Geoscientific
Model Development

Supplement of

WAYS v1: a hydrological model for root zone water storage simulation on a global scale

Ganquan Mao and Junguo Liu

Correspondence to: Junguo Liu (junguo.liu@gmail.com, liujg@sustech.edu.cn)

Supplementary Information

Supplementary Section: Parameter Sensitivity Analysis

In the WAYS mode, β and C_e are two crucial parameters that control the partitioning of precipitation into evaporation and runoff, thus affecting the water balance. Due to the incredibly high computation cost, only the sensitivity of the model simulation to these two parameters are tested. First, a pixel is selected randomly from the domain to demonstrate the impacts of parameter perturbation on simulated evaporation, runoff and RZWS. For each experiment, only one parameter is perturbed, and the other one is set to the calibrated value. The calibrated value for β and C_e is 0.17 and 1.67, respectively. The parameter is perturbed within the range randomly 1000 times during the experiment. Simulations are executed from 2009 to 2010 on a daily scale, while the results are shown on a monthly scale (see Figure S1). The model is more sensitive to parameter C_e than parameter β. The uncertainties caused by the parameter C_e are generally larger than those caused by the parameter β, especially for RZWS. These two parameters also have complementary effects on the model simulation, causing larger uncertainties for the simulation than one parameter.

To further investigate the uncertainties stemming from parameters on a global scale, a Monte Carlo simulation of 1000 samples is performed by perturbing the two parameters simultaneously. For both parameters, the normal distribution is used for the Monte Carlo perturbation. Simulations are executed from 2001 to 2010 on a daily scale. The coefficient of variation (CV) for each pixel is then calculated, which reflects the uncertainties (De Graaf et al., 2015). A high value of CV indicates relatively higher uncertainty caused by the parameters, while a low value of CV implies the opposite. Figure S2 shows that parameter-induced uncertainties of evaporation and runoff have similar patterns, while the magnitude is slightly higher for the runoff globally. This finding is consistent with the pixel-based sensitivity test (see Figure S1). The simulated RZWS has the largest uncertainties with the Monte Carlo simulation. Additionally, the uncertainties of RZWS show the opposite trend to the uncertainties of evaporation and runoff. In the northern part of Africa, the Arabian Peninsula, northwest of China and southern part of Australia, the uncertainties in evaporation and runoff are low. However, the uncertainties in RZWS are quite large in these regions.
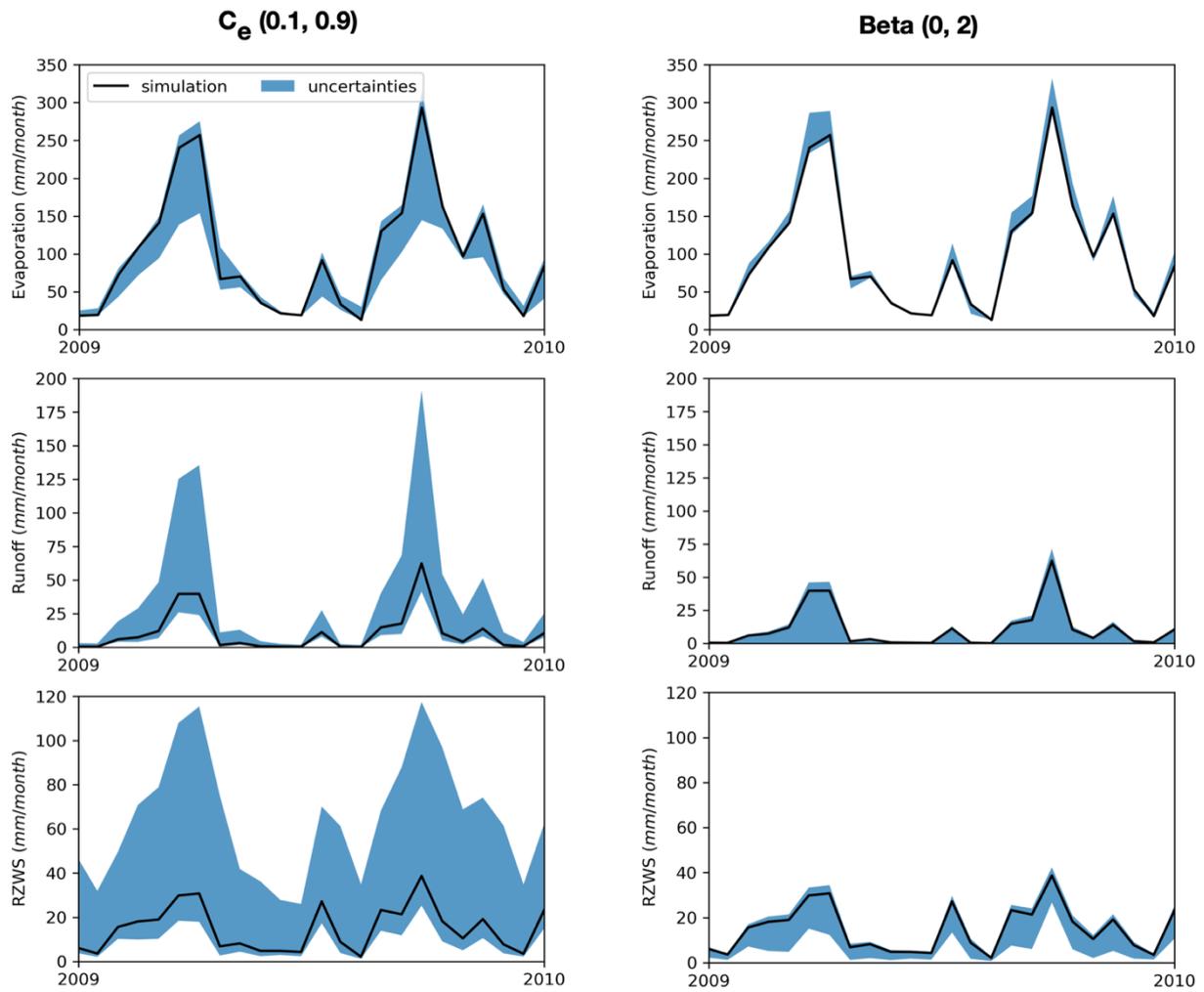
Figure S1. Sensitivity of simulated evaporation (top), runoff (middle) and RZWS (bottom) to parameter β and Ce in a randomly selected pixel within the domain. The black solid line represents the simulation based on the calibrated value. The blue area indicates the uncertainties induced from the perturbation of the parameter 1000 times.
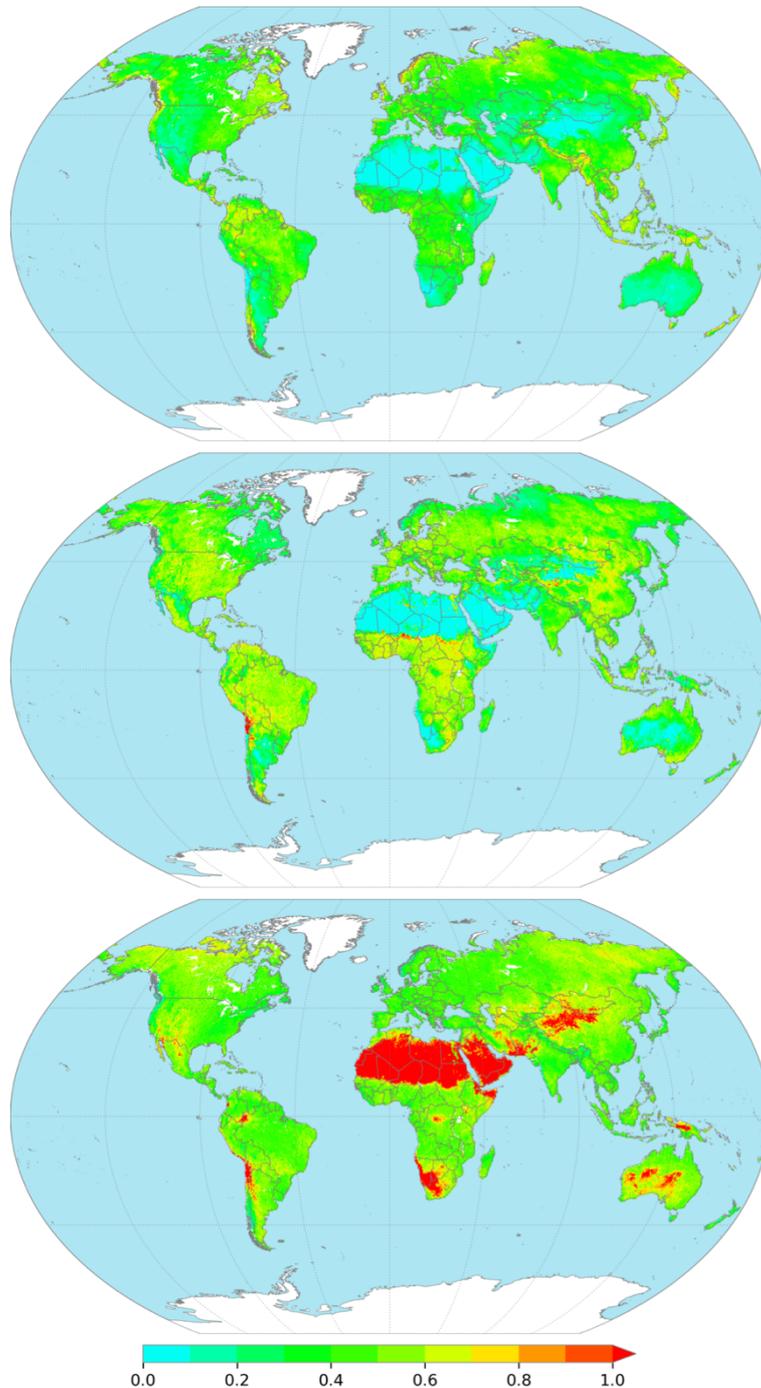
Figure S2. Coefficient of variation of model-simulated evaporation (top), runoff (middle) and RZWS (bottom) from 1000 Monte Carlo simulations with different parameter settings for β and Ce.

