at global scale. And it will be easy to compare other parallel works using the

same forcing. Related modification is shown in the reply to Comment 1.3.

To further understand the interactions between climate change, water cycle and human activities, a model that has been carefully evaluated is necessary, which integrates important mechanisms of CO₂, water and surface energy balances, ecological dynamics and anthropogenic processes, such as ORCHIDEE. In the introduction, the text is modified as (Page 3, Line 1-13): “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). ... 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. Besides land surface hydrology, ORCHIDEE 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, this evaluation of simulated SM controlled only by natural processes is useful to identify human effects (e.g., crops, irrigation and dam operation, etc) on water budget in regions where there is a large misfit between model and observation.”

1.5 “Concerning the validation as a whole, the model outputs for the four experiments are compared to, in addition to in-situ measurements, GLEAM and CCI soil moisture. However, these datasets (CCI and GLEAM) should also be compared to the in-situ measurements since the mere comparison does not result in any helpful answer on which of these datasets performs any better when compared to the actual ground measurements. Both GLEAM and the CCI dataset will likely have their own problems with accurately simulating soil moisture within certain areas. At the current state of the study they are used as a kind of additional ground-truth, which they most certainly are not (and in fact, as correctly noted, GLEAM shares some of the same input data with the forcing data used for the experiments).”

A: True. Comparison to in-situ measurement is an essential part of model validation. But this is not enough. Firstly, in-situ measurements cannot be used to estimate spatial and temporal variations of simulated SM. Secondly, because of spatial variations of climate variables and landscape, in-situ measurements can provide high accurate validation only if the atmospheric forcing is at the same spatial scale. A remote sensing product derived from multiple observations averaged or aggregated at daily time step is probably more comparable to model simulations obtained with meteorological reanalysis than local in-situ measurements.

We agree that ESA CCI and GLEAM SM should be validated before comparing to the simulations. In fact, the ESA CCI SM has been validated both at global scale (Dorigo et al., 2015) and in China (Peng et al., 2015; An et al., 2016). The GLEAM SM has been validated by ISMN as well (Martens et al., 2017). We will cite these works in the introduction of ESA CCI and GLEAM dataset. All in all, we use SM datasets for different purposes: 1) In-situ measurements are used for evaluation of fast variability,

mainly decrease and recharge of top and middle soil horizons from rain events; 2) ESA CCI and GLEAM are used mainly to evaluate seasonal, inter-annual and spatial patterns of SM. This will be explained in the revised Introduction (see reply to Comment 1.2).

2 Specific comments

2.1 “P2L21-32: Possibly add a sentence on soil moisture (products or raw data) data assimilation in the introduction, since the advantage / disadvantages of satellite based soil moisture products and land surface models are discussed. Data assimilation exactly tries to combine the strengths of these different types of data, such as in the GLEAM dataset.”

A: Thanks for the suggestion. A short description about data assimilation has been added in the introduction (Page 2, Line 30-33): “To overcome the uneven coverage of raw data, data assimilation is widely applied to analyze SM from in-situ or satellite observations (Reichle et al., 2007; Draper et al., 2012; Martens et al., 2016). Analyzed products help us understanding SM variation and its relation to climate (Taylor et al., 2012; Liu et al., 2015b, 2017). However, to capture changes of hydrological mechanisms for future projections, measurements are not enough.”

2.2 “P3L22-29: GSWP3 is very coarse, but downsampled. Could this be especially problematic in areas within China with complex terrain?”

A: Yes, coarse resolution is inaccurate over complex terrain regions from the Tibetan Plateau to the Sichuan Basin (Fig. 1). It is difficult to produce a high accurate forcing reanalysis dataset and simulated SM in these areas. Disagreements of simulated SM are shown in Fig. B11(a) in online supplementary. Mismatches of meteorological variables among forcing datasets are found in these regions as well (Fig. B11(c) and B11(e)). However, according to the results, the simulated SM driven by GSWP3 is not worse than others. So in our opinion, GSWP3 is not obviously less realistic in complex terrain regions. Inter-comparison should be applied among forcing datasets to further address this question, e.g. comparison with weather stations data, but it is beyond our ability and not the scope of this study.

