HAL
open science

# Evaluation of Groundwater Simulations in Benin from the ALMIP2 Project

Mehnaz Rashid, Rong-You Chien, Agnès Ducharne, Hyungjun J Kim, Pat J.-F. Yeh, Christophe Peugeot, Aaron Boone, Xiaogang He, Luc Seguis, Yutaro Yabu, et al.

## To cite this version:

Mehnaz Rashid, Rong-You Chien, Agnès Ducharne, Hyungjun J Kim, Pat J.-F. Yeh, et al.. Evaluation of Groundwater Simulations in Benin from the ALMIP2 Project. Journal of Hydrometeorology, 2019, 20 (2), pp.339-354. 10.1175/JHM-D-18-0025.1 . hal-02165400

**HAL Id: hal-02165400**

**https://hal.sorbonne-universite.fr/hal-02165400**

Submitted on 5 Jul 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Evaluation of groundwater simulations in Benin from the ALMIP2 project

Mehnaz Rashid, Rong-You Chien, Agnès Ducharne, Hyungjun Kim, Pat J.-F. Yeh, Christophe Peugeot, Aaron Boone, Xiaogang He, Luc Séguis, Yutaro Yabu, Moussa Boukari, Min-Hui Lo*

Mehnaz Rashid: Department of Atmospheric Sciences, National Taiwan University, Taipei, Taiwan (geo.mehnaz@gmail.com)

Rong-You Chien: Department of Atmospheric Sciences, National Taiwan University, Taipei, Taiwan (pshchien@gmail.com)

Agnès Ducharne: Sorbonne Université, CNRS, EPHE, UMR 7619 METIS, 4 place Jussieu, 75005 Paris, France (Agnes.Ducharne@upmc.fr)

Hyungjun Kim: Institute of Industrial Science, The University of Tokyo, Tokyo, Japan (hjkim@iis.u-tokyo.ac.jp)

Pat J.-F. Yeh: Department of Civil and Environmental Engineering, National University of Singapore, Singapore (patyeh2007@gmail.com)

Christophe Peugeot: HSM, IRD, Univ. Montpellier, CNRS, Montpellier, France (christophe.peugeot@ird.fr)

Aaron Boone: CNRM Météo-France/CNRS, Toulouse, France (aaron.a.boone@gmail.com)

Xiaogang He: (1) Institute of Industrial Science, The University of Tokyo, Tokyo, Japan (2) Department of Civil and Environmental Engineering, Princeton University, Princeton, NJ, USA (hexg@princeton.edu)

Luc Séguis: HSM, IRD, Univ. Montpellier, CNRS, Montpellier, France (luc.seguis@univ-montp2.fr)

Yutaro Yabu: Institute of Industrial Science, The University of Tokyo, Tokyo, Japan (yabu@rainbow.iis.u-tokyo.ac.jp)

Moussa Boukari: Laboratoire d'Hydrologie Appliquée Université d'Abomey-Calavi 01BP 4521 Cotonou – Bénin (moussaboukari2003@yahoo.fr)

Min-Hui Lo: Department of Atmospheric Sciences, National Taiwan University, Taipei, Taiwan (minhuilo@ntu.edu.tw)

*Corresponding Author: Dr. Min-Hui Lo

**Abstract**

A comprehensive estimation of water budget components, particularly groundwater storage (GWS) and fluxes, is crucial. In this study, we evaluate terrestrial water budget of the Donga basin (Benin, West Africa), as simulated by three land surface models (LSMs) used in the African Monsoon Multidisciplinary Analysis, Land Surface Model Intercomparison Project Phase 2 (ALMIP2): CLM4, CLSM, and MATSIRO. All three models include an unconfined groundwater component and are driven by the same ALMIP2 atmospheric forcing from 2005 to 2008. Results show that all three models simulate substantially shallower water table depth (WTD) with smaller seasonal variations approximately 1-1.5 m compared to the observed values that range between 4-9.6 m. While as the seasonal variations of GWS are overestimated by all the models. These seemingly contradictory simulation results can be explained by the overly high specific yield prescribed in all models. All models achieve similar GWS simulations but with different fractions of precipitation partitioning into surface runoff, baseflow and evapotranspiration (ET), suggesting high uncertainty and errors in the terrestrial and groundwater budgets among models. The poor performances of models can be attributed to bias in the hydrological partitioning (baseflow vs. surface runoff) and sparse sub-surface data. This analysis confirms the importance of subsurface hydrological processes in the current generation of LSMs and calls for substantial improvement in both surface water budget (which controls groundwater recharge) and the groundwater system (hydrodynamic parameters, vertical geometry).

## 1. Introduction

Approximately one-third of the global population currently lives in the water-stressed regions, and this figure is expected to rise in the future, because of climate change, expanding irrigation, and growing population (Alcamo et al. 2007; Döll 2009; Vörösmarty et al. 2000; Elliott et al. 2014; Taylor 2014; Jasechko and Taylor 2015). This will increase dependence on groundwater, a relatively reliable source of freshwater in water-deficient regions, with a recognized potential to enhance economic development and ensure food security (Wada and Bierkens 2014). Yet, the accompanying exploitation may threaten the future sustainability of groundwater resources (Gleeson et al. 2010; Gurdak 2017), and the declining groundwater levels can, in turn, have severe effects on the hydrological budget, irrigated agriculture, and ecosystems (Wada et al. 2010) that could even alter the regional hydro-climatology (Lo et al. 2013; Wey et al. 2015; Zeng et al. 2017). This is particularly true for Africa, which is the home to approximately one billion people and has the highest population growth rate in the world, with the population projected to reach 1.5-billion by 2050 (UN 2015). In addition, this region lacks proper quantification of freshwater resources (Taylor 2014); therefore, a comprehensive assessment of water resources particularly groundwater is essential for sustainable future water use.

The West African Monsoon (WAM) is the main water resource for West Africa and strongly influences water and food security of this region where approximately 300-million people live and depend directly on the rain-fed agriculture (Ndehedehe et al. 2016). Agricultural land use in this region has increased by 57% from 1975 to 2000 (Eva et al. 2006), but severe droughts were experienced during the recent decades (Le Barbé et al. 2002; Mahé and Paturel 2009; Lebel et al. 2009; Tadesse et al. 2008), and climate projections do not provide a clear picture of future rainfall trends over the region. However, enhanced runoff capabilities and floods of Sahelian watersheds,

due to land clearing and past drought impacts, may persist or increase due to rainfall intensification (Panthou et al. 2014; Sylla et al. 2016). In this context, groundwater is a vital source of water in West Africa, primarily for improving food security (Pavelic et al. 2012; Altchenko and Villholth 2015). Therefore, climate models with a detailed representation of the groundwater dynamics are required to address the complex interactions between the climate system and the water-food-energy nexus (Ringler et al. 2013; Rasul and Sharma 2015). This involves land surface models (LSMs) with reliable simulations of GWS variations and subsurface hydrological processes. However, the assessment of water and energy fluxes simulated by LSMs are generally challenging due to the scarcity of observed data, particularly for the subsurface hydrological processes (Lo et al. 2010; Cai et al. 2014).

