Table S1. 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 ($m^3.m^{-3}$)		
		China	Yangtze	Yellow	China	Yangtze	Yellow
ESA CCI	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
	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

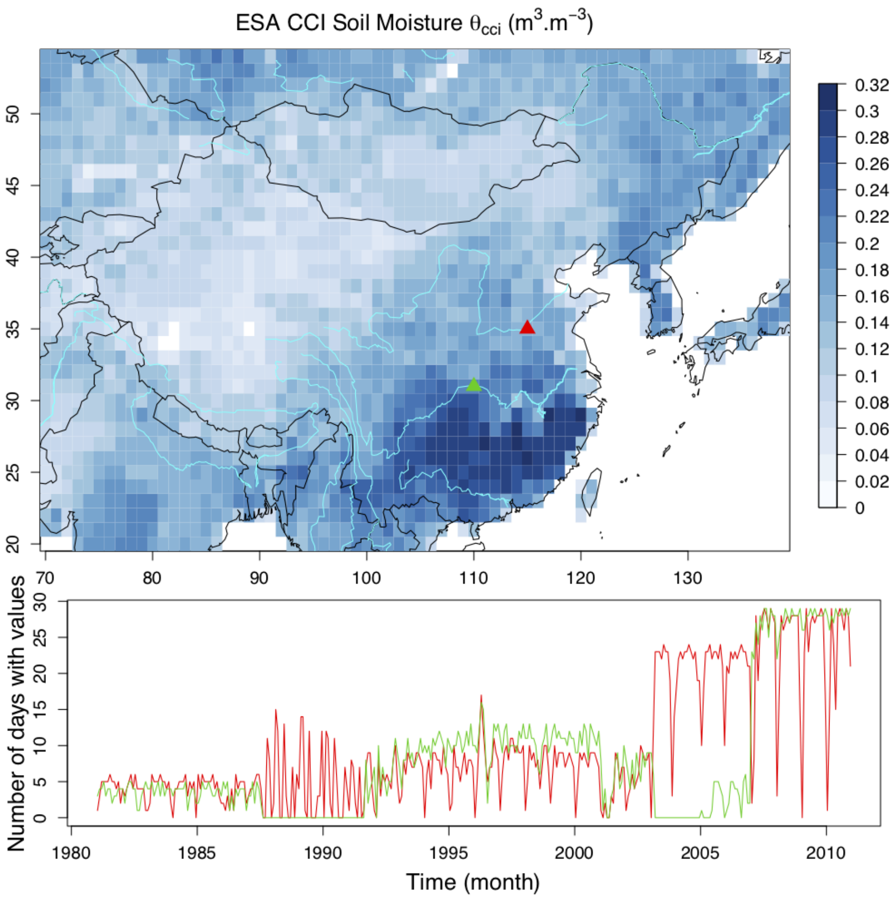

Figure S1. Top panel: annual averaged ESA CCI SM from 1979 to 2010. Bottom panel: fraction of days with available records per month at two grid cells shown as triangles in the top panel.