2.3 “P4L2-8: Is PGF still being updated? Maybe add this information to the other datasets too, or to the motivation of choosing these specific datasets.”

A: The last update of the PGF is on 13th July 2014 (<http://hydrology.princeton.edu/data.pgf.php>). The version information of each dataset will be added. The motivation of using these forcing data is added in the introduction. Please see our reply to Comment 1.3.

2.4 “P4L18-22: WFDEI, why only available until 2009? Both corrected with GPCC v5 and v6?”

A: The WFDEI_GPCC we have access to is only available until 2009 (version 31 July 2012). The version we used was only corrected with GPCC v5. The initial WFDEI forcing should be bias corrected for rainfall for use in ORCHIDEE. However, this work had not been completed when we ran the simulations.

Currently the latest WFDEI-GPCC is available. We can re-do the simulation with the latest WFDEI forcing if the reviewer strongly recommend. However, please note that there will be only one year (2010) extension due to the constraint of other forcing (e.g., GSWP3 is available only until 2010). Moreover, there is no difference of the version we used and the latest version, except for the period length ([http://www.eu-watch.org/gfx_content/documents/README-WFDEI%20\(v2016\).pdf](http://www.eu-watch.org/gfx_content/documents/README-WFDEI%20(v2016).pdf)).

2.5 “P5L20: GLEAM has 0.5° resolution? I thought 0.25. It’s 0.25° in Table 2, please recheck.”

A: True, it is 0.25°. Corrected.

2.6 “P5L28 GLEAM assimilates GLDAS? I’m not so sure about this. I think it’s somehow used for the background error estimation within the assimilation scheme, but please check this.”

A: We made a mistake here. The GLDAS was only used in the GLEAM data assimilation system to estimate the errors on annual basis (Martens et al., 2017). We corrected as (Page 7, Line 13-17): “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) and ESA CCI SM (ESA Climate Change Initiative Soil Moisture) 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) was used to estimate the errors of these SM products.”

2.7 “P6L26: 13 PFTs are grouped, did not understand. Only three land cover classes?”

A: No. We used 13-PFT map including one bare soil. Each PFT has its own parameterization. PFT fractions are assigned to three soil tiles corresponding to bare soil, short vegetation (grass and crop PFTs) and forests (all tree PFTs). So each grid cell can include up to three soil tiles. The soil moisture budget of each soil tile is calculated separately, but different PFTs in the same soil tile interact as they share the same soil moisture source. In the manuscript, we modified as (Page 8, Line 2-4): “Each grid cell can include up to three soil tiles: bare soil, trees and grasscrops, 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.”

2.8 “P7L10. Why aggregate results to 1 degree? This likely deteriorates the impact of ‘high resolution’ forcing datasets, such as WFDEI. Rather upsample coarser data by simply multiplying grid cells?”

A: Agreed. In the revised manuscript, we will sample the simulated SM driven by PGF from 1° to 0.5° by a nearest neighbour method and re-do all analysis and plots at 0.5°. Values of metrics will be updated as well. But there is no significant impact on our results and conclusion.

2.9 “P7L15: ‘distributed to the first half of the forcing time step..’ why the first half.”

A: This is the default setting of ORCHIDEE. To avoid underestimation of infiltration, precipitation amount should be assigned in the half of forcing time step. The ‘first half’ is the default setting of ORCHIDEE-MICT.

2.10 “P8L1-3: Which time shift was used? Between UTC and local time (several time zones) between model and in-situ measurements. Not vital but good to know.”

A: We used UTC for all SM records. In the manuscript, we added (Page 9, Line 9-10): “In addition, the timing of all SM datasets is uniformed to the Coordinated Universal Time (UTC).”

2.11 “P8L24: What is the exact motivation for the lag analysis? It does not seem to give any added value. How do you know one or the other are better in temporal terms? You are comparing two models.”

A: We agree that we are not able to explain which one is better by the comparison of the dynamics of SM decrease after rain events. The differences of autocorrelation can tell us the uncertainty of simulated SM from runoff, drainage and transpiration loss after a rainfall. For example, a large mismatch is found in the Yellow RB (Fig. 6(h)), which indicates that some unknowns existed in this region strongly influenced decrease of SM after rain events. In the paper, we discussed (Page 13, Line 24-27): “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 25 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 .” Furthermore, we demonstrate that the R_k curves vary among our simulations. It indicates that the autocorrelation is not only determined by the model but also by forcing data, such as precipitation intensity and frequency, which underlines our motivation: which atmospheric forcing is suitable for further hydrological study in China.