Since the groundwater recharge module in WAYS is based on the work of Döll and Fiedler (2008), the values are taken directly from it. These values are also used for other global groundwater recharge simulation-related works, e.g., Müller Schmied et al. (2014). However, as these are indeed empirical parameter values, uncertainty/sensitivity analysis is necessary for $R_{s,max}$. Three pixels from different soil types are selected for the $R_{s,max}$-induced uncertainty investigation. Figure S3 shows the grouped soil texture classes for this study based on the FAO Harmonized World Soil Database and the selected pixels for the

uncertainty analysis. Pixel 1, pixel 2 and pixel 3 represent the soil type of clay, loam and sand, respectively.

$R_{s,max}$ (mm/day) directly influences the matrix flow (contributes 100% to groundwater recharge with a certain time lag) based on equation 12 in Table 1 in the manuscript, as it controls the maximum groundwater recharge for different soil types. Consequently, it will also impact the preferential flow, as the runoff is partially split into matrix flow and the rest to the preferential flow. Therefore, parameter $R_{s,max}$ will have light effects on the runoff generation but could have considerable impacts on the matrix flow and preferential flow. Thus, the sensitivity of the simulated preferential flow and matrix flow to the maximum groundwater recharge $R_{s,max}$ is investigated, and the results are shown in Figure S4 (pixel with clayey soil), Figure S5 (pixel with loamy soil) and Figure S6 (pixel with sandy soil). The sensitivities of WAYS to $R_{s,max}$ are checked by perturbing the parameter. We set the simulation with soil texture-specified $R_{s,max}$ as the control run, and perturbed $R_{s,max}$ by -80%, -50%, -20%, 20%, 50% and 80% for the sensitivity test. Figure S4 (bottom plot) shows the impacts of the values of $R_{s,max}$ on the simulated daily matrix flow with the soil type of clay. With the increase in $R_{s,max}$, the simulated daily matrix flow has a higher peak, while the opposite is observed with the decrease in $R_{s,max}$. It changes the scale of the simulated matrix flow but not its shape at the daily scale. Moreover, due to the change in daily simulation, the monthly simulation of the matrix and preferential flow are affected accordingly, as seen in Figure S4 (top and middle plots). The results show that parameter $R_{s,max}$ has opposite impacts on preferential flow and matrix flow, which is logical because both are part of the runoff. A similar phenomenon is found in daily simulated time series. Thus, for pixel 2 and pixel 3, only monthly simulated matrix and preferential flow are shown to visualize the uncertainties stemming from $R_{s,max}$ for loamy and sandy soil. The simulated matrix flow time series with a decreased value of $R_{s,max}$ shows are found to have larger uncertainties than time series with an increased value of $R_{s,max}$, because the maximum value of matrix flow is not only determined by $R_{s,max}$ but also the groundwater recharge factor f_s.
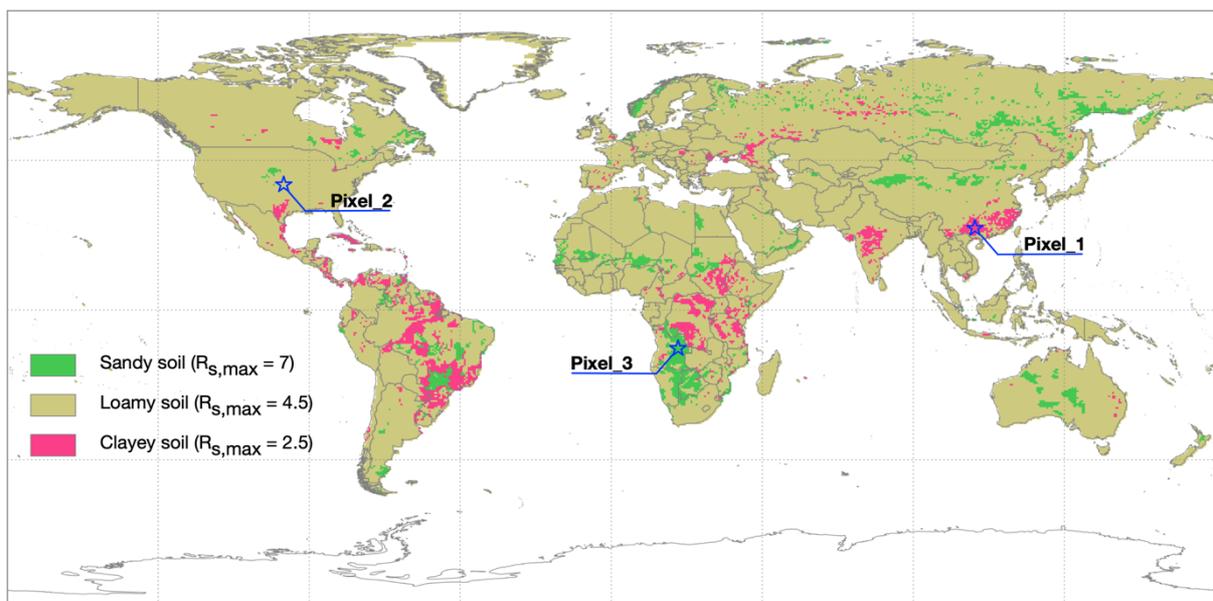
Figure S3. Grouped soil texture classes for the study based on the FAO Harmonized World Soil Database.
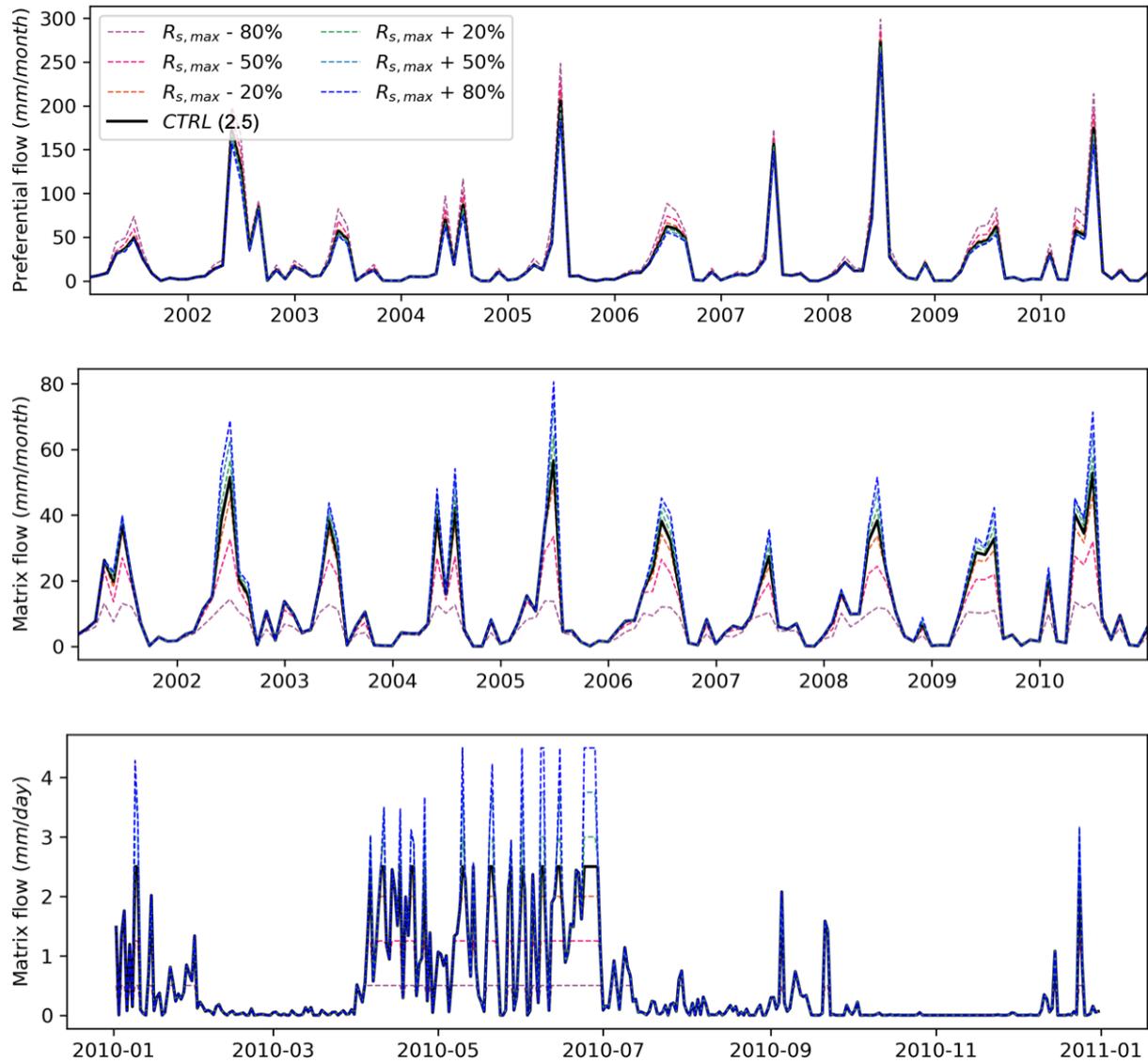
Figure S4. Sensitivity of the simulated preferential flow and matrix flow to the maximum groundwater recharge $R_{s,max}$ for the pixel with clayey soil.
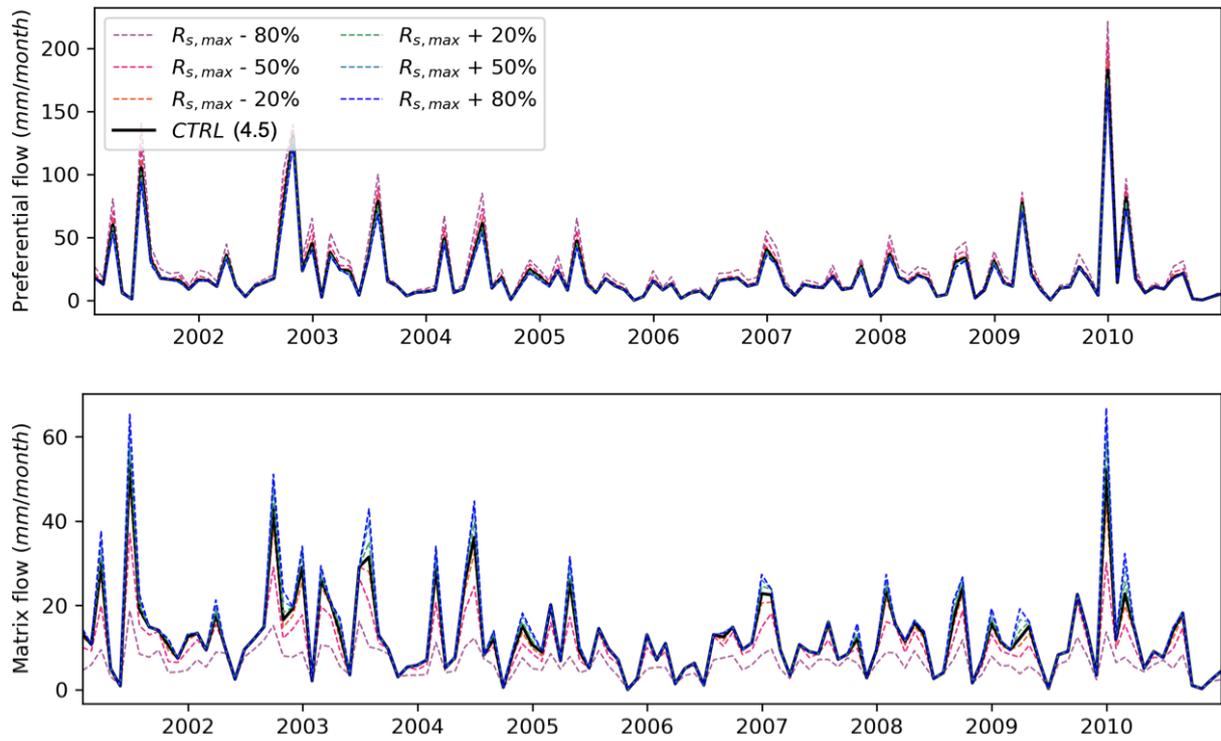
Figure S5. Sensitivity of the simulated preferential flow and matrix flow to the maximum groundwater recharge $R_{s,max}$ for the pixel with loamy soil.
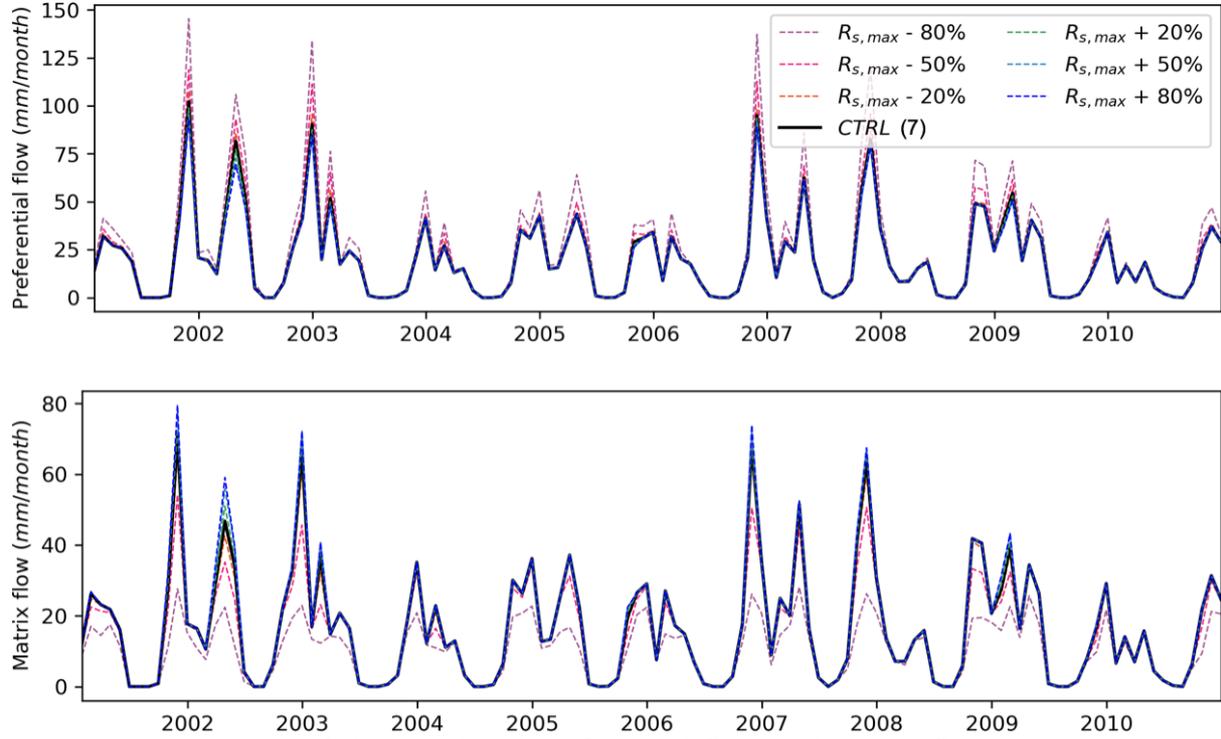
Figure S6. Sensitivity of the simulated preferential flow and matrix flow to the maximum groundwater recharge $R_{s,max}$ for the pixel with sandy soil.