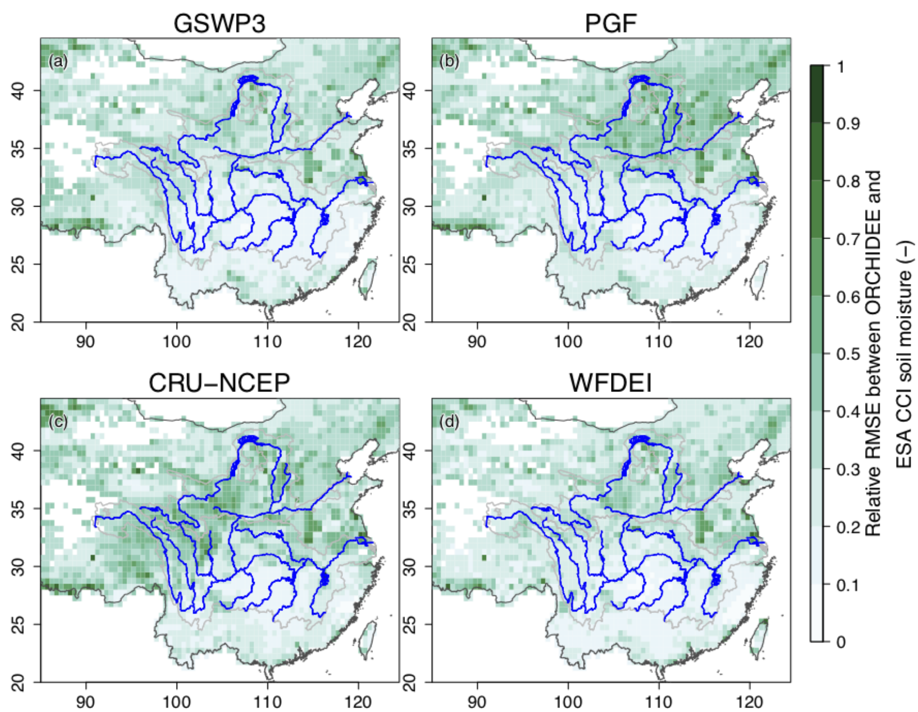

Figure S2. Relative RMSE between the daily ESA CCI SM and the corresponding ORCHIDEE SM.

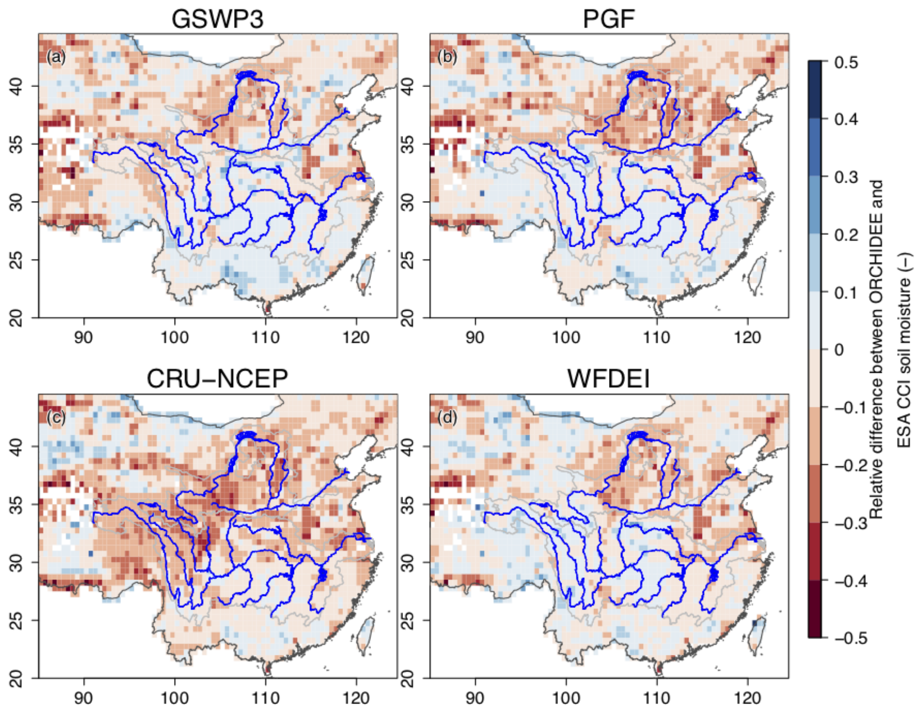

Figure S3. Relative difference ($\Delta_{orc\text{-}cci}$) between the daily ESA CCI SM and the corresponding ORCHIDEE SM.