2.12 “P9L12: Monthly values of other variables also considered ... How?”

A: We wanted to explain that not only P and N_p , but also other meteorological variables are included as indicator as well. We modified as (Page 10, Line 15-20): “We look at different climate variables to explain SM differences among simulations. ... Other meteorological indicators are incoming short/long wave radiation (R_s/R_l), air temperature (T_a), air humidity (Q_a) and wind speed (W).”

2.13 “P9L23: These two rivers are the main ones? How much of China do these two watersheds cover? Maybe obvious for some but more background on the study region could be valuable (here and / or in the introduction).”

A: Yes. Yangtze and Yellow are the two largest rivers in China. The watersheds of them cover 23% area of China. More importantly, the two basins cover most of agricultural and industrial regions in China. Motivation of the study area has been added in the Introduction as our reply to Comment 1.4.

More detailed explanation has been added in an extra subsection “Study area” in Section 2, as (Page 4, Line 12-21): “China has multiple climate regimes, which makes hydrological simulations influenced by different variables in different regions. The land water budgets in China is influenced by anthropogenic factors, such as irrigation (Puma and Cook, 2010), afforestation (Peng et al., 2014; Liu et al., 2015a), 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 river basins are of main 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 (grey 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 the two watersheds to assess SM dynamics not only in China but also at catchment scale.”

2.14 “This seasonality is computed across the boundaries depicted in Figure 1? A little more geographically distributed information would be helpful.”

A: Not exactly. The resolution of our simulations is coarser than the GIS data shown in Fig. 1. Thus in the analysis, the specific masks of the two river basins (not illustrated in the manuscript) do not perfectly cover the basins shown in Fig. 1. To avoid this confusion, we wrote (Page 4, Line 21-23): “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.”

2.15 “P10L3: The soil moisture patterns do not necessarily match the annual mean precipitation patterns, maybe mention something about obvious monthly differences, or stronger evaporation using a specific forcing dataset? Soil moisture is not just the result of precipitation but also the other input data and model internal mechanics. No in-depth analysis is needed but some additional maps or statistics for the most important other water balance variables, e.g. evapotranspiration, or at least some sentences on the issue would be helpful. The GLEAM model you are comparing to is actually primarily developed for evapotranspiration.”

A: True. SM does not only depend on MAP. Precipitation frequency and intensity, and evapotranspiration influence SM patterns. And all of them depend on the input: atmospheric forcing. This is the third question we addressed through this study. Considering the length of the paper, Section 4.1 will be removed in the revised version, as it is not tightly related to the topic. The comparison of simulated ET to GLEAM ET has been performed and will be shortly discussed in the autocorrelation analysis (Page 13, Line 23-27): “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 .”

2.16 **“Comparison with ISMN and PKU in-situ data’ seems to be a summary of the model performance for all four forcing datasets when compared to in-situ measurements. It should be noted that these are the average statistics for all carried out experiments. Instead of the next section being ‘Comparison with GLEM ...’ I as a reader would expect a more detailed analysis to follow (or the other way round), which now seems to be in section 4.3 and 4.4. Thus I would recommend to first do the in-depth comparison to in-situ measurements, followed by a comparison to other datasets thereafter. As stated at the beginning, I strongly believe that GLEAM and the CCI dataset should be validated against the in-situ measurements if you want to quantify which model actually performs better in which geographical area.”**

A: The comparison between simulated SM and in-situ measurement is a key part of validation. We will introduce the priorities of these comparisons in the Introduction (see reply to Comment 1.2). However we don’t agree to move the contents in Section 4.3 and 4.4 ahead of the comparison using GLEAM and ESA CCI SM datasets. There are three research questions in this study (see Introduction). The Section 4.2 presents comparison results in order to address question: Is ORCHIDEE able to reproduce reasonable SM dynamics? The Section 4.3 is aim to demonstrate which forcing is suitable for hydrological studies in China. And the Section 4.4 is for the third research question. We think the current order is logical for readers to follow our steps to address the three questions one by one.

