

We would like to thank William Chris for his interest in this topic and for the comments to improve our manuscript. Based on the comments some calculations have been performed. Our point-by-point response to the comments is given in the following (**Comments in black, Answers in blue** and **the content related to the changes in the revised manuscript are marked in orange.**):

**This study tries to develop a 'global' hydrological model. The authors are lack of good understanding on hydrological processes, and the methodology they used are not appropriate at all. The authors overclaimed their contribution. The manuscript is poorly written. Some of the figures are not clear. The current manuscript cannot be accepted, and should be returned to the authors to make it a better work.**

Thank you for the comment. In this work, we have extended a widely used lumped model, FLEX, into a distributed model that can be used on a global scale. In addition, a climate-derived root zone storage capacity (RZSC) is integrated into the developed model WAYS to capture the spatial heterogeneity of the rooting systems. We demonstrate the benefit of a climate-derived RZSC to the hydrological model for simulation, especially the capacity of root zone water storage (RZWS) simulation. Thus, we believe the methodology we have used is appropriate and that the hydrological processes conceptualized in WAYS are proper. Based on the comments from the three referees as well as from the short comments, we have further improved the manuscript text as well as the figures and tables.

**My comments are as below: 1. The methodology used are not appropriate at all. The authors compared their runoff simulation against some global model simulation and composite runoff data. We all know the global runoff simulation/composite runoff data are designed for global studies, and can have very large uncertainty on each river basin. They cannot use these data to verify their simulation, especially for a study aiming to develop a 'new model'. Thus, the comparison between the authors' simulation and other runoff data that the authors used means nothing: the authors cannot claim their model is good. The authors should compare their simulation against hydrological gauge observation which is not difficult to collect at all. I doubt the authors' results. They may choose to avoid the comparison against hydrological gauge observation purposely because their model is suffering fatal flaws. For a paper developing a hydrological model, comparison against in situ gauge observation is extremely important. The authors should not skip this step. In addition, the authors compare their runoff simulation after calibrating their model, whereas the other models in ISIMIP are not calibrated. Thus, the comparisons are useless because the models in ISIMIP have large uncertainty which have already been unrevealed in several recent studies by the ISIMIP group (perhaps the authors missed these very important publications).**

Thank you. In fact, the simulated runoff of WAYS is first compared to the reference data ERA-Interim/Land runoff. The performance of WAYS in runoff simulation is evaluated based on the comparison between ERA-Interim/Land data and WAYS simulation.

Since WAYS uses the same driving data as the ISIMIP2a models and the ISIMIP2a simulations are widely discussed in many studies, an additional comparison between WAYS and the

ISIMIP2a models can provide added-value for evaluating our model. Therefore, the ISIMIP2a simulations are also shown in the results section together with the ERA-Interim/Land data. We do mention the purpose of inclusion of the ISIMIP2a simulations in the results (page 13, line 15: “Since WAYS uses the same driving data as the ISIMIP2a models and the ISIMIP2a simulations have been widely discussed in many studies (Schewe et al., 2014; Müller Schmied et al., 2016; Gernaat et al., 2017; Zaherpour et al., 2018), we also perform a comparison between WAYS and the ISIMIP2a models to further evaluate our model.”). However, this is not mentioned in the validation strategy section of the manuscript. We have now further clarified this issue in the revised manuscript (please see “Authors’ change in the manuscript.”).

Indeed, the ISIMIP2a models are not calibrated. We have mentioned this issue in the manuscript (page 13, line 30: This result occurs partly because some of the ISIMIP2a models are not calibrated at all (Zaherpour et al., 2018), whereas WAYS is calibrated to a Composite Monthly Runoff data set that assimilates the monitored river discharge (Fekete et al., 2011).). We have revised the captions of the related figures (Figures 2 and 3 in the revised manuscript) to note this issue.

To enhance the validation part of this manuscript, we have additionally evaluated our results with observed discharge from the Global Runoff Data Centre (GRDC). Since WAYS does not currently have a native runoff routing module, a third-part runoff routing tool, CaMa-flood, is applied to route the WAYS simulated runoff (Yamazaki et al. 2011). Given that the manuscript is already quite extensive, the discharge comparison is not a direct evaluation of WAYS but of both WAYS and CaMa-Flood. This information has been added in Supplementary Information (SI). Since this comment shares the similar opinion with Referee #3 (comment 2), we would like to refer to the responses to the comments of referee #3 to avoid repetition, as the response is long. The corresponding revision to the manuscript can also be found there.