The African Monsoon Multidisciplinary Analysis (AMMA) project (Redelsperger et al. 2006) was aimed to understand the daily-to-interannual variability of the WAM dynamics and terrestrial water budget. In this framework, the AMMA Land Surface Model Intercomparison Project (ALMIP) has been designed to gain a better understanding of the skills of state-of-the-art LSMs using data-scarce West Africa as a test bed. The Phase 1 of ALMIP (ALMIP1) was focused on a large domain in West Africa using 3-hourly meteorological forcing data at the 0.5° spatial resolution (Boone et al. 2009a; Xue et al. 2012). A major achievement was the use of an ensemble of offline LSM simulations with the same climate forcing data to produce a multi-model analysis of land surface fluxes from 2004 to 2007. The use of three different forcing datasets showed that the precipitation forcing is crucial for properly simulating the water and energy budgets in the region. The ALMIP1 LSM simulations were compared with satellite estimates of total water storage variations (Grippa et al. 2011). The compared LSMs fairly captured the seasonal variability of the total water storage, despite an underestimation of the seasonal amplitude by the multi-model

mean. Interestingly, the only model with an explicit groundwater description, namely the Catchment LSM (Koster et al. 2000), showed a larger amplitude, in better agreement with satellite gravimetry data.

Due to the coarse spatiotemporal resolution of the models the comparison with in situ data from the heavily instrumented supersites from the AMMA-CATCH observing system in Mali, Niger, and Benin (www.amma-catch.org, Lebel et al. 2009) was difficult. This was a major motivation for the Phase 2 of ALMIP (Boone et al. 2009b), which is focused on three meso-scale domains including the above supersites, with high-quality half-hourly meteorological forcing data at 0.05° resolution (approximately 5 km). In addition, the three participating LSMs namely the Community Land Model version 4 (CLM4) (Oleson et al. 2010; Lawrence et al. 2011), the Catchment LSM (CLSM) (Ducharne et al. 2000; Koster et al. 2000), and the Minimal Advanced Treatments of Surface Interaction and Runoff (MATSIRO) (Takata et al. 2003) include an unconfined aquifer module for simulating water table dynamics. The current study is thus focused on evaluating the groundwater simulations in these ALMIP2 models, in one of the ALMIP2 representative sites, the Donga River basin in northern Benin, where measurements of WTD and streamflow are available from 2005 to 2008 (Kamagaté et al. 2007; Séguis et al. 2011; Hector et al. 2015). Furthermore, the Donga basin does not exhibit significant anthropogenic groundwater withdrawal (GW withdrawal is less than 1% of the total water consumption, Vouillamoz et al. 2015), that would otherwise influence the results due to the lack of such representation in the evaluated models. Comparisons of simulated GWS and the related water fluxes with corresponding in situ and satellite observations are carried out with two complementary objectives: (1) understanding the strengths and shortcomings of the groundwater simulations by these LSMs; (2) providing some insight into the water budget of the studied basin.

## 2. Study area

This study focuses on the Donga River basin in Benin (Figure 1), one of the AMMA-CATCH representative sites (www.amma-catch.org). The monitored river basin has an area of 586 $km^2$, and the topography ranges between 317 m and 508 m above sea level with a subdendritic drainage pattern. The climate is subhumid, with a mean annual precipitation of 1200 mm (Séguis et al. 2011), concentrated in a single rainy season extending from March to October, and approximately 60% of precipitation occurring between July and September (Depraetere et al. 2009). Streamflow starts in June–July, 2 to 3 months after the beginning of the rainy season, and declines by the end of the rainy season (October – November). The groundwater system is composed of a permanent unconfined aquifer located in a weathered crystalline formation (saprolites), with a thickness of 10 to 25 m over the metamorphic bedrock, overlaid with a seasonal perched groundwater, which contributes most of the baseflow (Kamagaté et al. 2007; Séguis et al. 2011; Hector et al. 2015). Inhabitation is mainly rural, and domestic water demand is supported by groundwater from boreholes. Land use and land cover are composed of wooded savannah along with agricultural land with sorghum, maize, yam, kassava, cotton, and cashew nuts as major crops.

## 3. Models

### *a. CLM4*

CLM4 (Oleson et al. 2010; Lawrence et al. 2011) coupled with an unconfined groundwater model (Niu et al. 2007) includes representations of various hydrological processes, namely soil hydrology, runoff generation, groundwater dynamics, and snowpack. Runoff generation is calculated as an exponential function of the grid-mean WTD, using simple TOPMODEL based runoff parametrization (SIMTOP) (Niu et al. 2005). The 3.8-m soil depth is divided into 10 hydrologically

active layers with varying soil layer thicknesses. The water table and soil moisture are coupled through net groundwater recharge (gravity drainage − capillary rise); moreover, the representation of groundwater dynamics is a simple extension of a one-dimensional diffusion equation, with a full hydraulic connection between the saturated and unsaturated zone, but no horizontal groundwater fluxes between neighboring grid cells. Soil water is calculated using a revised numerical solution of the one-dimensional Richards equation (Zeng and Decker 2009). Groundwater recharge from the soil column and WTD are updated using the parameterizations given in Niu et al. (2007) and Zeng and Decker (2009). Sub-surface runoff (baseflow) is formulated as:

$$Baseflow_{CLM4} = qdrai_{max} \times exp(-f * WTD) \quad (1)$$

$qdrai_{max}$ is 0.0055 mm $s^{-1}$ (global constant derived from sensitivity tests) and $f$ is the decay factor in the SIMTOP scheme with the value of 2.5 $m^{-1}$ (Niu et al. 2005). The WTD is in units of meters. In addition, the remained excess water after saturating the entire soil column is treated as the saturation excess baseflow, and is added into the baseflow generated by the SIMTOP scheme (Oleson et al. 2010). The unconfined aquifer is assumed to have a storage capacity of 5000 mm, with a 25-meter depth to bedrock (Oleson et al. 2010).