However, we recognize that the paragraphs of comparison with in-situ measurements were less detailed than others. In the revised manuscript, we provide a more explicit explanation to the IMSN-PKU section and reduced the GLEAM comparison and Section 4.4, as recommended by reviewer #2. Regarding the validation of ESA CCI and GLEAM SM, please check our replies to Comment 1.2 and 1.5.

2.17 **“The main finding that WFDEI performs best among forcing data is not so surprising when compared to some other studies. Again, more emphasis should be put on why this study is important specifically for China. Maybe compare the outcome of the study to other studies.”**

A: True. Motivation of the study area has been added in the Introduction. Please check our reply to comment 1.4.

2.18 **“Table 2: Correlations are stated as being significant. Was the autocorrelation of the datasets taken into account? Also valid for the correlation at the individual stations.”**

A: No. This table does not include autocorrelation. Because the k -lag is an array of correlations and is difficult to show in a table. Moreover, as the answer to Comment 2.11, the autocorrelation analysis provides another aspect of the dynamics of simulated SM after rainfall. It is hard to distinguish which forcing performs better in this metrics also

reflecting model processes.

We found that the initial caption of Table 2 is confusing. We checked the median p -value of each comparison and all of them are below 0.05. However, in a few grid cells (or stations), the p -value is > 0.05 . We added the quantile of the samples with significant correlation in Table 2, and modified the caption as (Page 28, Line 1-2): “The subscripts of correlation coefficients indicate the quantile of stations (samples) with significant correlation (p -value < 0.05).”

2.19 “Figure 10: Should include description of variable names. Use same variable names in Figures 9 and 10.”

A: We agree. In the revised manuscript, the description of the abbreviations has been added in the caption of Fig. 10 (Fig. 9 in the revision).

Table R1: Median of metrics in the comparison between ESA CCI and ORCHIDEE SM for period 1981-2009. The subscripts of correlation coefficients indicate the quantile of stations (samples) with significant correlation (p -value < 0.05).

| Dataset | Simulations | Correlation | | | RMSE ($\text{m}^3 \cdot \text{m}^{-3}$) | | |
|---------|-------------|----------------------|----------------------|-------------------|---|---------|--------|
| | | China | Yangtze | Yellow | China | Yangtze | Yellow |
| ESA | GSWP3 | 0.33 _{0.96} | 0.29 _{0.97} | 0.45 ₁ | 0.06 | 0.06 | 0.06 |
| | PGF | 0.21 _{0.90} | 0.22 _{0.95} | 0.24 ₁ | 0.07 | 0.07 | 0.07 |
| CCI | CRU-NCEP | 0.37 _{0.96} | 0.34 _{0.97} | 0.46 ₁ | 0.07 | 0.08 | 0.07 |
| | WFDEI | 0.47 _{0.98} | 0.42 _{0.97} | 0.57 ₁ | 0.06 | 0.06 | 0.06 |