Supplementary Section: Model Evaluation against GRDC observed discharge

To further evaluate the model performance, we have evaluated our results with observed discharges from the Global Runoff Data Centre (GRDC). The CaMa-flood model is the only available open-source global runoff routing model (http://hydro.iis.u-tokyo.ac.jp/~yamadai/cama-flood/) that is capable of simulating backwater effects, which is important for plain regions, making it a popular choice for many studies (Hirabayashi et al., 2013; Mateo et al., 2014; Pappenberger et al., 2012).

The GRDC stations along a river were selected with interstation areas larger than 7000 km^2 to omit catchments with hydrological processes that are not properly represented by global hydrological models operating at a 0.5° resolution (Hunger and Döll, 2008). In total, 154 stations are selected for major river basins worldwide. For discharge simulation, the CaMa-flood is run at a 0.5° resolution to maintain consistency with the WAYS simulated runoff. The WAYS_CRU simulation is used for routing due to global coverage of the data. The discharge is simulated for the 1971-2010 period.

For the evaluation, the simulated discharge is compared with the GRDC data at each selected station depending on the data availability. Since the observations provided by GRDC are on a monthly time scale, the simulated data are also aggregated to the monthly scale for the comparison. The correlation coefficient and Nash-Sutcliffe efficiency coefficient are calculated, while the correlation coefficient between the simulated discharges and GRDC station records are visualized in Figure S7.
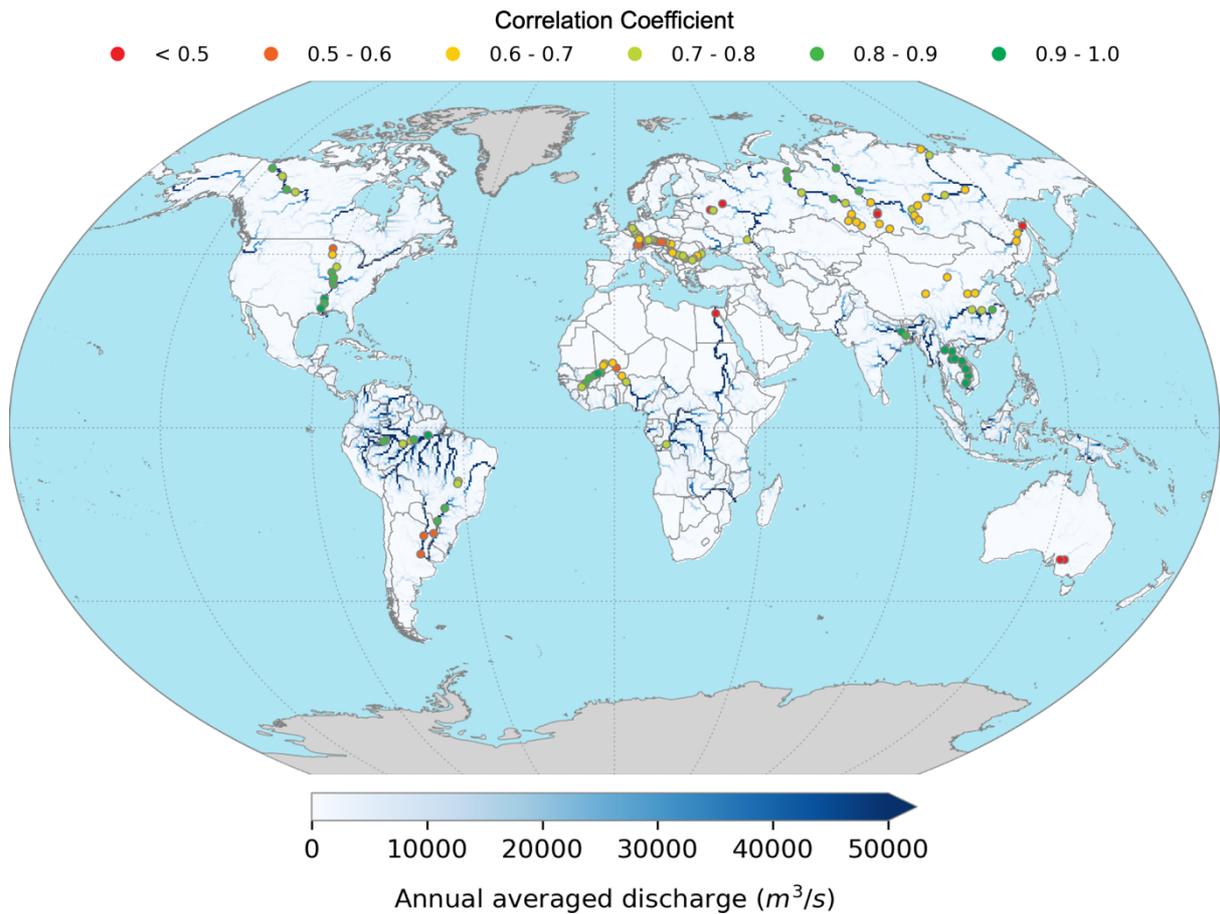
Figure S7. The evaluation of simulated discharge by comparison with the GRDC observations. The discharge is simulated by the CaMa-flood model, and the WAYS simulated runoff based on the RZSC data ($S_{R,CRU-SM}$) is used as the input data for routing. The background of the figure is the annual averaged discharge for the 1971-2010 period. The point indicates the correlation coefficient between simulated discharge and GRDC observations. The location of the points implies the location of the GRDC station. Different colors at the points represent the magnitudes of the correlation coefficient.