### Authors’ change in the manuscript.

#### Page 11, Line 5: (the changes is marked as blue)

In this study, the ERA-Interim/Land runoff data are used for validation of the runoff simulation, and the Normalized Difference Infrared Index (NDII) is used for the validation of the WAYS model for root zone water storage simulation. Considering the time period of coverage of both data sets (ERA-Interim/Land: 1979-2010, NDII: 2000-present) and the study period (1971-2010) of this work, the period 2001-2010 is selected as the validation period. For runoff evaluation, ISIMIP2a simulations are also included, as they use the same climate forcing as our study in the same period. The purpose of inclusion of the ISIMIP2a simulations for comparison can be found in the model evaluation section (see Section 4).

**2. The authors are lack of basic knowledge about remote sensing. The root zone can be more than 10 meters in depth. The sensors used in the NDII studies cannot penetrate the earth ground up to 10 meters, and even one or two meters are suffering large uncertainty because the attenuations of signals with increase in depth. This is why the most state-of-**

**the-art soil moisture products just provide data in the surface 5/10 cm. Thus, the comparison against NDII based data is not appropriate at all. Because this paper is to develop a model, the authors should use in situ observation which is not difficult to collect. I don't understand why the authors choose to skip the comparison against gauge observation.**

Thank you for the comment. As we stated in the manuscript (page 2, line 8 in the revised manuscript), remote sensing itself can only detect the soil water in the surface layer. However, NDII is not a direct observation from the satellite but a Normalized Difference Index, similar to NDWI, NDVI, and so on. It is calculated based on infrared reflectance (NIR) and shortwave infrared reflectance, and it reflects the water stress in the root zone layer and, thus, can be used as proxy data for RZWS rather than RZWS itself. The NDII-related information is interpreted in detail in Section 3.3.2 in the manuscript.

We do not compare the situ observations because there are no observations available for RZWS (Sriwongsitanon et al., 2016). The observation you mentioned is probably the soil moisture at a certain depth, which differs from RZWS.

**3. The authors do not have a good understanding on hydrological processes. i). Vegetation plays a vital role in runoff variations especially in densely vegetated regions (e.g., the Amazon, Congon and some regions in the Yangtze, Mekong, Ganges, Mississippi Rivers et.al.) through the transpiration processes. At the leaf and canopy scales, the mechanisms of transpiration are also different. LAI, fPAR, CO<sub>2</sub>, wind, solar radiation, stomatal conductance all are influencing transpiration. The authors did not consider the stomatal influence at all (as shown in the Figure 1 and Table 1). Without comprehensively considering the transpiration processes, how the model developed can predict water resources availability, especially many recent studies have unravelled that the earth is greening and CO<sub>2</sub> concentration is increasing. Thus, the model developed by the authors has fatal flaws, and this paper cannot be accepted.**

Thank you for the comment. We agree with the reviewer that vegetation plays a vital role in runoff variations, especially in densely vegetated regions, and the mechanisms of transpiration are also different at the leaf and canopy scales. However, the model we developed in this study is a conceptual hydrological model with a conceptualized structure to mimic the hydrological cycle. This design differs from land surface models, dynamic vegetation models or physically based hydrological models, which could have more functions with physical meanings (Bierkens, 2015). The conceptual hydrological model, however, has its own advantages in practicability and computation efficiency (Devia et al., 2015). The transpiration is of course considered by conceptual models, while some of them calculated the total evaporation without separating the evaporation into different fluxes. Moreover, conceptual models are widely applied for water related applications, e.g., runoff simulation and water scarcity analysis, especially on a global scale (Döll et al., 2003; Döll and Fiedler, 2008; Hanasaki et al., 2008; Wang-Erlandsson et al., 2014). Thus, a well-developed conceptual model, such as WAYS, should be proper for predicting water resource availability.