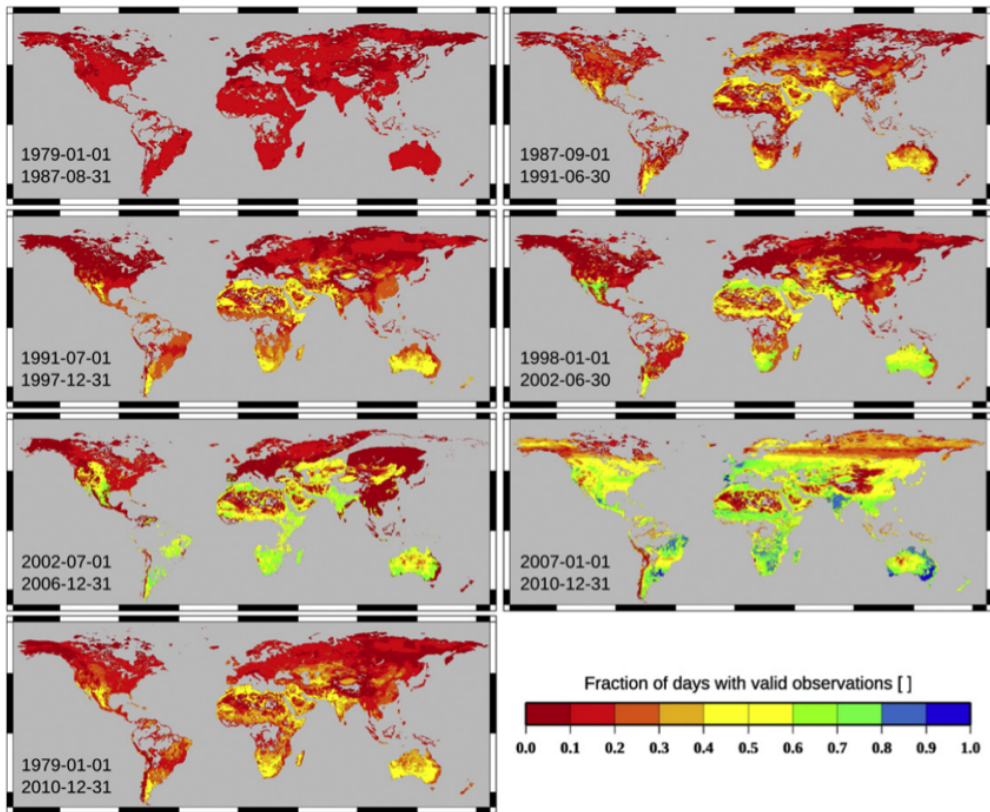


Figure R1: Number of days with available data per month of the ESA CCI soil moisture product.

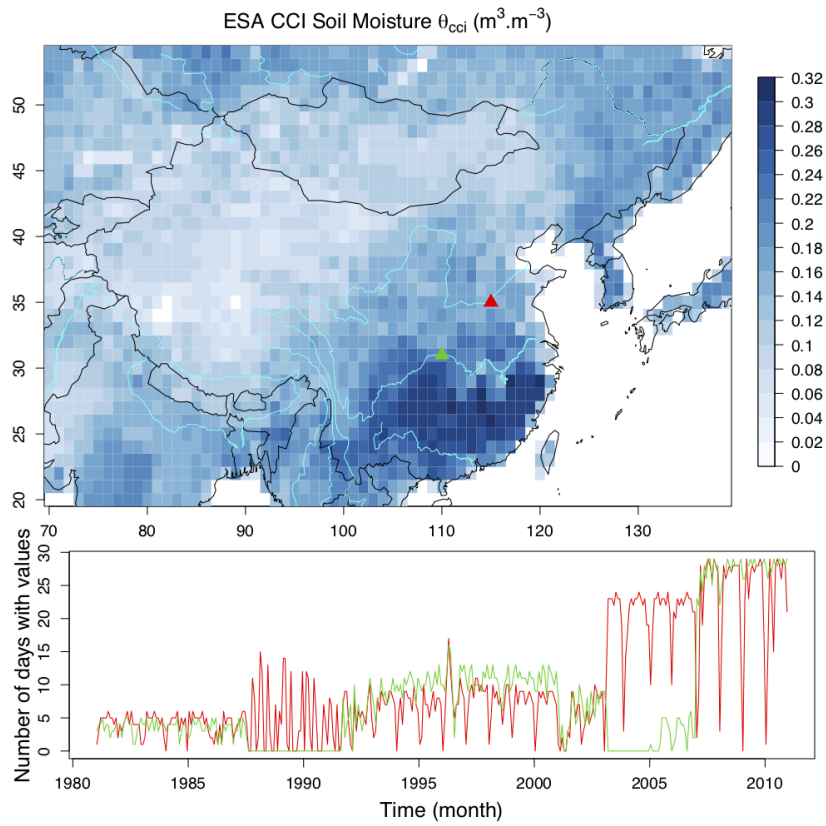


Figure R2: Top panel: annual averaged ESA CCI soil moisture from 1979 to 2010. Bottom panel: fraction of days with available data per month in two grid cells shown in the top panel.