The simulation shows a generally good correlation with the GRDC observations, while poor performance in the discharge simulation is also found in a few stations. The errors between the simulated discharge and observations could be caused by both the WAYS model for runoff simulation as well as the CaMa-Flood model for runoff routing, as the CaMa-Flood model itself also shows different performances in basins across the world (Yamazaki et al., 2011). The relatively low performance of WAYS is found in middle-high latitudes compared with low-middle latitude regions. This result could be explained by the relatively simple snow-melt module in the WAYS model, which thus could consequently produce low-quality runoff for river routing in cold regions. In Australia, only two GRDC stations in the Murray Darling basin are selected for the evaluation, and the correlation coefficient between simulated discharge and GRDC station is less than 0.5, indicating the large difference between them.

Figure S8 shows the histogram of the data points within different intervals of the correlation coefficient. Only in 7.2% of the stations are the correlations between simulation and

observation less than 0.5. For more than half the stations, the correlations are higher than 0.7. The results show a generally good correspondence between the simulated and observed discharge. The generally good performance in the discharge simulation confirms the strong capacity of WAYS for runoff generation.
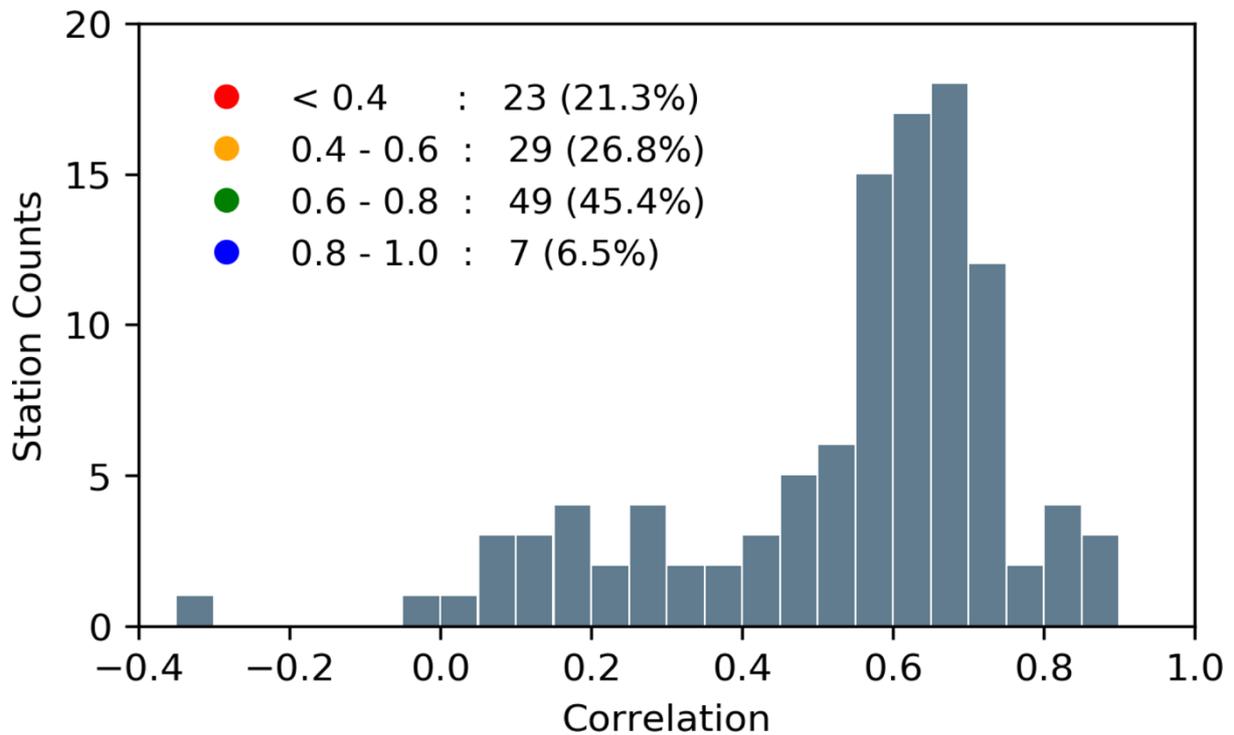
Figure S8. Histogram showing the percentage of data points within different intervals of the correlation coefficient.

Supplementary Section: Evaporation Evaluation against LandFluxEVAL data

The model simulation is further compared to a gridded data set, LandFluxEVAL data, for evaporation evaluation. The LandFluxEVAL data are a merged benchmark synthesis product of evaporation on a global scale and a combination of land-surface model simulations, remote sensing products, reanalysis data and ground observation data (Mueller et al., 2013). The LandFluxEVAL data are used in many studies as reference data for evaporation evaluations (Lorenz et al., 2014; Martens et al., 2017; Wartenburger et al., 2018). Since the LandFluxEVAL data are only available at 1-degree spatial resolution, the WAYS simulated evaporation is aggregated to 1 degree to match the resolution of the reference data. The evaluation is executed for 1989-2005 based on the availability of the LandFluxEVAL data. For the spatial evaluation, the WAYS simulation based on RZSC ($S_{R,CRU-SM}$) is used due to the global coverage of the RZSC product. For latitudinal comparison, both runs of WAYS simulated evaporation are used.

A promising relationship between WAYS simulated evaporation and LandFluxEVAL evaporation is found both in spatial pattern and in latitudinal average (see Figure S9). The generally high correlation coefficient (Figure S9, a) confirms the good performance of the WAYS model. However, relatively poor performance is also found in some regions in

Europe, North America and South America (Amazon basin). It can also be seen that the spatial pattern of WAYS simulated annual averaged evaporation follows that of LandFluxEVAL data, while overestimations are found in regions, e.g., the Amazon basin and southeast Asia. The latitudinal evaluation shows that both WAYS simulations (WAYS_CRU and WAYS_CHIRPS) display a slight overestimation.
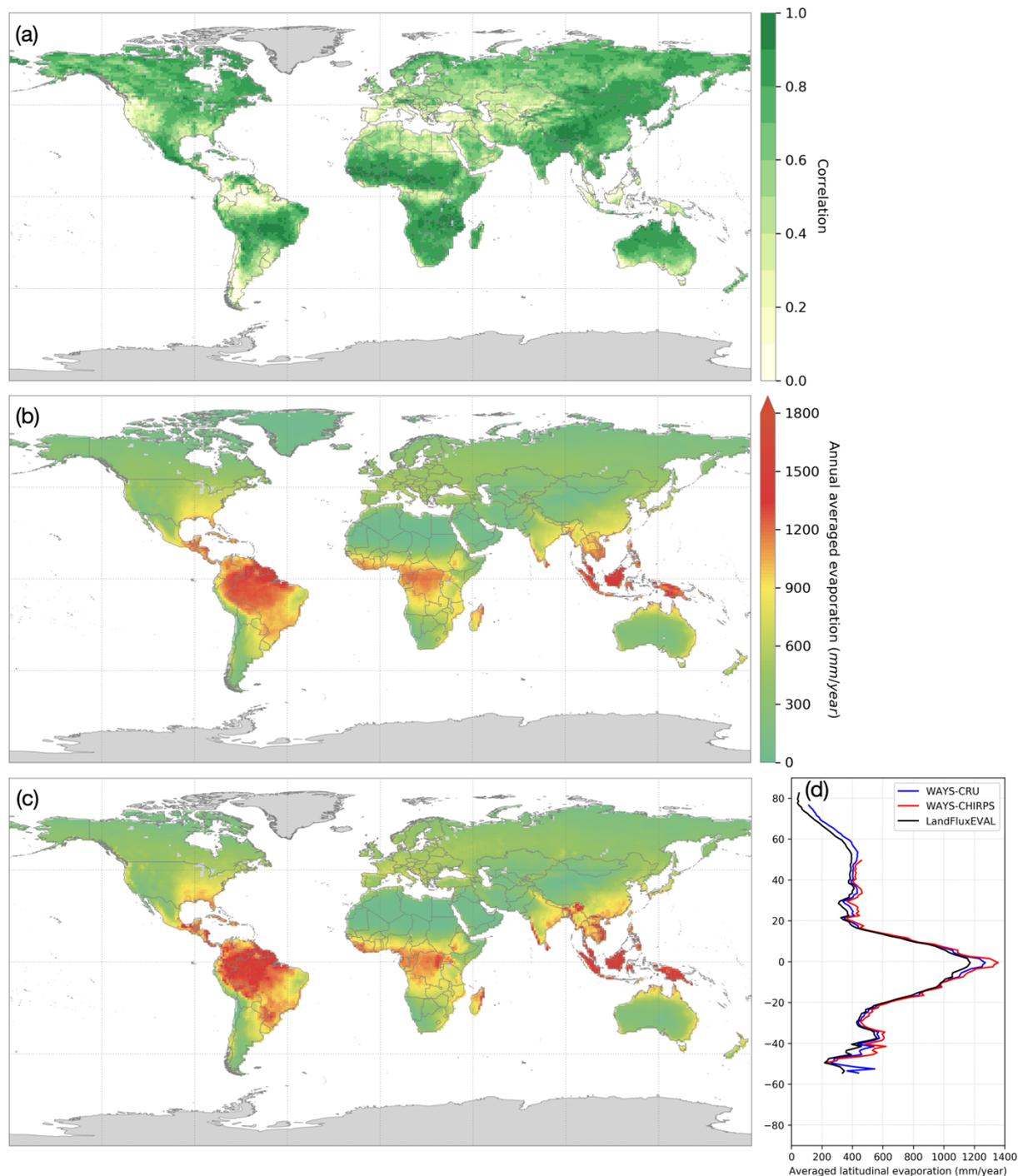
Figure S9. Validation results of the evaporation of WAYS simulation against the LandFluxEVAL data (1989-2005). (a) The calculated correlation coefficient between LandFluxEVAL data and WAYS simulation, (b) the annual averaged evaporation of LandFluxEVAL data, (c) the annual averaged evaporation of WAYS simulation based on RZSC $S_{R,CRU-SM}$, and (d) the comparison of the averaged latitudinal evaporation for WAYS model runs as well as the LandFluxEVAL data.