### *b. CLSM*

CLSM (Ducharne et al. 2000; Koster et al. 2000) belongs to a generation of LSMs that rely on the hydrological model TOPMODEL (Beven and Kirkby 1979) to account for lateral water fluxes along the topography, as well as their influence on the small-scale variability of soil moisture, runoff, and ET and thus, on the larger scale water budget. The simulated domain is discretized into elementary units, each including a shallow WTD. They are designed to be watersheds, but they

can be grid cells, as in this study with the 0.05° ALMIP2 mesh. In each grid cell, the spatial distribution of the topographic index is used as a template to laterally distribute the WTD around its mean value, which varies in time with the water budget. According to the concept of TOPMODEL, the WTD distribution controls the saturation excess runoff and baseflow to the streams. This distribution also influences the water exchanges with the root zone (recharge and capillary rise, both defining the net recharge). The resulting soil wetness in the root zone and surface layer (2 cm) drives the infiltration-excess (Hortonian) runoff and ET described using classic soil–vegetation–atmosphere transfer formulations, mostly derived from the Mosaic LSM (Koster and Suarez 1992). For computational efficiency, all processes related to the spatial WTD distribution are approximated by empirical functions of mean WTD.

In this framework, it is noteworthy that GWS and specific yield are not defined in a standard way in CLSM. As detailed in Ducharne et al. (2000), GWS is quantified by the mean saturation deficit assuming a hydrostatic soil moisture profile above the spatially distributed WTD, and this mean saturation deficit ($M_D$ in mm) is linked to the mean WTD ($d^*$ in m) by a quadratic function: $M_D = A\,(d^*+B)^2$. The parameters A and B are adjusted in each grid-cell to match the results of high-resolution 3D hydrodynamic calculations accounting for the full WTD distribution, and they implicitly depend on assumed porosity, suction, hydraulic conductivity. The specific yield is approximated by $M_D/d^*$, with a spatio-temporal range between 0.084 and 0.147 and a mean ca 0.12, further used for comparison with the other models.

The version of CLSM used in ALMIP2 is the same as that of the ALMIP1 intercomparison (Boone et al. 2009a; Grippa et al. 2011). To accurately constrain topography-related parameters, the topographic index (TI) is calculated from the 3-arc-sec resolution (ca 90 m at Equator) hydrologically conditioned elevation data of HydroSHEDS (Lehner et al. 2008); therefore, the

required TI distribution comprises 3600 values in each 0.05° cell (instead of 25 values only if we had kept the 1-km elevation used for ALMIP1). The mean WTD and baseflow, as in TOPMODEL, depend on the mean TI, which also depends on the digital elevation model (DEM) resolution. To prevent from changes in the simulations that would only result from a change in DEM resolution, the choice has always been made in the CLSM to scale the mean TI values to the ones of a 100-m DEM (Ducharne et al., 2000). The scaling law of Wolock and McCabe (2000) used in ALMIP1 only works to rescale the mean TI (noted $TI^*$) from a 1-km to a 100m resolution, and we rather used here the more generic scaling method of Ducharne (2009), according to which the DEM resolution is a scaling factor: $TI^*_{100m} = TI^*_{90m} + \ln(100/90)$. Other differences compared with the ALMIP1 simulations lie in the model's parameters, based on different soil and vegetation datasets.

*c. MATSIRO*

The original version of MATSIRO (Takata et al. 2003) was designed to serve as the land surface scheme of the general circulation model from the Model for Interdisciplinary Research on Climate (Hasumi and Emori 2004; Watanabe et al. 2010). Energy balances at the ground and canopy surfaces are considered in snow-covered and snow-free areas separately within a grid cell according to the subgrid snow distribution. Water evaporation is calculated using the bulk transfer function, and transpiration is estimated using the SiB2 parameterization scheme (Sellers et al. 1986, 1996). The surface and subsurface runoff processes are based on those of a simplified TOPMODEL (Stieglitz et al. 1997; Beven and Kirkby 1979). To incorporate an explicit representation of groundwater dynamics for more realistic hydrological simulations, Koirala et al. (2014) incorporated a groundwater scheme (Yeh and Eltahir 2005) with globally varying parameters into the original MATSIRO to more effectively consider the soil moisture flux exchange between an unsaturated soil layer and the underlying unconfined aquifer near the water

table. Given that the depth of an unsaturated soil zone can vary depending on the WTD, the soil column is extended to 40 m in the updated MATSIRO (Koirala et al. 2014). After the inclusion of groundwater dynamics, model processes have been demonstrated to yield an improved simulation of river discharge (Koirala et al. 2014) and groundwater withdrawals and depletion (Pokhrel et al. 2015). We use MATSIRO with groundwater representation based on (Yeh and Eltahir 2005 a, b; Koirala et al. 2014).

## 4. Data and methods

### *a. Simulation design and forcing data*

The simulations were performed for ALMIP2 intercomparison at a spatial resolution of 0.05° over the Benin mesosite. A common set of forcing data was provided to all ALMIP2 participants; hence the same half-hourly meteorological forcing is used to drive all three LSMs for the period 2005–2008. We used here in-situ rainfall data from the AMMA-CATCH rain gauge network interpolated with the Thiessen method. Downwelling radiative fluxes are obtained from the Surface Analysis Satellite Applications Facility (Land-SAF) in Meteosat Second Generation satellite series (Geiger et al. 2008). The European Centre for Medium-Range Weather Forecasts provides the remaining required meteorological forcing variables, namely air temperature, specific humidity, and surface pressure. The data are resampled from a 3-km resolution to a 0.05° resolution at 30-minute time steps for use in ALMIP2 studies. Other required parameters such as land cover, surface roughness, leaf area index, albedo, and soil properties, are provided by the ECOCLIMAP-II-Africa database (Thibaut et al. 2010). Notice that the vegetation map in CLM4 uses intrinsic parameters based on MODIS. Further details concerning the ALMIP2 protocol can be found in Boone et al. (2009b), He et al. (2015), Getirana et al. (2017), and Grippa et al. (2017). The original model outputs are at

the native ALMIP2 time step (30-minute), and we aggregated it to monthly resolution for the current analysis.

### *b. Validation data*

WTD measurements are provided by the AMMA-CATCH observation system, and the average time series of WTD data from 2005-2008 are constructed from 24 wells distributed throughout the Donga basin (Figure 1) (Séguis et al. 2011). The wells monitor the moderately deep saprolite aquifer, and the average WTD ranges between 4 and 9.6 m below ground level.