In addition, the continuous greening of the earth as well as the increased CO<sub>2</sub> concentration are indeed important issues. However, they are beyond the scope of this study as they are more related to climate change analysis.

**ii). The infiltration capacity of soil plays an important role in controlling the volume of surface runoff and subsurface runoff, and also influences root zone water storage. The infiltration capacity of soil is related to soil type, and has clear physical meaning. The authors considered the infiltration as shown in the Figure 1. However, the authors did not report how they determine this important parameter value. If the authors used the values related each soil type, they did not report which soil map distribution data and which hydraulic property datasets of the soil types are used. If the authors calibrated the parameter values, the authors should be aware of that if it is appropriate to calibrate because the results may be wrong after calibrating some parameters with clear physical meaning. The authors are afraid of reporting the calibrated parameter values and the parameter ranges used in the calibration. The authors stated they calibrated their model for good runoff simulation. I am afraid that they calibrated their model for good runoff simulation with the cost of losing the physical meaning of important parameters. Perhaps the authors choose to not show the important information purposely in order to get their paper published. No, absolutely no. The authors have to show which parameters are calibrated, the parameter ranges used in calibration and calibrated parameter values.**

Thank you for the comment. The precipitation partitioning function in the WAYS model is based on a widely used beta function of the Xinanjiang model (Zhao, 1992). It is a conceptualized runoff generation function that consists of empirical parameters. The model is a conceptual model and is different from the physically based model. The model parameter must be calibrated before simulation. Indeed, the physically based model is usually run without calibration as the parameters it uses have corresponding physical meanings. However, the physical model and conceptual model are two different methods without any conflicts between each other.

Moreover, we have added a table to illustrate the parameters used as well as the parameter ranges if calibration is needed. Please refer to the changes in the revised manuscript (page 11, Table 2). We will share the calibrated parameters together with the code for the model after the paper is accepted for publication.

Since the calibrated parameters are spatially distributed and are not appropriate to show in tables, we provide the spatial patterns of two key parameters ( $\beta$ ,  $C_e$ ) that are calibrated, as these two parameters mostly affect the partitioning of precipitation (see Figure S21 and Figure S22). The rest of the calibrated parameters are uploaded to the response thread in a netCDF file as a supplementary document.