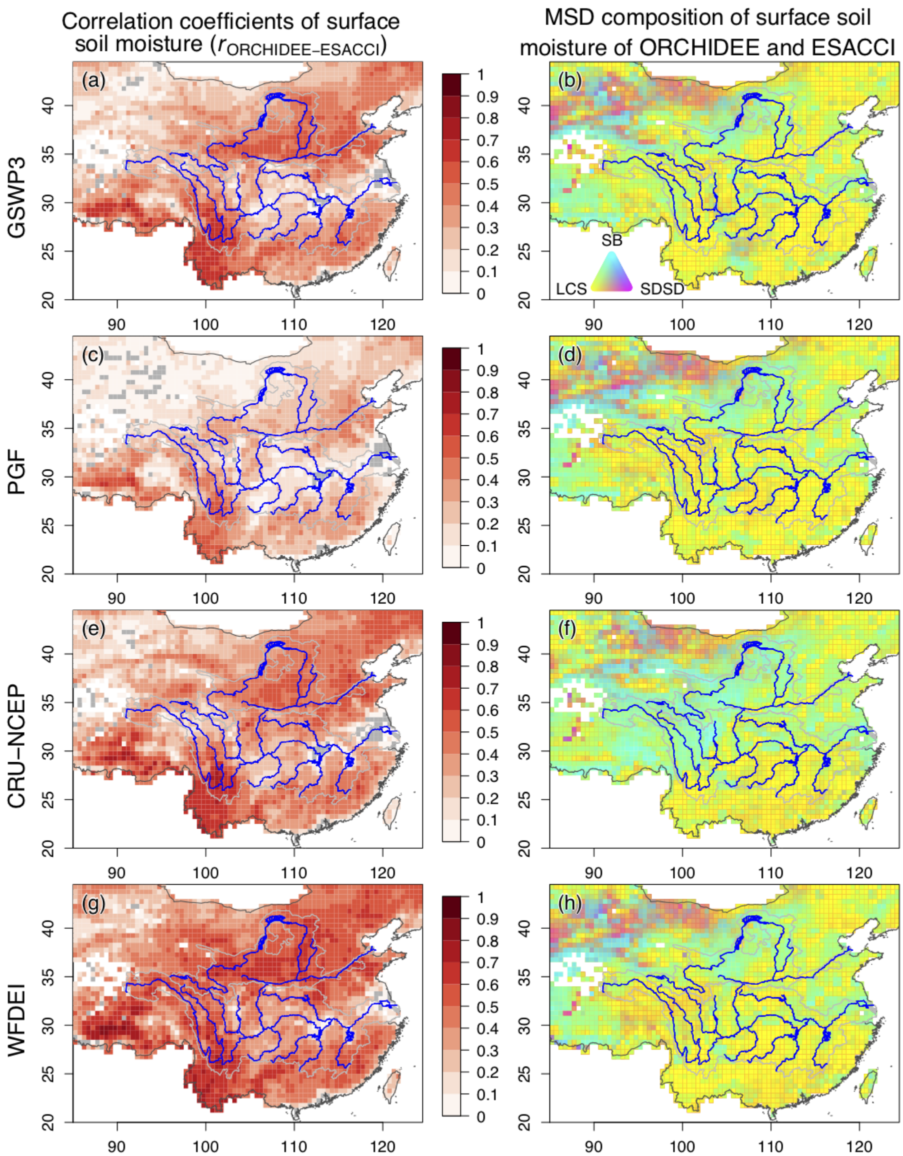

Figure S4. Left panel: Correlation coefficients of the ESA CCI SM and the corresponding ORCHIDEE SM from 1981 to 2009. Gray pixels indicate non and negative correlation. Right panel: decomposition of the MSD between the daily ESA CCI SM and the corresponding ORCHIDEE SM (Eq. 1). Cyan, magenta, and yellow indicate the fractions of SB, SDSD, and LCS respectively.