Streamflow measurements ($m^3/s$) are available for the period 2005–2008 at 15-minute time steps at the outlet of the river basin (station Donga-Pont, shown in Figure 1). Streamflow is converted from $m^3/s$ to mm/day by time-aggregation and normalization by the watershed area. Monthly ET simulations are compared with two satellite-based datasets, namely the Moderate Resolution Imaging Spectroradiometer (MODIS) product MOD16 and Global Land Evaporation Amsterdam Model (GLEAM). The global monthly ET product MOD16 works on the Penman–Monteith framework and data are available at a spatial resolution of 0.05° (Mu et al. 2011). GLEAM works on the Priestley and Taylor methodology using multiple satellite sensor products for generating daily ET data at a spatial resolution of 0.25° (Miralles et al. 2011; http://www.gleam.eu/#downloads).

## 5. Results

### *a. Total water budget*

Table 1 shows that the three LSMs simulate noticeably different water budget partitions. For ET, the simulations of MATSIRO (79% of multi-year mean precipitation *P*) and CLM4 (72% of *P*)

are much larger than that of the CLSM (57% of *P*), and they are closer to the ET estimates obtained from the two satellite datasets (66% to 71% of *P*). For streamflow (Q), the simulations of the CLSM (42% of *P*) and CLM4 (28% of *P*) are larger than that of the MATSIRO (18% of *P*), and they are all larger to much larger than the corresponding streamflow observations (12% of *P*). Regarding the total runoff partitioning between surface runoff and baseflow, the three models also present contrasted results. The MATSIRO and CLM4 simulate nearly two-thirds of total runoff as baseflow, whereas the CLSM simulates all runoff as surface runoff. The simulated groundwater recharge (taken here as the net recharge, i.e. the downward recharge reduced by upward capillary fluxes to the unsaturated zone) is a small fraction of precipitation in all three models, but it decreases with the importance of surface runoff: 8, 4, and 1 % of *P* for the MATSIRO, CLM4, and CLSM, respectively.

These results are confirmed by the monthly time series of the main water budget components (Figure 2). In particular, the CLSM significantly underestimates ET and overestimates runoff, compared with the other two models and the corresponding observations. Besides, the strong seasonal cycle (summer peak and winter trough) of ET, as displayed by four other ET datasets in Figure 2b, is largely dampened in the CLSM simulations. Both satellite-based ET datasets demonstrate consistent mean ET estimates and ET-to-precipitation ratios (Table 1): the mean ET estimated from GLEAM (MODIS) is 2.15 (2.31) mm/d, and the ET-to-precipitation ratio is 66% (71%) for GLEAM (MODIS). The seasonality of observed ET is found to closely follow that of precipitation, with the highest ET occurring during the rainy season (Figure 2a and 2b). Despite a slight overestimation, mostly during the dry season and the beginning of the rainy season, the MATSIRO and CLM4 adequately reproduce the seasonal variability of satellite-based ET observations (Figure 2b), contrarily to the CLSM. This model underestimates/overestimates ET

during the rainy/dry season, and as the forcing data are the same for all three models, we attribute these defects to the overestimation of surface and total runoff during the rainy season, and excessive capillary rise during the dry season, probably linked to the too shallow WTD(discussed in section 6).

Regarding total runoff simulations, all models are consistent with the observations in that no streamflow occurs in the absence of rainfall, but they show an excessive runoff production at the beginning of the rainy season, which is particularly strong for the CLSM. The early runoff generation biases are also indicated in Getirana et al. (2017). Peak flows and their interannual variability are well captured by CLM4 despite minor timing errors in 2006 and 2008. The CLSM, too, captures the interannual variability of peak flows, although they are strongly overestimated, especially during the first two years, which are drier (less annual mean runoff and precipitation than the last two years, Table 1). The MATSIRO, in contrast, underestimates the peak runoff, which balances the overestimation found at the monsoon onset, so the annual mean runoff is close to observations, while the other two models exhibit a positive bias.

The monthly changes of total water storage (dS/dt) are similar in all three models and consistent with the corresponding observations, except for the slight overestimation of the peak in the summer of 2008 by both the MATSIRO and CLM4. Although all three models exhibit similar monthly dS/dt, the partitioning between monthly ET and runoff differs significantly and this discrepancy can be attributed to differences in the parameterizations linking these fluxes to the various water storage components (groundwater and soil moisture) and/or in the parameters on each model.

*b. Groundwater storage*

Figure 3 presents a comparison of the average seasonal cycles of WTD obtained from the three models and observations. A considerable discrepancy can be observed between all model simulations and observations. The observed WTD shows a sharp rise during the peak rainy season, because of substantial groundwater recharge (as analyzed below). WTD reaches to maximum (approximately 4 m) in September, one month after the maximum rainfall, and starts to decline rather steadily to the deepest observed level of the year (approximately 9.6 m) in March–April, before the onset of the monsoon. By contrast, the WTD simulated by all three models displays a lag of 1 to 2 months, and much lower seasonal amplitudes (approximately 1–1.5 m, one order of magnitude smaller than the observations), around a shallower mean depth. The CLM4 simulated WTD ranges from 2.8 m in July to 3.9 m in October, whereas the CLSM and MATSIRO simulated WTD are in the ranges of 2–3.1 and 3.2–4.7 m, respectively. Important differences between the simulated and observed WTD result from uncertainties in the groundwater system properties, particularly the depth to bedrock and the porosity. To examine the effect of these uncertainties, we measured the anomalies in the GWS (with respect to its mean over the period, ΔGWS), based on the following equation:

$$\Delta GWS = Sy \times \Delta h \qquad (2)$$

where $S_y$ is the specific yield and Δh (in meters) is the anomaly of the WTD (with respect to its mean over the period). The specific yield, also called drainable porosity, indicates the fraction of the bulk aquifer volume that can be drained out under gravity. In unconfined aquifers, it is equivalent to the effective porosity and makes the link between the changes in WTD and GWS. The GWS here is defined using Eq. 2. On the basis of magnetic resonance sounding, Séguis et al. (2011) estimated specific yield in the Donga catchment to be approximately 0.02. The corresponding values are much higher in the three models (specific yield is 0.20 in CLM4, 0.12 in

CLSM, and 0.08 in MATSIRO). It is noteworthy that specific yield is an explicit parameter of CLM4 and the MATSIRO, but not of the CLSM. In this model, the link between GWS and WTD is based on TOPMODEL's concepts (Ducharne et al. 2000). The resulting relationship is not linear but leads to an equivalent specific yield *ca* 0.12 in the Benin mesosite and the Donga basin.