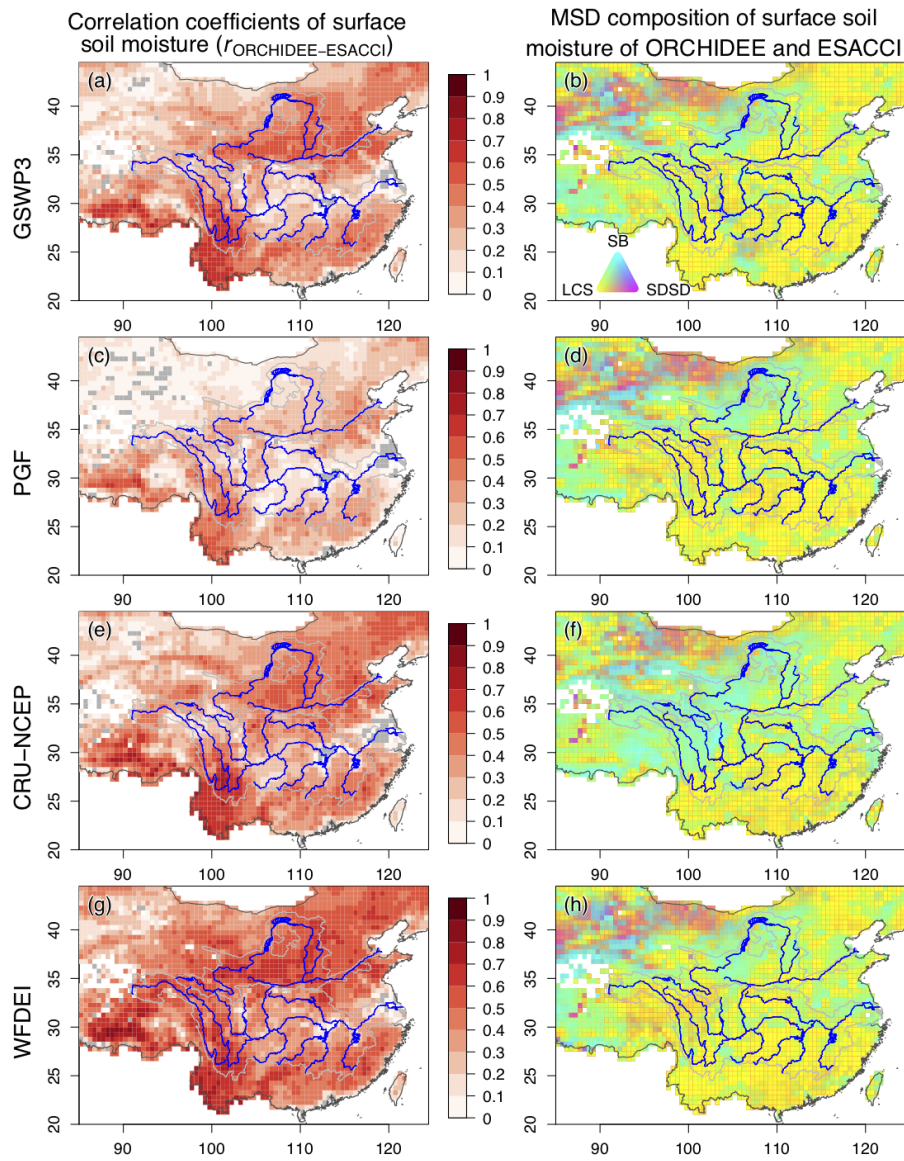


Figure R3: Left panel: Correlation coefficients of the ESA CCI SM and the corresponding ORCHIDEE SM from 1981 to 2009. Gray pixels indicate non and negative correlation. Right panel: decomposition of the MSD between the daily ESA CCI SM and the corresponding ORCHIDEE SM. Cyan, magenta and yellow indicate the fractions of SB, SDSD and LCS respectively.

Bibliography

- An, R., Zhang, L., Wang, Z., Quayle-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>, URL <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>, URL <http://linkinghub.elsevier.com/retrieve/pii/S0925857417304147>, 2017.
- 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>, URL <http://linkinghub.elsevier.com/retrieve/pii/S0048969717329467>, 2018.
- 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>, URL <http://www.nat-hazards-earth-syst-sci.net/9/1995/2009/>, 2009.
- 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>, URL <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., 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.