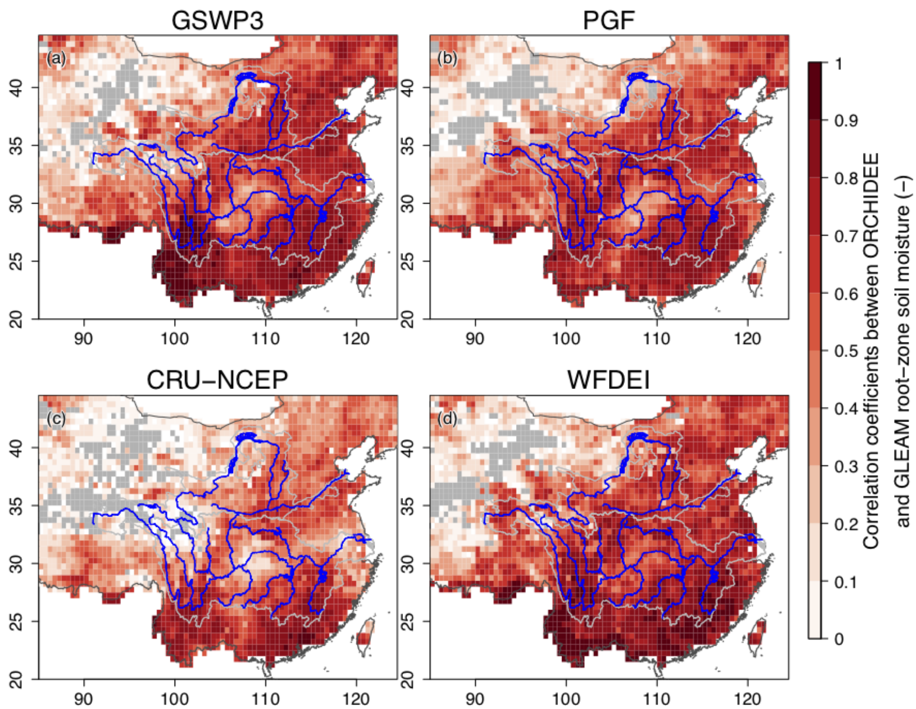

Figure S5. Pearson correlation coefficients between GLEAM root-zone SM and corresponding ORCHIDEE SM. Gray pixels indicate non and negative correlations.

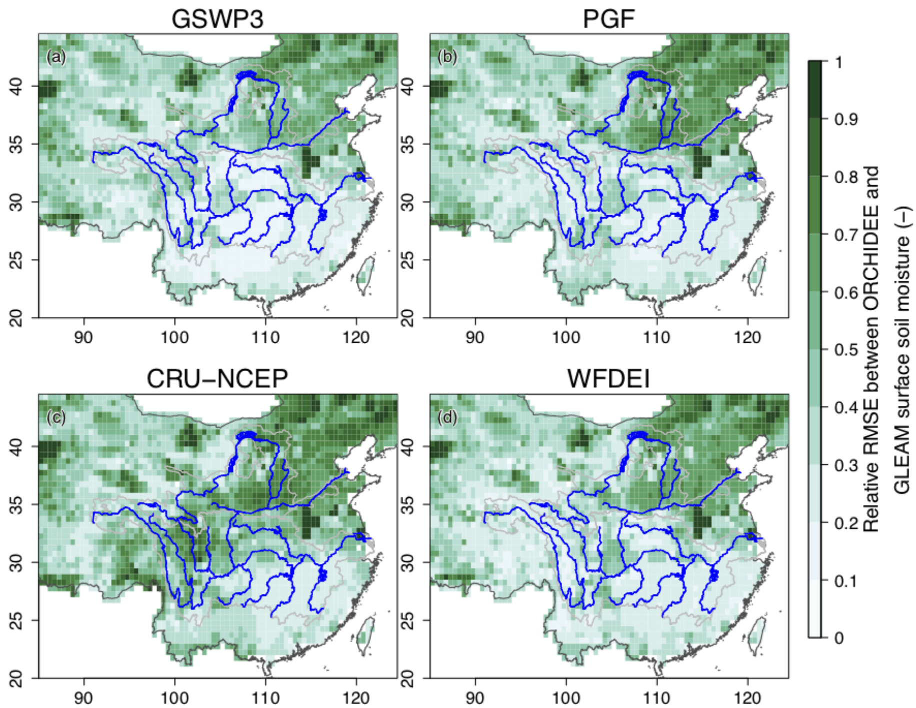

Figure S6. Relative RMSE between the daily GLEAM surface SM and the corresponding ORCHIDEE SM.