Figure 4 presents a comparison of the simulated (by all three models) and observed monthly ΔGWS from 2005–2008 (averaged over the entire Donga basin) along with the monthly rainfall. Both observations and simulations exhibit strong seasonal variations in GWS, which respond to the alternation of clearly distinguishable wet and dry seasons. Like for the WTD, the observed ΔGWS rises rapidly at the onset of the rainy season in June-July, whereas the simulated ΔGWS by the models responds more slowly, with a lag of approximately 2 months relative to the observations, which is likely due to excessive early runoff generation in the models (Figure 2c). But in contrast to the WTD, the mean annual amplitude of ΔGWS is overestimated by the models: 134 mm for the MATSIRO, 206 mm for the CLSM, and 244 mm for CLM4 compared to only 112 mm based on the WTD observations and Eq. 2. This results from both a too strong decrease of GWS at the beginning of the rainy season, and a too strong GWS increase at the end of it, which is particularly marked during the two wettest years (2007 and 2008). These differences are attributed to different groundwater flow parametrizations (subsurface runoff or baseflow), different model-setup as the prescribed intrinsic parameters in each model are different e.g. specific yield, depth to bedrock, aquifer geometry, and different partitioning bias among hydrological fluxes such as surface runoff and infiltration which reduces early recharge to the groundwater system.

This overestimated annual amplitude of ΔGWS is also revealed by Figure 5. This Taylor diagram (Taylor 2001) provides a statistical summary on how well the simulations match observations in

terms of temporal correlation (angle) and standard deviation (the radial distance gives the normalized standard deviation, i.e. the standard deviation of model simulations divided by that of observations) at monthly timescale. For all three models, the normalized standard deviation is greater than 1 (Figure 5), showing that the variability of model simulations is higher than that in observations, in agreement with overestimated seasonal amplitude of ΔGWS. The correlation coefficient around 0.8 is consistent with the correct seasonality of the simulations, despite a slight phase shift (Figure 4). This delay is slightly higher in MATSIRO, which shows the lowest correlation coefficient among the three LSMs. Notice that the three models exhibit very similar temporal characteristics in the total runoff with the correlation coefficient around 0.85 in the Taylor diagram, but with rather different seasonal amplitudes. Figure 5 also indicates the lower ET seasonality simulated by the CLSM compared to the MATSIRO and CLM4.

*c. Groundwater budget*

Figure 6 shows the monthly variations in GWS (its time derivative), surface runoff, groundwater recharge, and baseflow for the period 2005–2008 as simulated by all three models. For comparison, the monthly changes of observed GWS (based on Eq. 2) are also plotted in Figure 6a. For all three models, the seasonal dynamic of the GWS is rather well captured, although the decrease of GWS after the annual peak (negative monthly GWS change) tends to start too late. The monthly changes of GWS simulated by the MATSIRO and CLSM are close to each other and consistent with the observations, while CLM4 considerably overestimates these monthly changes.

According to the groundwater budget equation, changes in GWS must be balanced by the difference: groundwater recharge - baseflow. Thus, the overestimation of GWS variations by CLM4 may be explained, at least partially, by its much higher recharge (for both peak recharge in

winter and negative recharge in summer) than the other two models, so much higher that they are plotted separately (Figure 6c and 6d). Another important difference between the three models is that only CLM4 and the CLSM simulate negative groundwater recharge (occurring when capillary rise outweighs gravity drainage) during the dry months. The negative recharge in the MATSIRO is very weak, which helps this model to simulate the highest mean recharge (Table 1). The three models simulate between 1% and 8% of annual precipitation as recharge (8% for MATSIRO, 4% for CLM4, and 1% for the CLSM). In comparison, the recharge estimates from a model-based analysis based on the calibration of routing scheme on observational river discharge, Kamagaté et al. (2007) and Séguis et al. (2011) report higher mean recharge rates, between 10% and 17% of annual precipitation.

For baseflow, the simulations of all three models differ significantly. CLM4 simulates the largest amount of baseflow (108 mm/yr, 38% of total runoff), followed by the MATSIRO (68 mm/yr, 32% of total runoff), whereas the CLSM simulates only a negligible amount of baseflow (4 mm/yr, 0% of total runoff). The baseflow ratio in all models is less than 40% of total runoff, so considerably lower than the estimate of 70%–90% reported by other studies (Kamagaté et al. 2007; Séguis et al. 2011; Hector et al. 2015). Accordingly, all three models overestimate surface runoff (ranging between 62% and 100%, Table 1), compared with the estimate of 10%–30% from the aforementioned studies. In addition, CLM4 simulates markedly higher baseflow values in the wet years of 2007 (119 mm/yr) and 2008 (216 mm/yr), four to five times greater than the other two models, because of its correspondingly high GWS changes (Figure 6a).

For these three LSMs with groundwater parametrization, the simulated groundwater budget largely depends on the effective groundwater rating curve (i.e., the relation between WTD or ΔGWS and baseflow), shown in Figure 7. When baseflow is confronted to the WTD (Figure 7a), CLM4

exhibits the steepest rating curve so the simulated baseflow in CLM4 is dissipated more rapidly (Figure 6e). CLM4 generates a significantly higher baseflow simulation than does the MATSIRO when the water table is shallower than 3 m, whereas the CLSM generates zero baseflow when the water table is deeper than 2 m. After multiplying by the corresponding specific yield, Figure 7b indicates similar relations between the ΔGWS and baseflow for the MATSIRO and CLM4, although a larger range of ΔGWS in CLM4, corresponding with higher baseflow that is also clearly shown. Thus, the different baseflow simulations (Figure 6e) can be explained by the different ΔGWS simulated among the models (Figure 4). The discrepancies in the rating curves partially explain the large sensitivity of simulated groundwater-related fluxes and GWS changes (Figure 6).

## 6. Discussion

The closure of the water budget is not perfect when based on observational estimates (Table 1), which reveal uncertainties on the observed fluxes that can amount 13%–18% of the mean precipitation. The uncertainty on observed rainfall is estimated at approximately 10% (relative error, Vischel et al. 2011). The main source of uncertainties probably arises from the satellite ET, due to the coarse spatial resolution of satellite data (MOD16: 0.05° and GLEAM: 0.25°) compared to the smaller dimensions of the forest clumps in the area. This needs to be kept in mind when comparing models to observations, but the main modeling errors revealed by the above analyses are large enough to stand against observational uncertainties.