Supplementary Section: Investigation of the spatial pattern of simulated RZWS

Figure S10 shows the spatial distribution of the annual averaged RZWS simulated by WAYS in the 1971-2010 period. It shows that RZWS is high in low-middle latitudes, while RZWS in middle-high regions is relatively low. RZWS represents the water content that is stored in the root zone as well as the available water for plants. Therefore, for lower or middle latitudes, the available water for plants is relative higher than for high-latitude areas.

To further investigate the soil water condition, we have calculated the root zone soil wetness by dividing RZWS by RZSC, and the results are shown in Figure S11. The root zone soil wetness follows the spatial patterns of RZWS in general. However, differences can also be found in regions such as Europe, South America and the eastern part of North America.
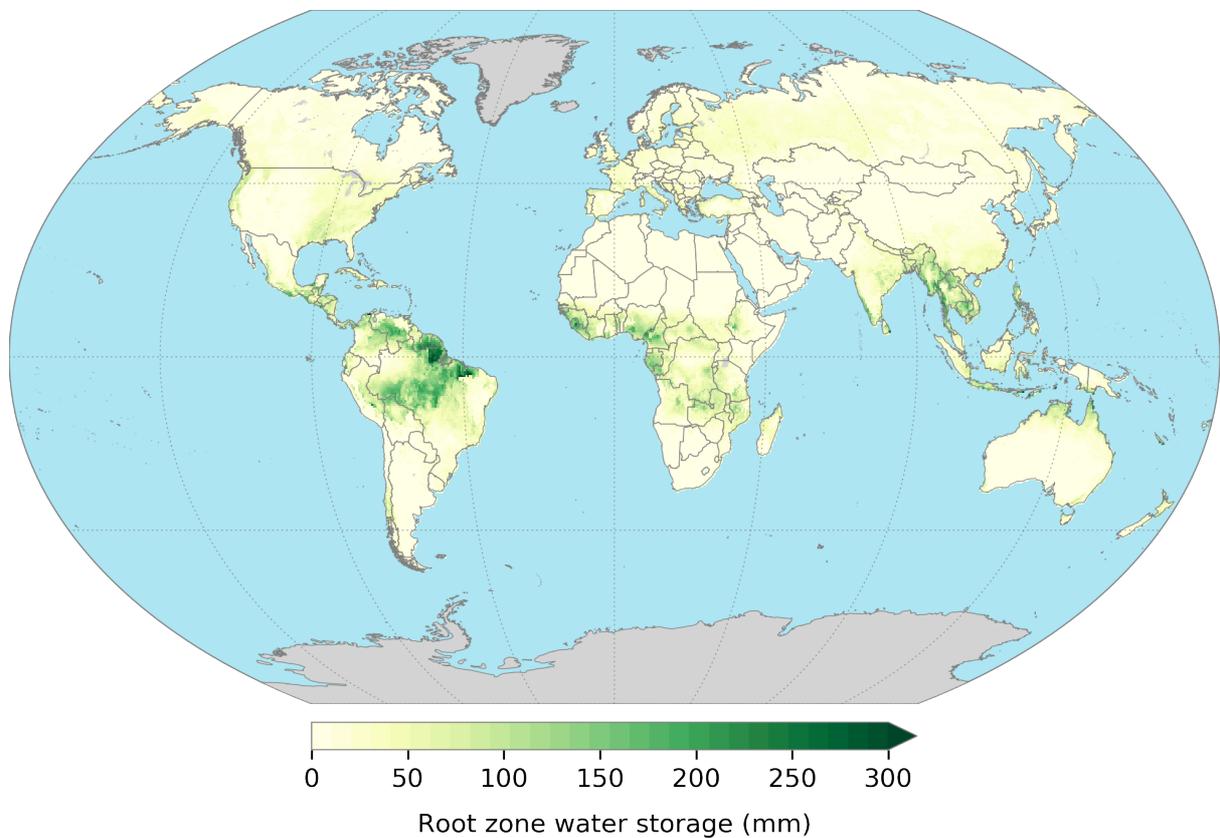
Figure S10. The spatial distribution of the annual averaged RZWS simulated by WAYS in the 1971-2010 period.
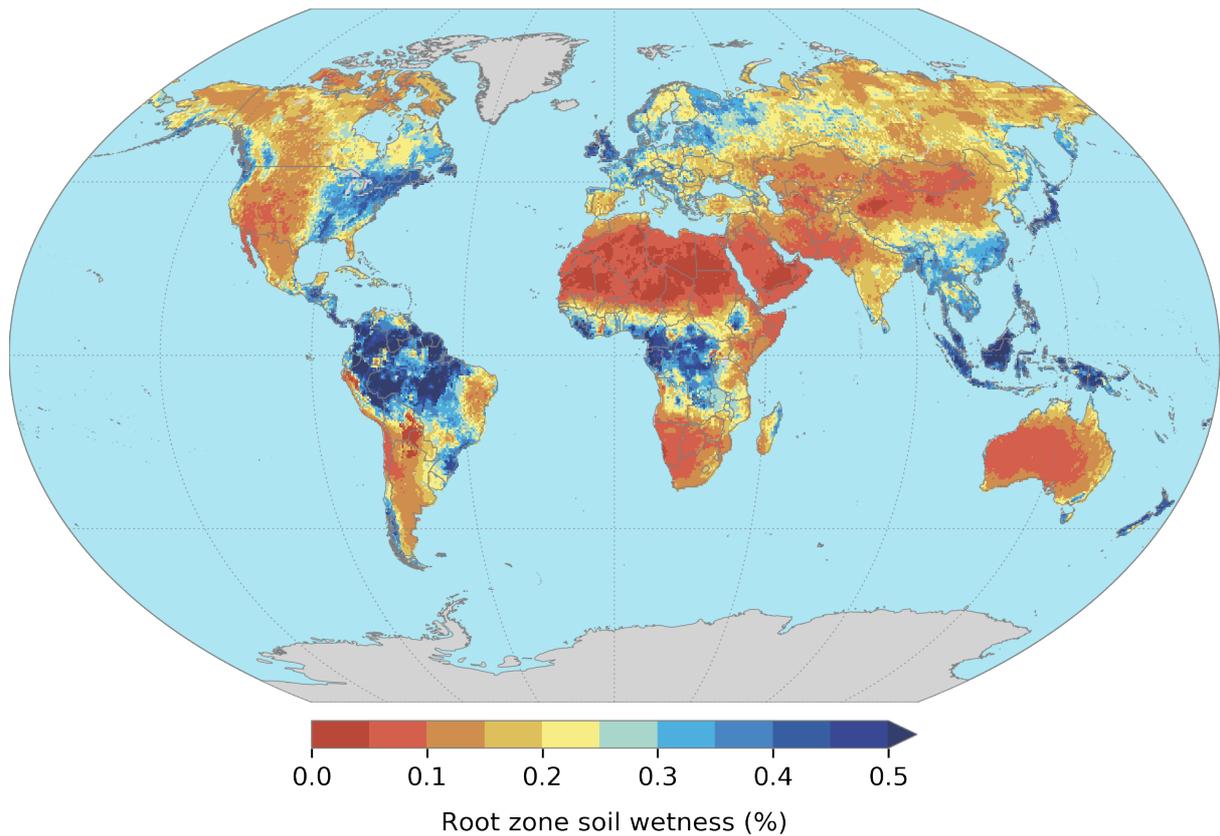
Figure S11. The spatial distribution of the annual averaged root zone soil wetness in the 1971-2010 period.

The simulated RZWS is shown in Figure S12 to reveal dynamics in different latitudes and in different months. The latitudinal averaged RZWS again confirms that the RZWS are relatively plentiful compared with the high latitudes. However, a decreasing trend can also be found by moving from a low latitude to the equator. Two simulations of the WAYS model show similar fluctuation along the latitudes, while the simulation with $S_{R,CHIRPS}$ is slightly higher. Figure S12 shows that RZWS in low-middle latitudes has a larger monthly variation than in other regions, while the Northern Hemisphere and Southern Hemisphere show opposite changing trends. In the low latitudes in Northern Hemisphere, the RZWS peak occurs in May-June and the off-peak in October-November. In the Southern Hemisphere, the RZWS off-peak occurs in May-June and the peak in October-November.
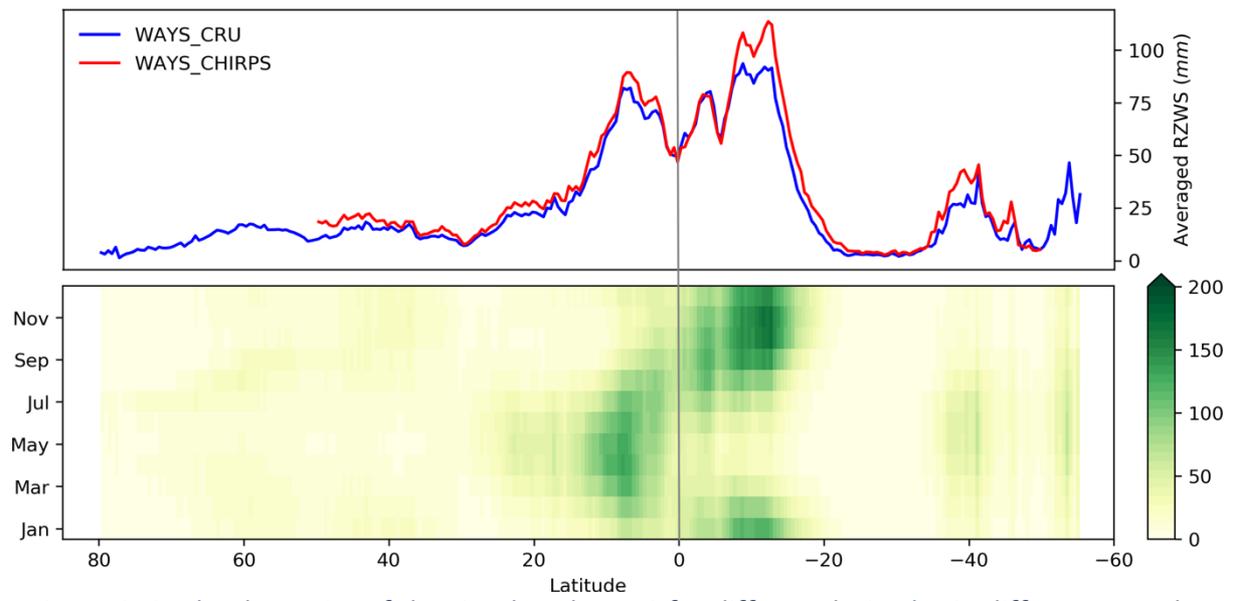
Figure S12. The dynamics of the simulated RZWS for different latitudes in different months.