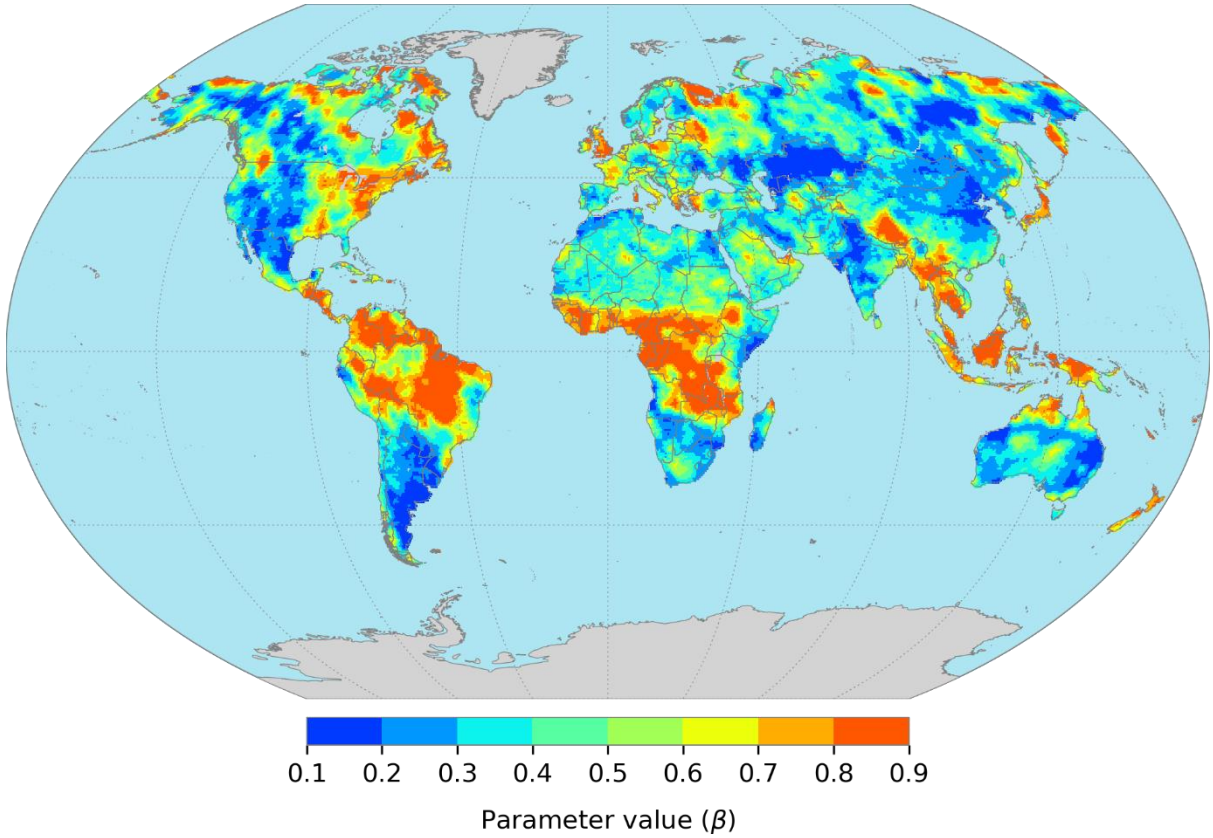


Figure S21. The spatial distribution of the model parameter  $\beta$

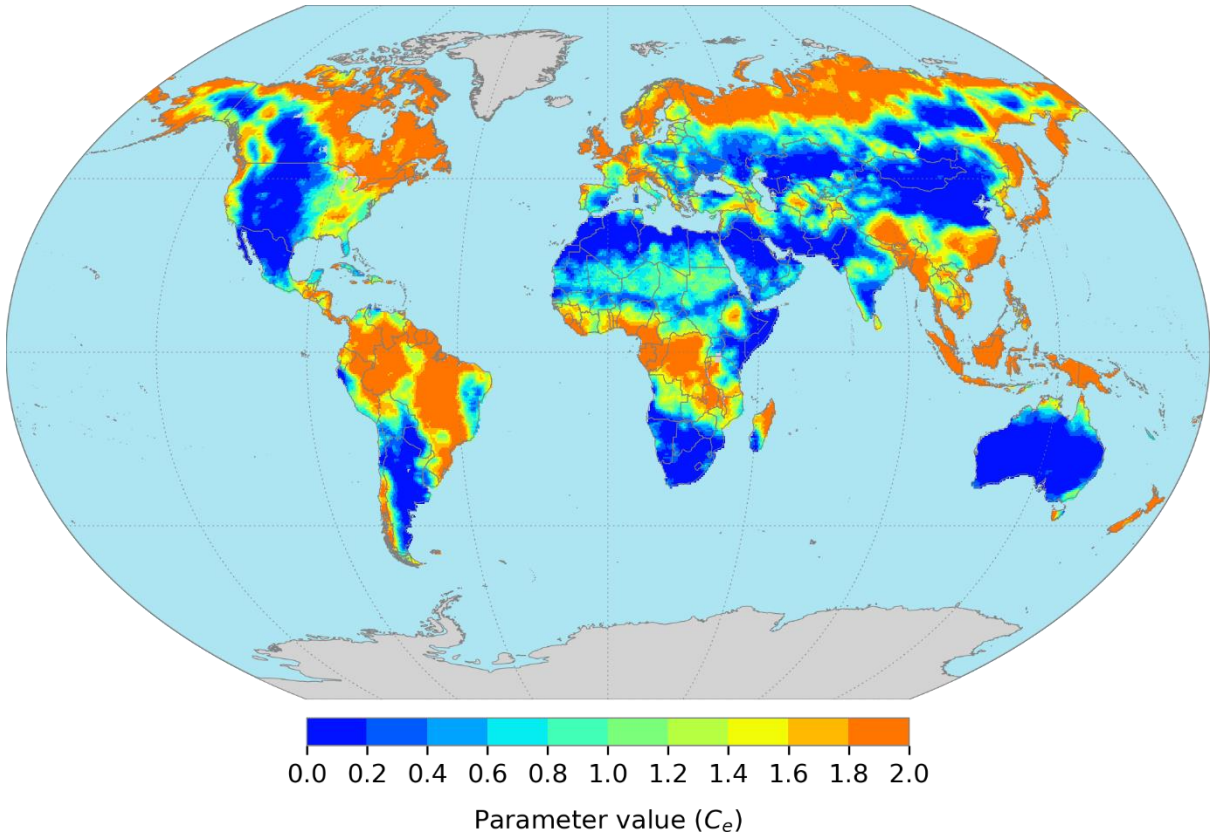


Figure S22. The spatial distribution of the model parameter  $C_e$

Authors' change in the manuscript.

Page 12: (the following table is added)

Table 2. Parameter ranges of the WAYS model

Parameter	Range	Literature	Parameter	Range
$S_{i,max}$	distributed	Wang-Erlandsson et al. (2014)	$\beta$	(0, 2)
$S_{rz,max}$	distributed	Wang-Erlandsson et al. (2016)	$C_e$	(0.1, 0.9)
$R_{s,max}$	7/4.5/2/5 (Sand/Loam/Clay)	Döll and Fiedler (2008)	$K_f$	(1, 40)
$K_s$	100	Döll et al. (2003)	$K_{ff}$	(1, 9)
$f_s$	distributed	Döll and Fiedler (2008)	$S_{ftr}$	(10, 200)
$F_{DD}$	distributed	Müller Schmied et al. (2014)	$T_{lag}$	(0, 5)
$T_t$	0	Müller Schmied et al. (2014)		

**4. The model developed is not a global scale model at all. Because the authors did not use soil map and related soil hydraulic parameter values, the use of the model must rely on calibration to determine some of its parameter values on river basin scales. Therefore, it cannot be a global scale model. It is still a river basin scale model, and the authors just applied the model in several large-scale river basins (without any river basins in most of the regions of Canada, Europe, Middle East, Russia, Mongolia). The used river basins just cover a small proportion of global land surface.**

Thank you for the comment. The WAYS model actually uses many global parameters for hydrological simulation, e.g., RZSC, land cover, DEM, digital maps of the slope, soil texture, geology and permafrost information (see page 6, line 5 in the revised manuscript).