Compared to the observational data, the simulated ΔGWS in all three LSMs shows a strong seasonal variability with a lag of approximately 2 months, especially CLM4 with larger variations of ΔGWS during the wet years. The models also simulate substantially shallower WTD than observed, with smaller seasonal variations. These seemingly paradoxical conclusions can be

reconciled given the overly high specific yield in the models. The results also show that the three models achieve rather similar simulations of ΔGWS by means of different precipitation partitioning between runoff and ET, implying different shortcomings and errors with respect to the total and groundwater budget. Therefore, improvement in model set-up and parameters, particularly specification of regional aquifer properties, is crucial and can be achieved by rigorous experiments and calibrations. The LSMs used in current study despite their similar GWS variability, do have multiple partitioning differences. For example, CLSM simulates an overly shallow WTD due to the specification of an insufficient depth to bedrock (4 m). This is a major explanation for the overestimated ET during the dry season, by means of excessive capillary rise, as also noted in temperate climate (Habets et al. 2013). The TOPMODEL concept embedded into the CLSM leads to saturation-excess runoff when the WT reaches the soil surface, but this does not happen in the studied simulation (Figure 7). The overly high surface runoff (Figure 3b) rather results from excessive infiltration-excess runoff (Horton mechanism), as also reported in the Mali mesosite of ALMIP2 (Grippa et al. 2017). This latter error propagates to excessively weak ET and recharge by CLSM during the rainy season, especially at the onset of the monsoon, which prevents capturing the sharp rise of the WT and ΔGWS (Figures 3 and 4). In contrast, the MATSIRO and CLM4 realistically reproduce the magnitude and seasonal variations of observed ET (Figure 2b). The MATSIRO and CLM4 have a much deeper depth to bedrock than the CLSM (25 m thick in CLM4, 40 m in MATSIRO), but they also fail to capture the phase of the WTD and ΔGWS, notably the sharp rise of July and August.

Eventually, a major problem shared by all three models is their overestimation of total runoff at the beginning of the rainy season, which delays the recharge to the groundwater system, consistent with the findings of Getirana et al. (2017). Furthermore, Grippa et al. (2017) compared ALMIP2

simulations with observations from the arider ALMIP2 site in Mali and also reported that the simulated total runoff was overestimated by most models with the early runoff generation and reduced infiltration, despite the different hydrological functioning of this Sahelian area. This explains why the models simulate an excessive GWS decrease at the beginning of the rainy season (Figures 4 and 6a). Correspondingly, the excessive increase of GWS at the end of the rainy season (in the same figures) may be attributed to the underestimation of baseflow and ET by the three models.

To further explore how the simulated hydrological fluxes affect ΔGWS, we changed the surface runoff generation parameterization in CLM4 by performing two experiments (Figure S1): 1) In the first experiment (EXP1) the generating factor for the surface runoff parameterization was decreased to 10 times compared to the controlled simulations (CON-SIM). 2) In the second experiment (EXP2), the generating factor of surface runoff parameterization was increased by 50% compared to the CON-SIM. In the experiments, the changes in surface runoff leads to changes in infiltration, recharge and baseflow. The results show that the total runoff simulations along with ΔGWS do not show significant differences because of the counter-balance effect of recharge and baseflow on ΔGWS. Also, the increased baseflow is compensated by the reduced surface runoff in EXP1 and vice versa in EXP2; thus, the total runoff remains nearly unchanged in both the experiments. Therefore, while the modification in parameterization can partially improve the simulations of certain water budget fluxes (e.g., recharge and baseflow), the modified simulation may not improve the simulation of overall water budget and ΔGWS. Further evaluations with a combination of other improved model parameterization and more suitable specification of groundwater parameters are required to explore for future improvement.

Furthermore, a series of sensitivity tests were carried out in CLM4 to evaluate the impact of the specification of specific yield on ΔGWS simulations. The results show that although ΔGWS is considerably influenced by these modifications (Figure S2), the modification of the specific yield value to the in-situ value (0.02) does not result in improved GWS simulations. Furthermore, there is negligible change in the total runoff simulations and WTD (Table S1). It is clear that groundwater simulation should be sensitive to the specific yield, but its value is not the only parameter which causes the uncertainty in groundwater simulations. Other factors of uncertainty sources such as the depth to bedrock, the bias in hydrological fluxes partitioning, and the uncertainty in model structures (e.g., non-representation of the lateral groundwater flow), all need to be considered in current LSMs to best improve model simulations. We recognize these limitations in our modeling study and their improvement will be investigated in our future studies.

The analysis in this study overlooks a major feature of the Donga basin, namely the co-existence of two different water tables (Séguis et al. 2011): a deep one in the saprolites, mostly related to storage processes, and surveyed by the 24 wells analyzed in this study; a seasonal perched WT, mostly related to baseflow production, as illustrated by the fact that the WTD and GWS mostly decrease between November and April, when streamflow is negligible (Figure 2c). Using chemical signatures, Séguis et al. (2011) confirmed this differentiated behavior and estimated that approximately 70–90% of the total streamflow was provided by the seasonal perched WT. In contrast, the evaluated LSMs represent a single groundwater table, contributing to baseflow, and rather shallow (Figure 3). Combined with evidence that deep roots may contribute to groundwater loss through ET in the Donga catchment (Le Lay et al. 2008; Guyot et al. 2009; Richard et al. 2013; Mamadou et al. 2016), this sheds a new light on the partitioning of the output water fluxes along the year. Since streamflow is negligible at the end of the rainy and the dry season, the decrease in

GWS is likely caused by ET. This decrease is underestimated by the models, which even predict an increase of GWS at the end of the rainy season, suggesting that when the groundwater level is deep, the models either underestimate ET, or overestimate recharge, or both. The second option is consistent with the underestimation of baseflow from the perched water table compared to the estimates of Seguis et al. (2011). The fact that baseflow originates from the perched water table is also consistent with the timing of this flux, largely synchronous with the peak rainfall. Eventually, it may be argued that the baseflow from the perched water table shows many similarities with what models call surface runoff, so the overall simulated water budget is rather satisfactory, at least for the MATSIRO and CLM4. In contrast, the CLSM suffers from excessive dry season ET and excessive runoff, which may be explained, at least partially, by the too shallow water table. Furthermore, two-dimensional groundwater representation is essential as highlighted by Zeng et al. (2016). However lateral groundwater flow is not represented in current LSMs, therefore, its representation in future can improve model simulations.

## 7. Conclusions

This study aims at evaluating the terrestrial water budget simulated by three LSMs (CLM4, CLSM, and MATSIRO) over the Donga basin (Benin, West Africa). This well-monitored 586-km$^2$ sub-basin in the upper Ouémé basin is the most humid among the three AMMA-CATCH study sites in the ALMIP2 project. All three LSMs are driven by the same atmospheric forcing data, and their simulations are evaluated against the same set of observational data. A particular emphasis of this study is placed on groundwater budget simulations since all three models include an unconfined aquifer representation, which remains a rare feature in state-of-the-art LSMs.