(a)

(b)

Figure S13. Two global root zone storage capacity products at 0.5 degrees: (a) $S_{R,CRU-SM}$; (b) $S_{R,CHIRPS-CSM}$. Figures are produced based on the data provided by Wang-Erlandsson et al. (2016). Gray color indicates no data.

Figure S14. Mean latitudinal root zone storage capacity of $S_{R,CRU-SM}$ and $S_{R,CHIRPS-CSM}$

Figure S15. Histogram showing the percentage of data points within different intervals of the correlation coefficient.
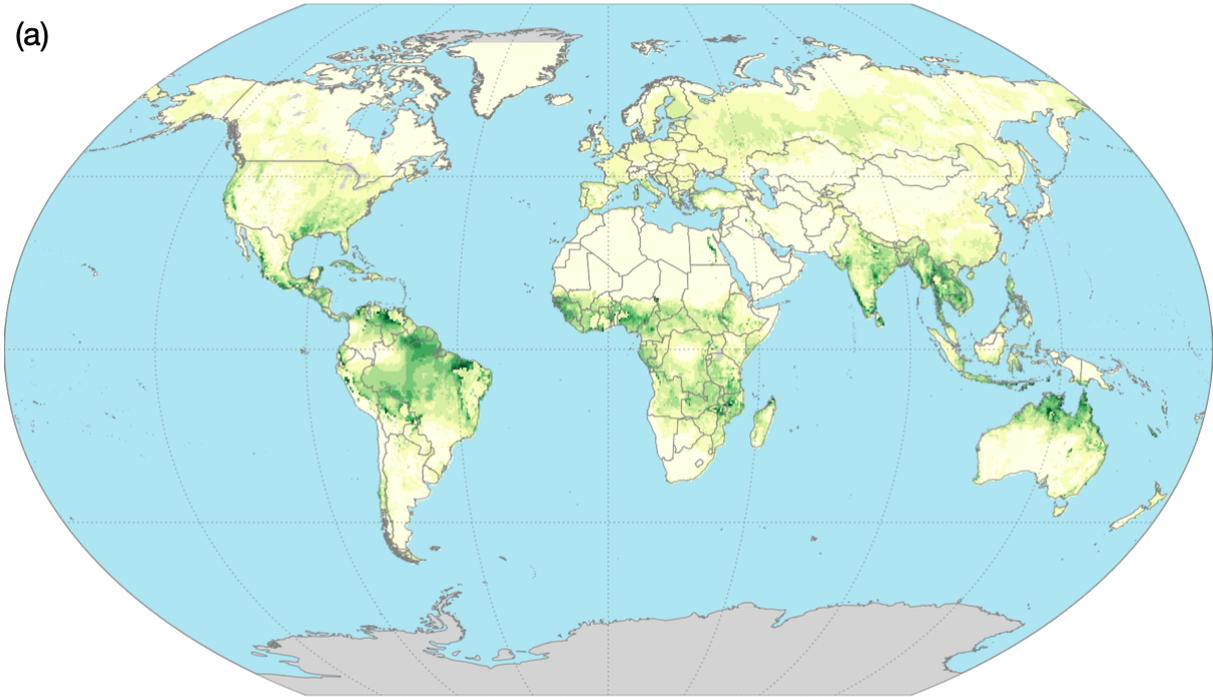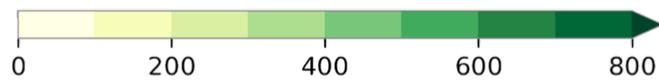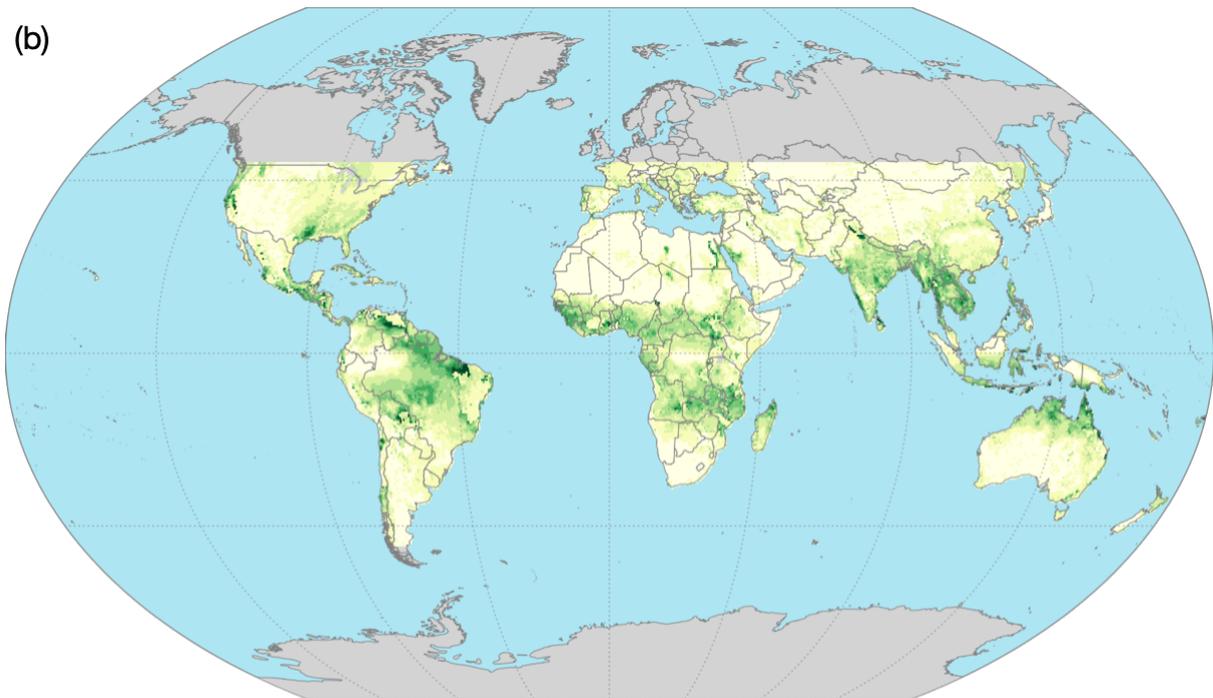
Root Zone Storage Capacity (mm)
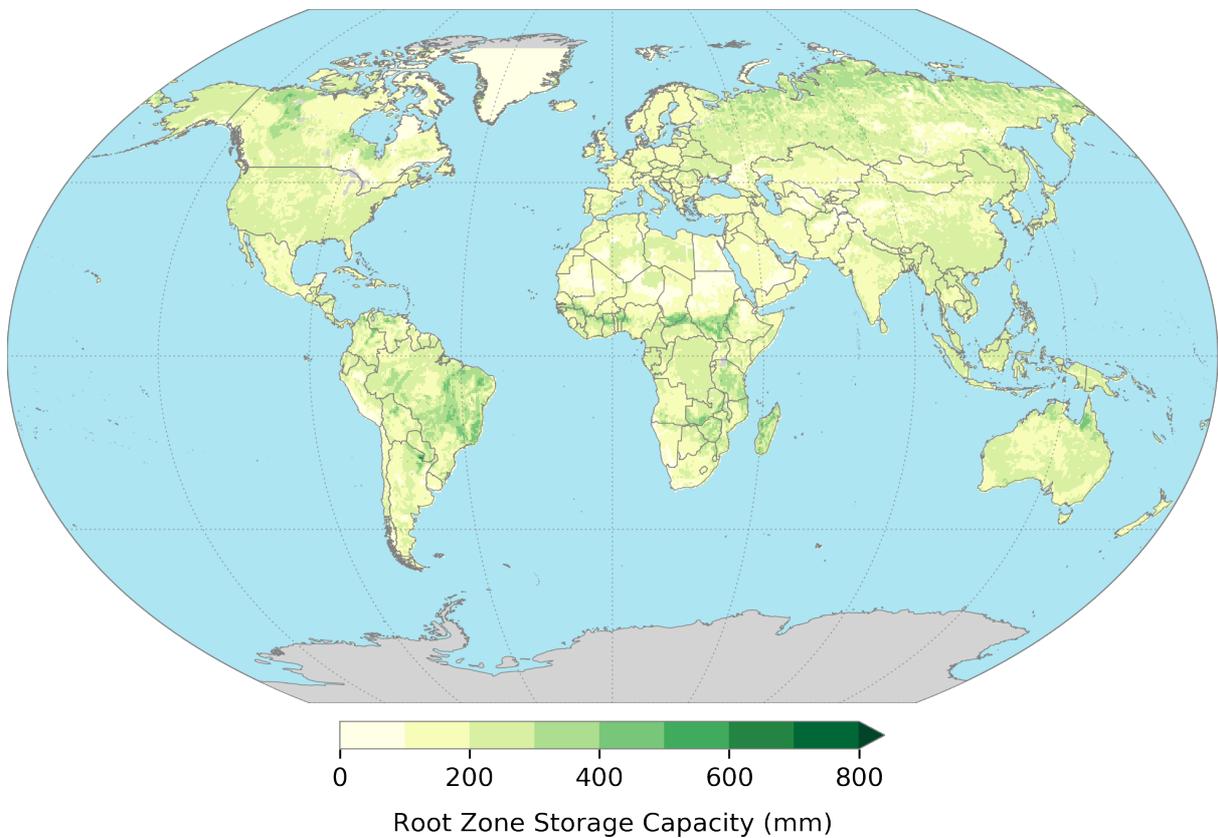
Figure S16. Spatial distribution of uncertain RZSC (SR,LOOKUP-TABLE).