Based on the comments from the referees as well as from the short comments, additional evaluations covering large areas are included in the revised manuscript. These include discharge comparison to GRDC observation, evaporation comparison to FLUXNET2015 and LandFluxEVAL data. The results of the discharge evaluation can be found in responses to comments of referee #3 (comment 2). For the evaporation evaluation, we would like to refer to the responses to the comments of referee #3 (comment 5) to avoid repetition, as the response is long. The corresponding revision to the manuscript can also be found there.

**5. The authors claimed they used 2000 iterations to calibrate their model. However, the authors did not explain the reason. Why 2000 iterations were used?**

Thank you. In fact, the number of iterations is recommended by the author of Dynamically Dimensioned Search (DDS) algorithm (Tolson and Shoemaker, 2009). We have clarified this in the revised manuscript.

Authors' change in the manuscript.

Page 13, Line 6: (the changes is marked as blue)

The criterion of fit for calibration is the Nash-Sutcliffe efficiency coefficient (NSE), and the DDS optimization algorithm is run with 2000 iterations for each grid cell for parameter estimation, as suggested by the author of DDS (Tolson and Shoemaker, 2007).

**6. The root zone storage variations are related to ground water level dynamics. Did the model simulate the ground water level changes? Please show the simulation results.**

Thank you. The current work did not consider the groundwater level changes as well as the capillary rise due to the lack of groundwater table information. We have discussed this issue in the revise manuscript.