Interestingly, the three LSMs simulate rather consistent GWS and total water storage dynamics despite many differences in their groundwater properties (e.g., specific yield, depth to bedrock, baseflow rates) and simulated water fluxes (ET, surface runoff, baseflow, recharge). The overall water budget is well captured by the MATSIRO and CLM4, but the CLSM fails to capture the contrasted seasonal cycle of observed ET, which reveals tight links, probably too strong, with the groundwater. The considerably overestimated ET during the dry season by CLSM can be attributed to its too shallow WT which overstimulates capillary rise to the overlying soil layers. The significant underestimation of ET at the beginning of the rainy season is likely coupled to the underestimation of infiltration induced by the excessive runoff. In contrast, the MATSIRO and CLM4 achieve a favorable agreement with the satellite-derived ET, despite the consistency with the CLSM in the simulated shallow WT and overestimated runoff at the beginning of rainy season.

Major improvements can be expected from the adequate estimation of the sensitive parameters through reliable regional information, either by direct parameter specification or through calibration. For instance, the above results suggest that the infiltration capacity should be reduced in the CLSM. Regarding the groundwater compartment, the specific yield and the baseflow rate are shown to significantly contribute to the dispersion of the simulated GWS and WTD among the models. The vertical geometry of the groundwater system is also relevant, as highlighted by Swenson and Lawrence (2015) and Brunke et al. (2016) for CLM4. This point is particularly important in the Donga basin to faithfully describe the origin of baseflow and ET between the perched water table and the deeper saprolite aquifer, as detailed in the Discussion. A two-layer aquifer model might thus be useful in this particular site, and in many others as argued by de Graaf et al. (2017).

## Acknowledgments

This study was supported by the MOST 104-2923-M-002-002-MY4 to National Taiwan University, and by the French Agence Nationale de la Recherche (ANR Grant No. ANR-14-CE01-0018-01). Mehnaz Rashid is also supported by the grant of MOST 107-2811-M-002-016. ALMIP2 was supported by the French National Programs EC2CO/INSU and LEFE/INSU. AMMA-CATCH is mainly supported by IRD and CNRS-INSU (France); the ground datasets are available at http://bd.amma-catch.org/.

## References

Alcamo, J., M. Flörke, and M. Märker, 2007: Future long-term changes in global water resources driven by socio-economic and climatic changes. *Hydrol. Sci. J.*, **52**, 247–275, doi:10.1623/hysj.52.2.247.

Altchenko, Y., and K. G. Villholth, 2015: Mapping irrigation potential from renewable groundwater in Africa – a quantitative hydrological approach. *Hydrol. Earth Syst. Sci.*, **19**, 1055–1067, doi:10.5194/hess-19-1055-2015.

Beven, K. J., and M. J. Kirkby, 1979: A physically based, variable contributing area model of basin hydrology / Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant. *Hydrol. Sci. Bull.*, **24**, 43–69, doi:10.1080/02626667909491834.

Boone, A., and Coauthors, 2009a: The AMMA Land Surface Model Intercomparison Project (ALMIP). *Bull. Am. Meteorol. Soc.*, **90**, 1865–1880, doi:10.1175/2009BAMS2786.1.

——, and Coauthors, 2009b: AMMA Land Surface Model In- tercomparison Project Phase 2 (ALMIP-2). Gewex News, Vol. 9, No. 4, International GEWEX Project Office, Silver.

Brunke, M. A., and Coauthors, 2016: Implementing and Evaluating Variable Soil Thickness in the Community Land Model, Version 4.5 (CLM4.5). *J. Clim.*, **29**, 3441–3461,

doi:10.1175/JCLI-D-15-0307.1.

Cai, X., Z. Yang, Y. Xia, M. Huang, H. Wei, L. R. Leung, and B. E. Michael, 2014: Assessment of simulated water balance from Noah, Noah-MP, CLM, and VIC over CONUS using the NLDAS test bed. *J. Geophys. Res. Atmos.*, **119**, 13751–13770, doi:10.1002/2014JD022113.

de Graaf, I. E. M., R. L. P. H. van Beek, T. Gleeson, N. Moosdorf, O. Schmitz, E. H. Sutanudjaja, and M. F. P. Bierkens, 2017: A global-scale two-layer transient groundwater model: Development and application to groundwater depletion. *Adv. Water Resour.*, **102**, 53–67, doi:10.1016/J.ADVWATRES.2017.01.011.

Depraetere, C., M. Gosset, S. Ploix, and H. Laurent, 2009: The organization and kinematics of tropical rainfall systems ground tracked at mesoscale with gages: First results from the campaigns 1999-2006 on the Upper Ouémé Valley (Benin). *J. Hydrol.*, **375**, 143–160, doi:10.1016/j.jhydrol.2009.01.011.

Döll, P., 2009: Vulnerability to the impact of climate change on renewable groundwater resources: a global-scale assessment. *Environ. Res. Lett.*, **4**, 1–13, doi:10.1088/1748-9326/4/3/035006.

Ducharne, A., 2009: Reducing scale dependence in TOPMODEL using a dimensionless topographic index. *Hydrol. Earth Syst. Sci.*, **13**, 2399–2412, doi:10.5194/hess-13-2399-2009.

Ducharne, A., R. D. Koster, M. J. Suarez, M. Stieglitz, and P. Kumar, 2000: A catchment-based approach to modeling land surface processes in a general circulation model: 2. Parameter estimation and model demonstration. *J. Geophys. Res.*, **105**, 24823–24838, doi:10.1029/2000JD900328.

Elliott, J., and Coauthors, 2014: Constraints and potentials of future irrigation water availability

on agricultural production under climate change. *Proc. Natl. Acad. Sci.*, **111**, 3239–3244, doi:10.1073/pnas.1222474110.

Geiger, B., C. Meurey, D. Lajas, L. Franchisteguy, C. Dominique, and J.-L. Roujean, 2008: Near real-time provision of downwelling shortwave radiation estimates derived from satellite observations. *Meteorol. Appl.*, **15**, 411–420, doi:10.1002/met.84.