Figure S17. The difference between $S_{R,CRU-SM}$ and $S_{R,LOOKUP-TABLE}$ ($S_{R,LOOKUP-TABLE}$ - $S_{R,CRU-SM}$).
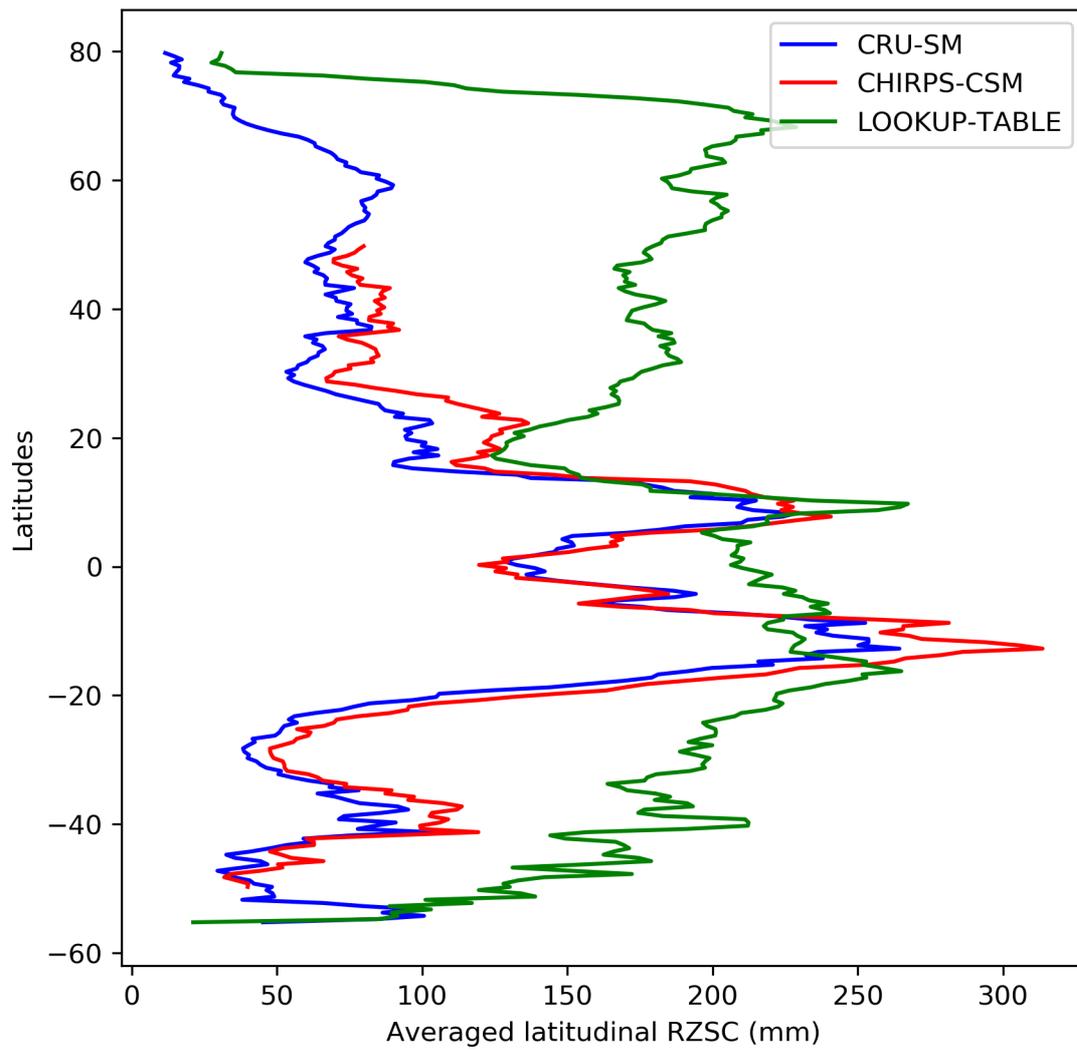
Figure S18. Latitudinal averaged RZSC of different products.
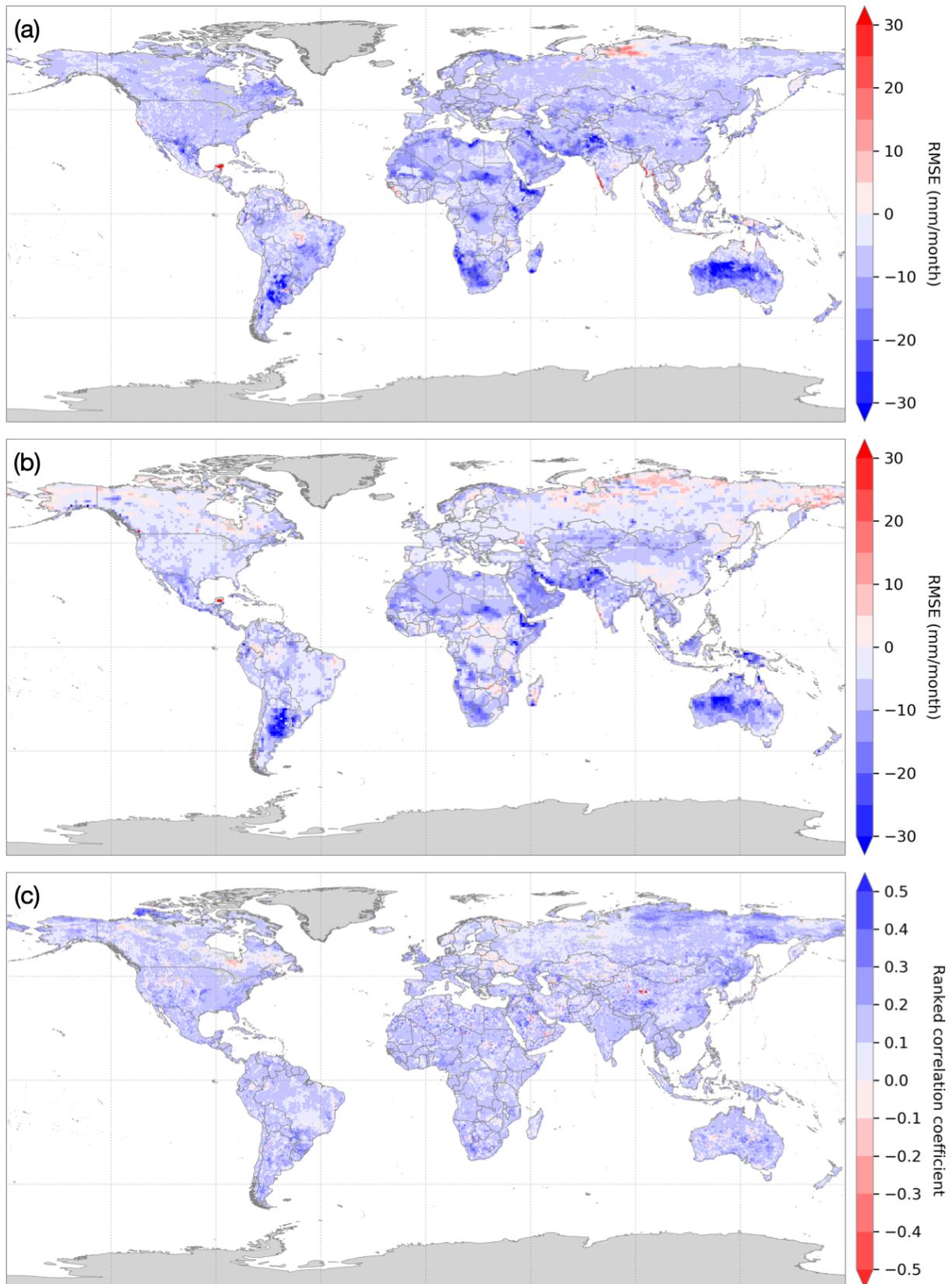
Figure S19. The impacts of RZSC on the model simulation. Blue color indicates the improvement of the simulated results by replacing the uncertain RZSC ($S_{R,LOOKUP-TABLE}$) with satellite data-derived RZSC ($S_{R,CRU-SM}$), while red color implies the opposite. (a) The result for runoff and the reference data for comparison is ERA-Interim/Land data (2001-2010, monthly), (b) the result for evaporation and the reference data for comparison is LandFluxEVAL data