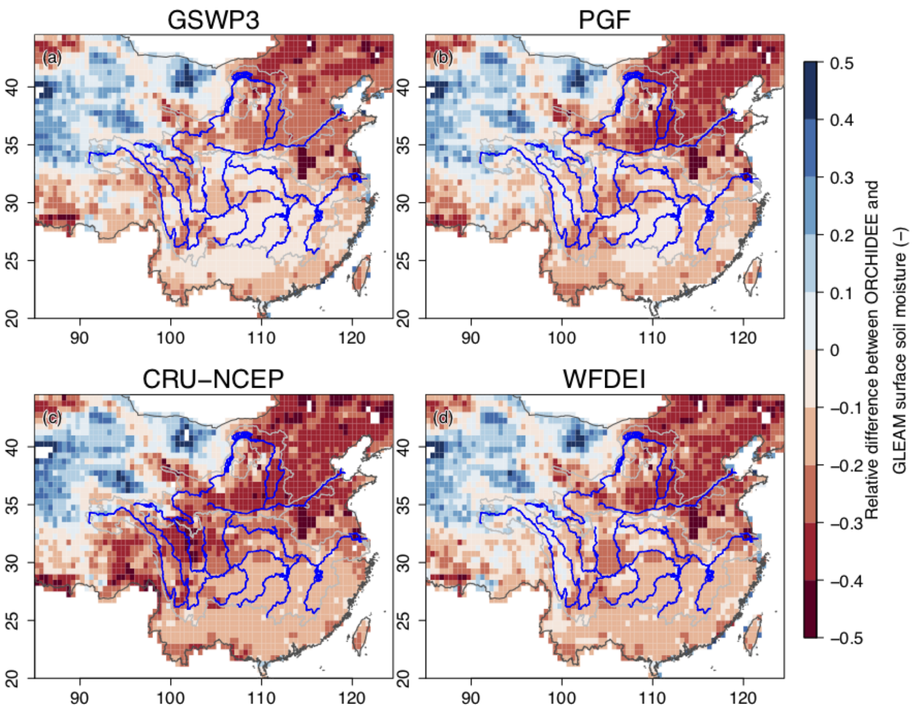

Figure S7. Relative difference ($\Delta_{\text{orc-gleam}}$) between the daily GLEAM surface SM ($\theta_{\text{s,gleam}}$) and the corresponding ORCHIDEE SM ($\theta_{\text{s,orc}}$). The $\Delta_{\text{orc-gleam}}$ is calculated by $(\theta_{\text{s,orc}} - \theta_{\text{s,gleam}})/(\theta_{\text{s,orc}} + \theta_{\text{s,gleam}})$.

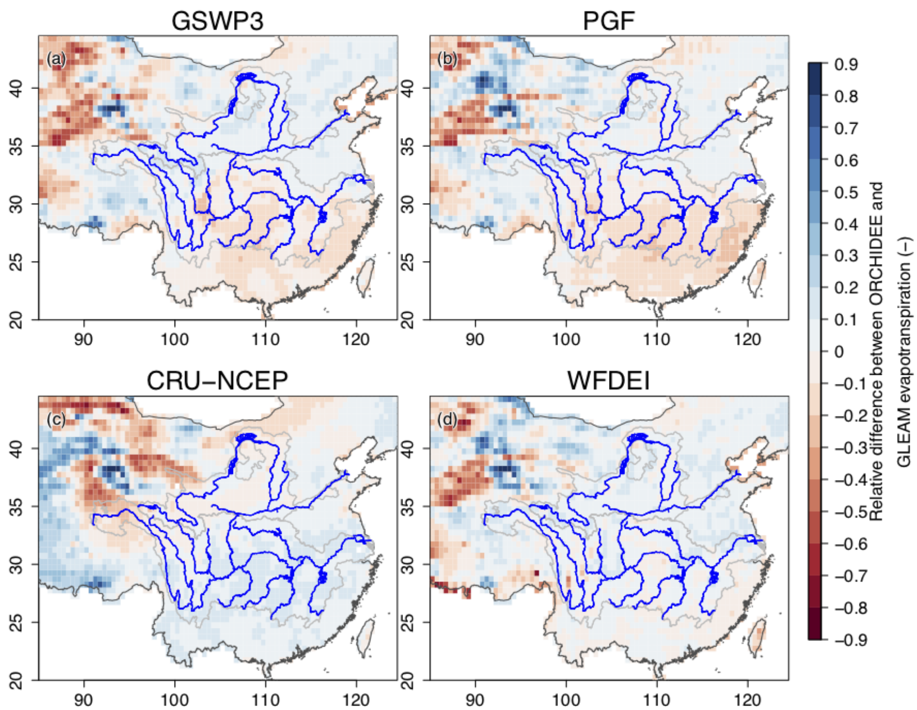

Figure S8. Relative difference between GLEAM and ORCHIDEE evapotranspiration. It is calculated by the same formula as Fig S7.