**Authors' change in the manuscript.**

**Page 25, Line 4: The following paragraph is inserted in the discussion part**

Moreover, the current study does not consider the groundwater access and irrigation mainly due to the lack of global information. The groundwater table information is crucial for capillary rise simulation (Vergnes et al., 2014). Capillary rise simulation without proper water table information could significantly overestimate the evaporation. Thus, the capillary rise flux is ignored in this study. A similar strategy has also been applied by other works due to the absence of the information on the global water table (Döll et al., 2003; De Graaf et al., 2015; Hanasaki et al., 2018). Observations of irrigation on the global scale are also not available (Leng et al., 2015). Although there are simulated irrigation data available on the global scale, the inherent uncertainties could be propagated in our model simulation. Therefore, irrigation is also not considered at this time. However, this neglect could potentially introduce biases into the model simulation in irrigated areas and deep rooted plant-distributed regions, as both irrigation and capillary rise are an additional supply of soil water recharge. The biases may cause an underestimation of evaporation, especially in the dry summertime (Vergnes et al., 2014). This underestimation could consequently affect the simulation of RZWS and runoff because of the interlinkage of these three elements (Rockström et al., 1999). It is found that ignoring the capillary rise could reduce soil water content in the root zone (RZWS), while the runoff will also be reduced (Vergnes et al., 2014). However, these shortcomings can be simply overcome once the global data are available.

**7. Please use scientific languages. The sub-titles of Section 2.4 and 2.5 are not appropriate in such as a scientific paper. The statements 'Fast- and Slow-' are vague.**

Thank you. We have now changed the Fast- and Slow- flow to preferential flow and matrix flow based on the related literature (Ali et al., 2018; Gao et al., 2019).

**Authors' change in the manuscript.**

The fast flow and slow flow are replaced by preferential flow and matrix flow in the entire manuscript.

**8. I agree with the reviewer 1 about the capillary mechanism which is missed by the authors. This indicates the authors are lack of good understanding on hydrological processes from another perspective. When we develop a new model, we try to incorporate new hydrological mechanism to advance our understanding on hydrological processes. However, the authors missed several very important hydrological processes which have already been recognised to be very important. Therefore, the 'developed' model cannot provide any new understanding on hydrology to us. I am afraid that the authors just copy other models' code, delete several important parts, replace a few equations and change computer language used in original code, and then the authors claim they develop a new model. No, this is not the right way to do research. I also wonder why the authors delete the capillary mechanism part from the original code. The authors should realize that they cannot just delete some codes of other's model, and make it look like a 'new model' in order to get the manuscript published. This is not real science. The authors must work hard to consider the capillary and vegetation transpiration mechanisms and using gauge data to validate their simulation. Otherwise, their model cannot be better (based on the physical processes considered) than other hundreds/thousands of models that already exist.**

Thank you. In fact, WAYS does include the capillary module from Gao et al. (2014a), a key publication on the FLEX model. At the current stage, it is, however, disabled due to the lack of global information on the groundwater table. A detailed explanation could be found in the responses to comment 6 of referee #1. The corresponding revision in the manuscript can also be found there.

**9. The manuscript is poorly written and needs to be largely reworked. There are many typos and grammar mistakes. Many sentences are vague and lack of support. The figures are not clear, e.g., Figure 4 and Figure 5, and one cannot distinguish the lines.**

Thank you. We have carefully checked the manuscript and corrected typos and grammatical mistakes. The revised manuscript has been edited by a professional academic language and manuscript service company. We have also further improved the manuscript as well as the figures and tables. Figure 4 and Figure 5 are reproduced with high resolution, and the lines are clearer now. Since short comment 10 suggested us to move some figures from the main text of the manuscript. Thus, the figures in the revised manuscript are re-sorted.

**10. Figure 2 is not your result. Please remove Figure 2. Using related references in the manuscript to refer to the data is ok.**

Thank you. Indeed, Figure 2 shows the spatial distribution of RZSC, which is obtained from Wang-Erlandsson et al. (2016). Since it a key parameter for the model we developed and spatial distribution information would be useful, we have move them to SI rather just placed them in the references. In addition, Figure 3, which shows the latitudinal averaged RZSC, has also been moved from the main text of the manuscript to SI, as it is also based on the results from Wang-Erlandsson et al. (2016).



It is also important to note that the RZSC is now updated based on the comment of Referee #3. Referee #3 suggested that RZSC should be updated by applying the Gumbel normalization, as Wang-Erlandsson et al. (2016) found that normalizing the RZSC using the Gumbel distribution by land cover type further improves performance.

**Authors' change in the manuscript.**

The Figures 2 and 3 are moved to the SI.

**All the references are included in the manuscript.**