- 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>, URL <http://doi.wiley.com/10.1029/2011GL050655>, 2012.
- 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.: Revegetation in China’s Loess Plateau is approaching sustainable water resource limits, *Nature Climate Change*, 6, 1019–1022, <https://doi.org/10.1038/nclimate3092>, URL <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>, URL <http://journals.ametsoc.org/doi/abs/10.1175/JHM-D-14-0068.1>, 2014.
- 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>, URL <http://link.springer.com/10.1007/s00382-011-1252-5>, 2012.
- Guimberteau, M., Ducharne, A., Ciais, P., Boisier, J. P., Peng, S., De Weirtdt, 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>, URL <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>, URL <https://www.hydrol-earth-syst-sci.net/21/1455/2017/>, 2017.
- Guimberteau, M., Zhu, D., Maignan, F., Huang, Y., Yue, C., Dantec-Nédélec, S., Ottlé, C., Jornet-Puig, A., Bastos, A., Laurent, P., Goll, D., Bowering, 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., Joetzjer, 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>, URL <https://www.geosci-model-dev.net/11/121/2018/>, 2018.
- He, J. and Yang, K.: China meteorological forcing dataset, Tech. rep., Cold and Arid Regions Science Data Center, Lanzhou, 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>, URL <http://linkinghub.elsevier.com/retrieve/pii/S003442571400337X>, 2014.
- 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>, URL <http://link.springer.com/10.1007/s11269-015-0957-3>, 2015.
- 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>, URL <http://doi.wiley.com/10.1002/hyp.10459>, 2015a.
- 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>, URL <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>, URL <http://content.sciendo.com/view/journals/johh/65/2/article-p123.xml>, 2017.
- 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.

- Peng, J., Niesel, J., Loew, A., Zhang, S. Q., and Wang, J.: Evaluation of Satellite and Re-analysis Soil Moisture Products over Southwest China Using Ground-Based Measurements, *Remote Sensing*, 7, 15 729–15 747, <https://doi.org/10.3390/rs71115729>, URL <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., Myrneni, 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>, URL <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, 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>, URL <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>, URL <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 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, D09108, <https://doi.org/10.1029/2006JD008033>, URL <http://doi.wiley.com/10.1029/2006JD008033>, 2007.
- 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>, URL <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>, URL <http://www.proc-iahs.net/374/53/2016/>, 2016.
- 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.
- 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>, URL <https://www.hydrol-earth-syst-sci.net/22/1811/2018/>, 2018.
- 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.
- 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>, URL <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>, URL <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 Wheeler, 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>, URL <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.
- 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>, URL <http://doi.wiley.com/10.1111/gcb.13983>, 2018.
- 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>, URL <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>, URL <http://doi.wiley.com/10.1002/2015JD023129>, 2015.
- 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.
- 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>, URL <https://www.hydrol-earth-syst-sci-discuss.net/hess-2018-88/>, 2018.
- 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.

Reply to Referee #2 for “Evaluation of ORCHIDEE-MICT simulated soil moisture over China and impacts of different atmospheric forcing data” on HESSD

Z. Yin on behalf of all co-authors

1 Major comments

1.1 **“The abstract contains some details that cannot be understood by reading the abstract only (it should be avoided). For instance, median R and RMSE are reported at page 1 - line 9 without mentioning with respect to which dataset they are computed. The reference to SB and LSC metrics is given but the reader is not able to understand what these metrics represent. Why are they used? Similarly for the discrepancies metric. I suggest mentioning in the abstract the results in general terms, without referring to metrics not know to the reader.”**

A: True. Details of the comparison (value of metrics) are removed from the abstract. Other sentences are also slightly modified to make them more clear and brief.

1.2 **“GLEAM contains several datasets included in the atmospheric forcing datasets. It is not only ERA-Interim but also GPCC through MSWEP product. Therefore, I expect a large agreement between GLEAM and modelled soil moisture, but it does not mean the soil moisture simulations are accurate, they are simply consistent with GLEAM soil moisture (as I expected). The corresponding results should be clarified and put in perspective.”**

A: Exactly. Some information is contained in both GLEAM and atmospheric forcing that we used, which may lead to a good agreement between GLEAM and simulated SM. We will discuss this issue in the revised Section of “Discussion and perspective”.