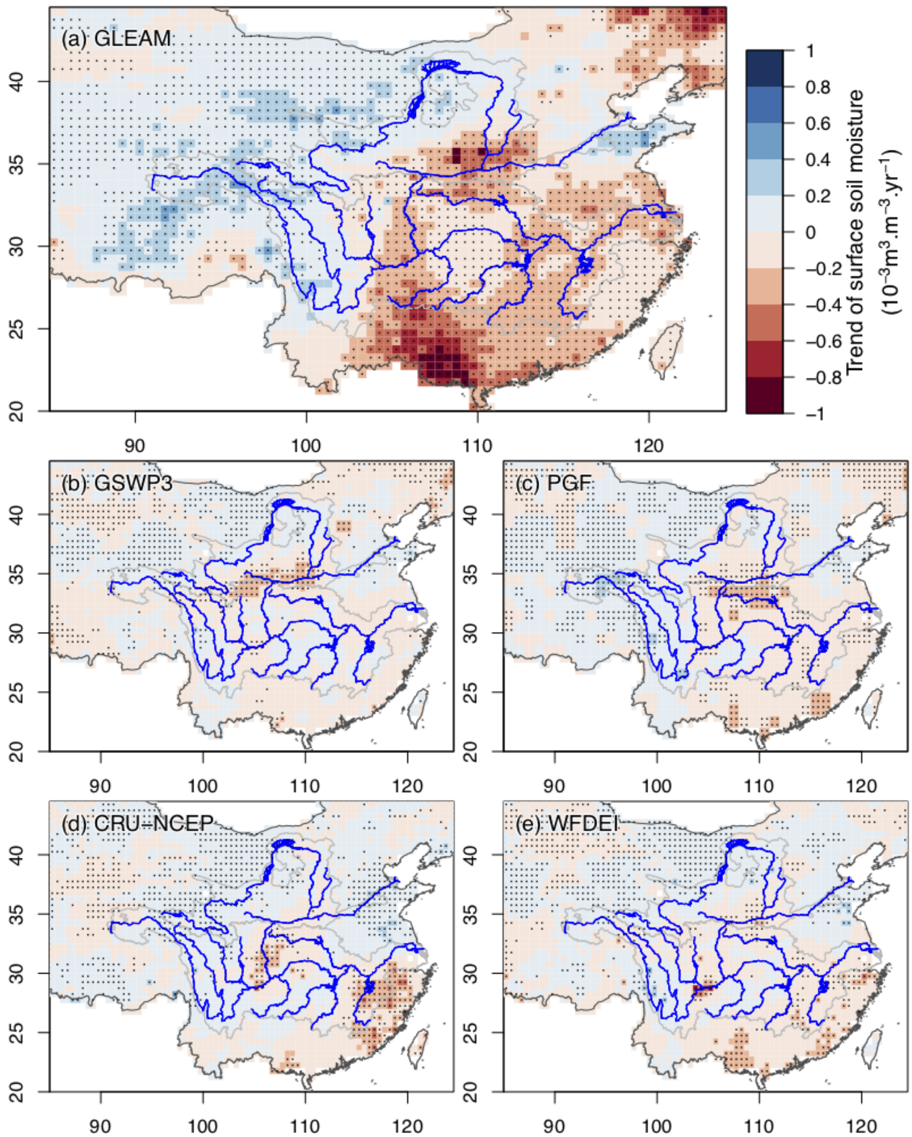

Figure S9. The trend of GLEAM and ORCHIDEE surface SM. Grid cells with dark circles indicate significant trend ($p < 0.05$).

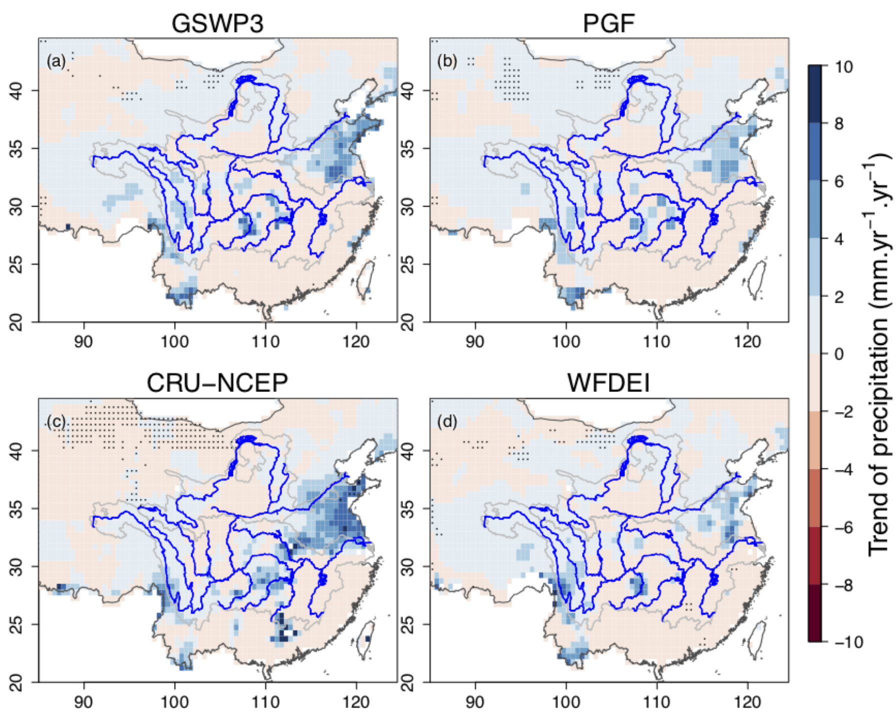

Figure S10. Trend of precipitation form the four forcing datasets. Grid cells with dark circles indicate significant trend ($p < 0.05$).