(1989-2005, monthly), and (c) the result for RZWS and the reference data is NDII data (2001-2010, 8 days).
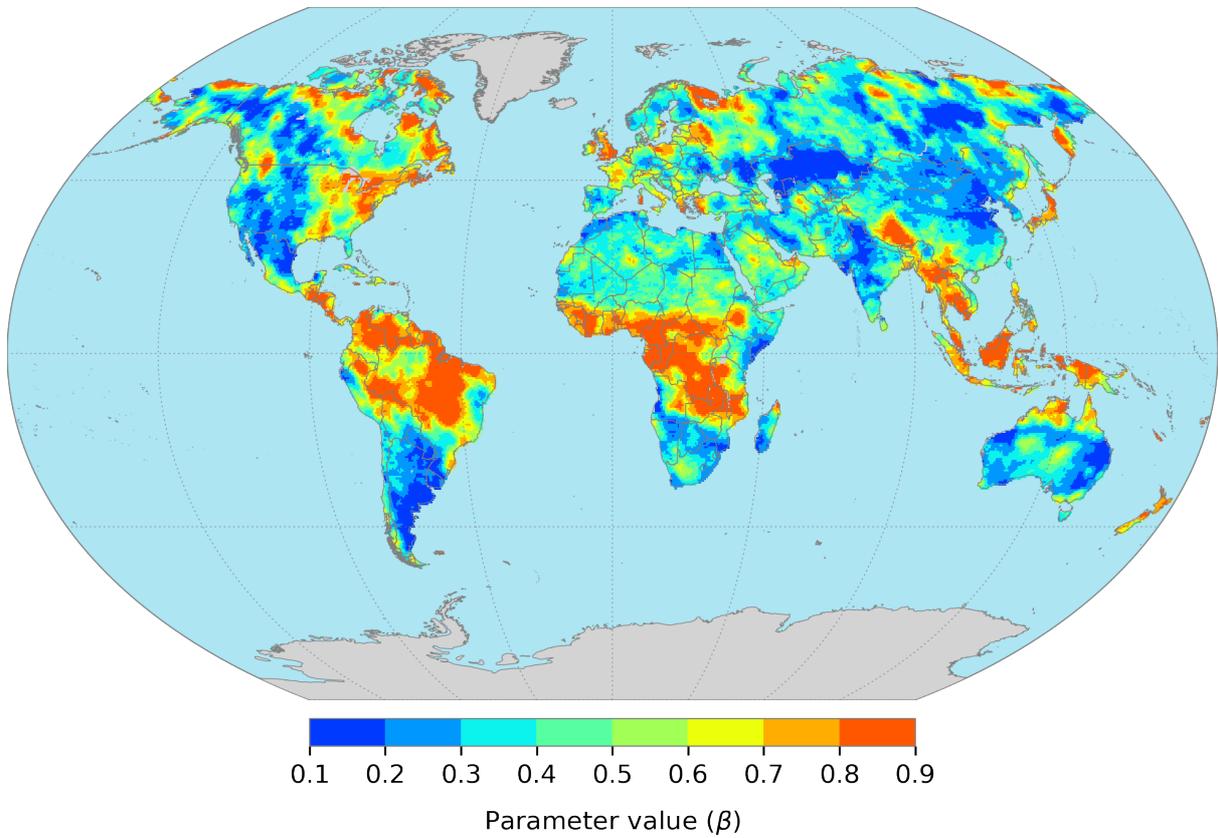
Figure S20. The spatial distribution of model parameter β.
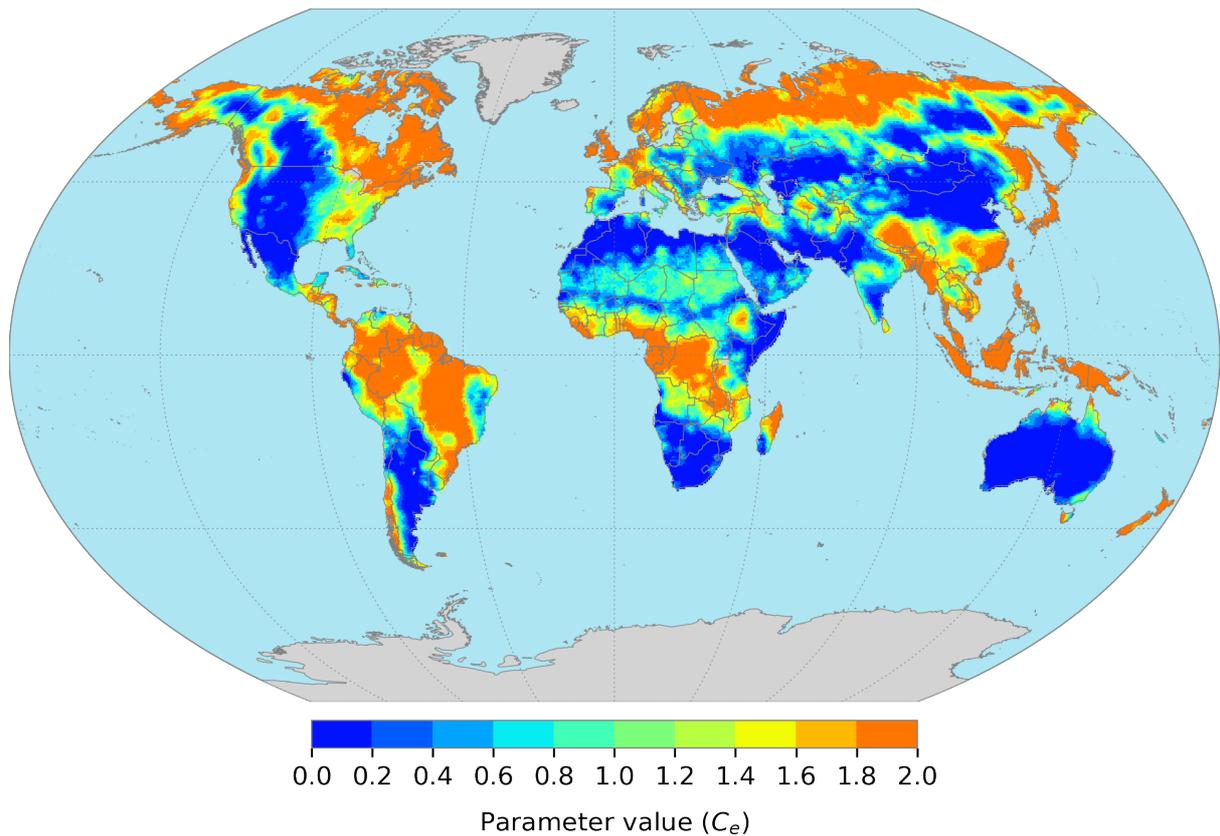
Figure S21. The spatial distribution of model parameter C_e.

References:

Döll, P. and Fiedler, K.: Global-scale modeling of groundwater recharge, Hydrol. Earth Syst. Sci., 12(3), 863–885, doi:10.5194/hess-12-863-2008, 2008.

De Graaf, I. E. M., Sutanudjaja, E. H., Van Beek, L. P. H. and Bierkens, M. F. P.: A high-resolution global-scale groundwater model, Hydrol. Earth Syst. Sci., 19(2), 823–837, doi:10.5194/hess-19-823-2015, 2015.

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H. and Kanae, S.: Global flood risk under climate change, Nat. Clim. Chang., 3(9), 816–821, 2013.

Hunger, M. and Döll, P.: Value of river discharge data for global-scale hydrological modeling, Hydrol. Earth Syst. Sci., 12(3), 841–861, doi:10.5194/hess-12-841-2008, 2008.

Lorenz, C., Kunstmann, H., Devaraju, B., Tourian, M. J., Sneeuw, N. and Riegger, J.: Large-Scale Runoff from Landmasses: A Global Assessment of the Closure of the Hydrological and Atmospheric Water Balances*, J. Hydrometeorol., 15(6), 2111–2139, doi:10.1175/JHM-D-13-0157.1, 2014.

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, Geosci. Model Dev., 10(5), 1903–1925, doi:10.5194/gmd-10-1903-2017, 2017.

Mateo, C. M., Hanasaki, N., Komori, D., Tanaka, K., Kiguchi, M., Champathong, A., Sukhapunnaphan, T., Yamazaki, D. and Oki, T.: Assessing the impacts of reservoir operation

to floodplain inundation by combining hydrological, reservoir management, and hydrodynamic models, Water Resour. Res., 50(9), 7245–7266, doi:10.1002/2013WR014845, 2014.

Mueller, B., Hirschi, M., Jimenez, C., Ciais, P., Dirmeyer, P. A., Dolman, A. J., Fisher, J. B., Jung, M., Ludwig, F., Maignan, F., Miralles, D. G., McCabe, M. F., Reichstein, M., Sheffield, J., Wang, K., Wood, E. F., Zhang, Y. and Seneviratne, S. I.: Benchmark products for land evapotranspiration: LandFlux-EVAL multi-data set synthesis, Hydrol. Earth Syst. Sci., 17(10), 3707–3720, doi:10.5194/hess-17-3707-2013, 2013.

Müller Schmied, H., Eisner, S., Franz, D., Wattenbach, M., Portmann, F. T., Flörke, M. and Döll, P.: Sensitivity of simulated global-scale freshwater fluxes and storages to input data, hydrological model structure, human water use and calibration, Hydrol. Earth Syst. Sci., 18(9), 3511–3538, doi:10.5194/hess-18-3511-2014, 2014.

Pappenberger, F., Dutra, E., Wetterhall, F. and Cloke, H. L.: Deriving global flood hazard maps of fluvial floods through a physical model cascade, Hydrol. Earth Syst. Sci., 16(11), 4143–4156, doi:10.5194/hess-16-4143-2012, 2012.

Wang-Erlandsson, L., Bastiaanssen, W. G. M., Gao, H., Jägermeyr, J., Senay, G. B., van Dijk, A. I. J. M., Guerschman, J. P., Keys, P. W., Gordon, L. J. and Savenije, H. H. G.: Global root zone storage capacity from satellite-based evaporation, Hydrol. Earth Syst. Sci., 20(4), 1459–1481, 2016.

Wartenburger, R., Seneviratne, S. I., Hirschi, M., Chang, J., Ciais, P., Deryng, D., Elliott, J., Folberth, C., Gosling, S. N., Gudmundsson, L., Henrot, A.-J., Hickler, T., Ito, A., Khabarov, N., Kim, H., Leng, G., Liu, J., Liu, X., Masaki, Y., Morfopoulos, C., Müller, C., Schmied, H. M., Nishina, K., Orth, R., Pokhrel, Y., Pugh, T. A. M., Satoh, Y., Schaphoff, S., Schmid, E., Sheffield, J., Stacke, T., Steinkamp, J., Tang, Q., Thiery, W., Wada, Y., Wang, X., Weedon, G. P., Yang, H. and Zhou, T.: Evapotranspiration simulations in ISIMIP2a—Evaluation of spatio-temporal characteristics with a comprehensive ensemble of independent datasets, Environ. Res. Lett., 13(7), 075001, doi:10.1088/1748-9326/aac4bb, 2018.

Yamazaki, D., Kanae, S., Kim, H. and Oki, T.: A physically based description of floodplain inundation dynamics in a global river routing model, Water Resour. Res., 47(4), 1–21, doi:10.1029/2010WR009726, 2011.