However, the GLEAM SM assimilates a set of satellite observations and ground measurements (Martens et al., 2017). In addition, it has high spatio-temporal integrity in comparison to in-situ and remote sensing SM. Therefore, GLEAM is used to evaluate spatial variation of simulated SM in a long time period. In introduction we will explain the aim of using GLEAM SM, as (Page 3 Line 31 – Page 4, Line 2): **“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 SM from CCI.”

1.3 **“Even though it is a satellite-based dataset (therefore, its accuracy might be not good enough), the use of ESA CCI soil moisture dataset is in my opinion good. However, why only 3 years? I agree with authors that ESA CCI soil moisture product is more accurate after 2007, but for modelling assessment, I would prefer to see a long-term comparison (1980-2017). It is highly needed and to me much more appropriate than using GLEAM.”**

A: Although ESA CCI has long time coverage, data availability (the fraction of days with available measurements) is very low in China until 2006 (as shown in Fig. R1 and R2 in the reply to Reviewer #1). Moreover, the data availability varies significant in both space and time. Thus we decided to only use the data from 2007 to 2009 for comparison. In the manuscript, we modified as (Page 7, Line 1-5): **“The data availability also varies along the period according to the number of available instruments available 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 before 2006 (Fig. B1). ... To provide a reliable validation, we only use the CCI data between 2007-2009.”**

We suggest to present the long time comparison in online supplementary (as Fig. R3 and Table R1 shown in the reply to reviewer #1). However, if both reviewers were aware the limitation of ESA CCI already and consider the long time period comparison more important, we will present it in the revised manuscript instead of the 2007-2009 comparison.

2 Moderate comments

2.1 **“Too many figures, also by considering the Appendix, have been presented in the paper. I would prefer a lower number of more focused figures that would help the reader to understand clearly the main results the authors want to convey. Please try to reduce the length of the paper, mainly the results section.”**

A: True. Section 4.1 will be removed. Section 4.2 (comparison between GLEAM and ORCHIDEE SM), 4.4 and 5.2 will be reduced. Figure 9 will be moved to supplementary. Figure A10, A11, A12 and A13 will be removed.

2.2 **“The sensitivity analysis linking soil moisture and meteorological variables seems to me not robust enough for being published on HESS. I might be wrong, but also the authors acknowledge this problem. I suggest removing or, at least, strongly reducing.”**

A: True. The results and discussions related this analysis will be reduced in the revised version. Several related figures will be removed (see reply to the previous comment).

3 Minor comments

3.1 **“Acronyms and symbols should be specified the first time they appear in the text, please check.”**

A: Revised. Explanation is given before the first appearance of each symbol.

3.2 **“Page 5, line 3: How is it assessed the quality of ISMN stations? Please clarify.”**

A: It should be “availability”. Corrected.

3.3 **“Page 5, line 17: I would not say ‘only’ 203 stations.”**

A: True. Corrected.

3.4 **“Page 7, line 19: Soil depths are not different in the four datasets. If I am right, please remove.”**

A: Here we talk about the SM datasets, not of forcing datasets or simulation outputs. Revised as (Page 8, Line 29): [“As the soil depths, periods and spatio-temporal resolutions are different in the four SM datasets \(Sect. 2.2\)...”](#)

3.5 **“Page 10, line 16: ‘an traditional’ should be ‘a traditional’.”**

A: True. Corrected.

3.6 **“Page 10, line 20: Why the magnitude of soil moisture is systematically underestimated? Please try to find an explanation.”**

A: Here we are talking about the comparison at Xuzhou station. To avoid misunderstanding, it is revised as (Page 11, Line 16): [“However the magnitude of \$\theta_t\$ is systematically underestimated as well \(Fig. 4\).”](#)

3.7 **“Page 12, lines 2-3: Again, why changes of precipitation regimes are not enough to predict changes in soil moisture? Please comment.”**

A: It is confusing. Based on the different trends of θ_s and P , we infer that the trend of precipitation amount cannot well explain the trend of SM. Modified as (Page 13, Line 33): [“The mismatch of \$\theta_s\$ and \$P\$ trends suggest that the change of precipitation amount is not the only driver of the trend of SM.”](#)

Bibliography

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.

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.