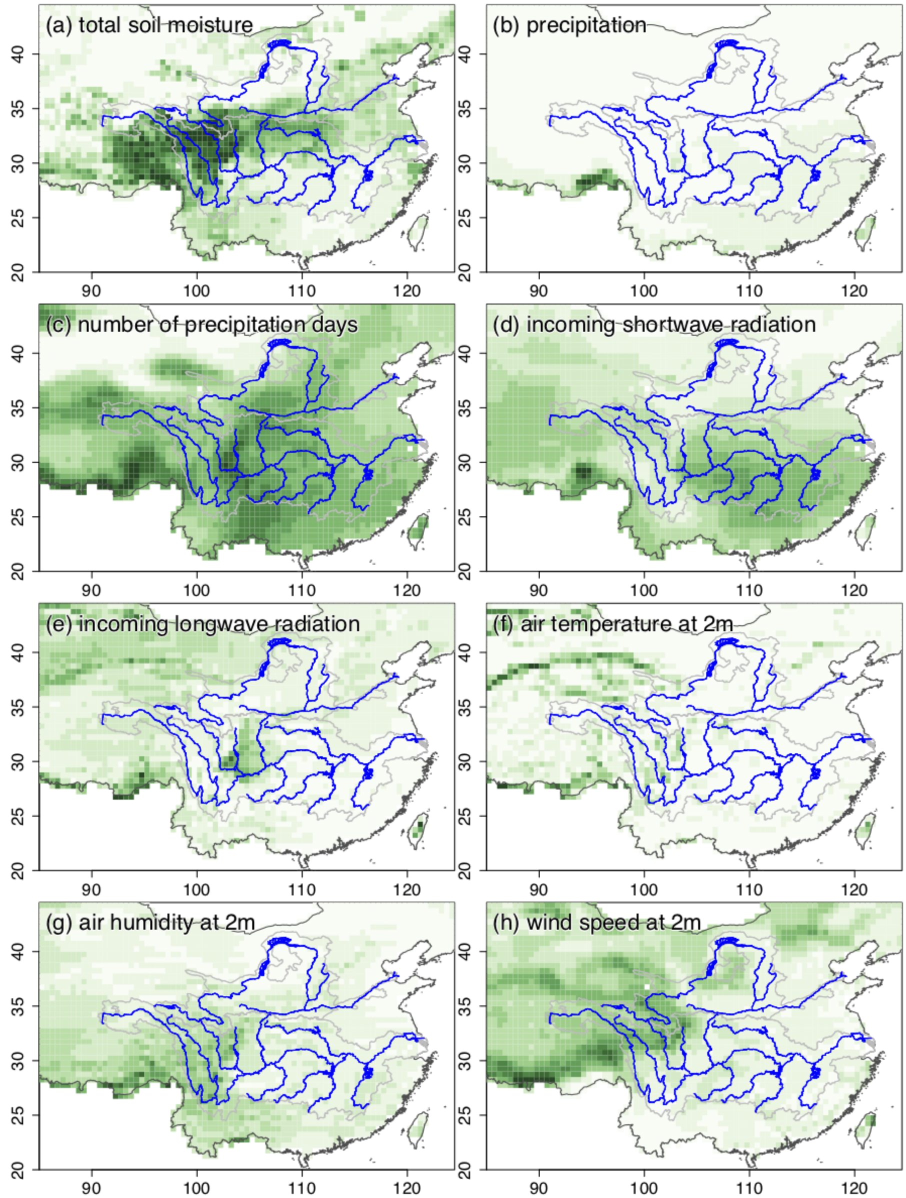

Figure S11. D of simulated SM and meteorological variables. In each grid cell, the monthly time series of specific variable are chosen from the four simulations and the difference is calculated by Eq. 6. The unit of D varies with specific variables. Thus only the spatial patterns can be compared, not the D value.

S1 Why relative difference is not suitable?

In this study we want to investigate the difference of which meteorological variable among the four forcing leads to the difference of simulated soil moisture. To address this question, we first calculated the averaged MSD D_x (see Eq. 6) of meteorological variables and SM. Then the correlation coefficients are estimated between them to discover the highest correlation of D between those meteorological variables and simulated SM. However one question is proposed about the methodology: Why not use the relative D for the analysis to remove potential impacts of magnitude of specific variable on its D? Below, we will discuss and compare the absolute and relative D in a simple example to prove that the relative value is not suitable.

We assume that a variable x from atmospheric forcing dominates the simulated variable y through ORCHIDEE. The relation between x and y can be written as a function:

$$y = f(x). \tag{S1}$$

Then for each grid cell m, there are N observations of x^m from N forcing (one observation from one forcing), recorded as $\{x_i^m\}$. The true value of x in grid cell m is $\hat{x}^m$. We assume that the magnitude of observed error correlates with the magnitude of the true value. Then the $\{x_i^m\}$ can be written as:

$$x_i^m = \hat{x}^m + \hat{x}^m \epsilon_i. \tag{S2}$$

Here ϵ_i is Gaussian noise, which obeys the normal distribution with a mean value of 0 and a uniform standard deviation in all grid cells. Then we have:

$$\sum_{i=1}^{N}\sum_{j=1}^{N}(\epsilon_i - \epsilon_j)^2 = 2N^2\sigma^2. \tag{S3}$$

Here σ^2 is the variance of $\{\epsilon_i\}$. We record the value as E^m:

$$\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}(\epsilon_i - \epsilon_j)^2}{\binom{N}{2}} = E^m\,. \tag{S4}$$

Note that E^m varies among the grid cells, which is distributed according to chi-squared distribution. And it is independent to the magnitude of x. The simulated variable y is calculated by Eq. S1. As the ORCHIDEE is a model, with given input, the output is determined by a set of equatons. So the uncertainty is only transferred from the input, as:

$$y_i^m = f(x_i^m). \tag{S5}$$

Now the averaged SMD of $\{x_i^m\}$ can be calculated as:

$$\begin{aligned}\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left(x_i^m-x_j^m\right)^2}{\binom{N}{2}} &= A\sum_{i=1}^{N}\sum_{j=1}^{N}\left[\hat{x}^m+\hat{x}^m\epsilon_i-\left(\hat{x}^m+\hat{x}^m\epsilon_j\right)\right]^2 \\ &= A\left(\hat{x}^m\right)^2\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\epsilon_i-\epsilon_j\right)^2 \\ &= \left(\hat{x}^m\right)^2 E^m \end{aligned} \tag{S6}$$

Here A is used to temporally record constants. As the same as Eq. S6, the averaged MSD of $\{y_i^m\}$ can be calculated as:

$$\begin{aligned}\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\left(y_i^m-y_j^m\right)^2}{\binom{N}{2}} &= A\sum_{i=1}^{N}\sum_{j=1}^{N}\left[f\left(\hat{x}^m+\hat{x}^m\epsilon_i\right)-f\left(\hat{x}^m+\hat{x}^m\epsilon_j\right)\right]^2 \\ &\approx A\sum_{i=1}^{N}\sum_{j=1}^{N}\left[\left(f\left(\hat{x}^m\right)+f'\left(\hat{x}^m\right)\hat{x}^m\epsilon_i\right)-\left(f\left(\hat{x}^m\right)+f'\left(\hat{x}^m\right)\hat{x}^m\epsilon_j\right)\right]^2 \\ &= A\sum_{i=1}^{N}\sum_{j=1}^{N}\left[f'\left(\hat{x}^m\right)\hat{x}^m\left(\epsilon_i-\epsilon_j\right)\right]^2 \\ &= \left[f'\left(\hat{x}^m\right)\hat{x}^m\right]^2 E^m \end{aligned} \tag{C7}$$

It is clear that the D of output $\{y_i^m\}$ depends on the magnitude of $\{x_i^m\}$, not on the magnitude of output. In other words, the Ds of input and output are correlated due to such uncertainty transport through a model. But the relative Ds, D over the value of specific variable, are probably not correlated as they removed the magnitude information of input and brought extra uncertainty of output.