Getirana, A., Boone, A., and Peugeot, C., 2017: Streamflows over a West African basin from the ALMIP-2 model ensemble. *J. Hydrometeorol*., **18**, 1831–1845, doi: 10.1175/JHM-D-16-0233.1

Gleeson, T., J. VanderSteen, M. a. Sophocleous, M. Taniguchi, W. M. Alley, D. M. Allen, and Y. Zhou, 2010: Groundwater sustainability strategies. *Nat. Geosci.*, **3**, 378–379, doi:10.1038/ngeo881.

Grippa, M., and Coauthors, 2011: Land water storage variability over West Africa estimated by Gravity Recovery and Climate Experiment (GRACE) and land surface models. *Water Resour. Res.*, **47**, 1–18, doi:10.1029/2009WR008856.

——, M., Kergoat, L., Boone, A., Peugeot, C., Demarty, J., Cappelaere, B., Gal, L., Hiernaux, P., Mougin, E., Ducharne, A., Dutra, E., Anderson, M., and Hain, C., 2017: Modelling surface runoff and water fluxes over contrasted soils in pastoral Sahel: evaluation of the ALMIP2 land surface models over the Gourma region in Mali. *J. Hydrometeorol.*, **18**, 1847-1866., doi: 10.1175/JHM-D-16-0170.1

Gurdak, J. J., 2017: Groundwater: climate-induced pumping. *Nat Geosci.*, **10,** 71. doi:10.1038/ngeo2885.

Guyot, A., J. M. Cohard, S. Anquetin, S. Galle, and C. R. Lloyd, 2009: Combined analysis of energy and water balances to estimate latent heat flux of a sudanian small catchment. *J. Hydrol.*, **375**, 227–240, doi:10.1016/j.jhydrol.2008.12.027.

Habets, F., J. Boé, M. Déqué, A. Ducharne, S. Gascoin, A. Hachour, E. Martin, C. Pagé, E. Sauquet, L. Terray, D. Thiéry, L. Oudin, P. Viennot, 2013: Impact of climate change on surface water and ground water of two basins in Northern France: analysis of the uncertainties associated with climate and hydrological models, emission scenarios and downscaling methods. *Climatic Change*, **121**, 771-785. Doi:10.1007/s10584-013-0934-x.

Hasumi, H., and S. Emori, 2004: K-1 Coupled GCM (MIROC) Description K-1 model developers K-1 Coupled GCM ( MIROC ) K-1 model developers. *K-1 Tech. Rep. 1*, 34.

He X., H. Kim, P.E. Kirstetter, K. Yoshimura, EC Chang, CR Ferguson, JM Erlingis, Y. Hong, T. Oki. 2015. The diurnal cycle of precipitation in regional spectral model simulations over West Africa: Sensitivities to resolution and cumulus schemes. *Weather and Forecasting*, **30**, 424–445, doi:10.1175/WAF-D-14-00013.1.

Hector, B., L. Séguis, J. Hinderer, J. M. Cohard, M. Wubda, M. Descloitres, N. Benarrosh, and J. P. Boy, 2015: Water storage changes as a marker for base flow generation processes in a tropical humid basement catchment (Benin): Insights from hybrid gravimetry. *Water Resour. Res.*, **51**, 8331–8361, doi:10.1002/2014WR015773.

Jasechko, S., and Taylor, R.,G .2015. Intensive rainfall recharges tropical groundwaters. *Environmental Research Letters* 10(12): 124015. doi: 10.1088/1748-9326/10/12/124015.

Kamagaté, B., L. Séguis, G. Favreau, J. L. Seidel, M. Descloitres, and P. Affaton, 2007: Processus et bilan des flux hydriques d'un bassin versant de milieu tropical de socle au Bénin (Donga, haut Ouémé). *Comptes Rendus - Geosci.*, **339**, 418–429,

doi:10.1016/j.crte.2007.04.003.

Koirala, S., P. J. F. Yeh, Y. Hirabayashi, S. Kanae, and T. Oki, 2014: Global-scale land surface hydrologic modeling with the representation of water table dynamics. *J. Geophys. Res. Atmos.*, **119**, 75–89, doi:10.1002/2013JD020398.

Koster, R. D., and M. J. Suarez, 1992: Modeling the Land Surface Boundary in Climate Models as a Composite of Independent Vegetation Stands. *J. Geophys. Res.*, **97**, 2697–2715, doi:10.1029/91JD01696.

——, A. Ducharne, M. Stieglitz, and P. Kumar, 2000: A catchment-based approach to modeling land surface processes in a general circulation model: 1. Model structure. *J. Geophys. Res.*, **105**, 24809–24822, doi:10.1029/2000JD900327.

Lawrence, D. M., and Coauthors, 2011: Parameterization Improvements and Functional and Structural Advances in Version 4 of the Community Land Model. *J. Adv. Model. Earth Syst.*, **3**, 1–27, doi:10.1029/2011MS000045.

Le Barbé, L., T. Lebel, and D. Tapsoba, 2002: Rainfall Variability in West Africa during the Years 1950 – 90. *J. Clim.*, **15**, 187–202, doi:10.1175/1520-0442(2002)015,0187:RVIWAD.2.0.CO;2.

Le Lay, M., G. M. Saulnier, S. Galle, L. Séguis, M. Métadier, and C. Peugeot, 2008: Model representation of the Sudanian hydrological processes: Application on the Donga catchment (Benin). *J. Hydrol.*, **363**, 32–41, doi:10.1016/j.jhydrol.2008.09.006.

Lebel, T., and Coauthors, 2009: AMMA-CATCH studies in the Sahelian region of West-Africa: An overview. *J. Hydrol.*, **375**, 3–13, doi:10.1016/j.jhydrol.2009.03.020.

Lehner, B., K. Verdin, and Jarvis A, 2008: New Global Hydrography Derived from Spaceborne Elevation Data. *Eos Trans. Amarican Geophys. Union*, **89**, 93–94.

Lo, M.-H., J. S. Famiglietti, P. J.-F. Yeh, and T. H. Syed, 2010: Improving parameter estimation and water table depth simulation in a land surface model using GRACE water storage and estimated base flow data. *Water Resour. Res.*, **46**, W05517, doi:10.1029/2009WR007855.

——, P. J. F. Yeh, and J. S. Famiglietti, 2008: Constraining water table depth simulations in a land surface model using estimated baseflow. *Adv. Water Resour.*, **31**, 1552–1564, doi:10.1016/j.advwatres.2008.06.007.

——, C. M. Wu, H. Y. Ma, and J. S. Famiglietti, 2013: The response of coastal stratocumulus clouds to agricultural irrigation in California. *J. Geophys. Res. Atmos.*, **118**, 6044–6051, doi:10.1002/jgrd.50516.

Mahé, G., and J. Paturel, 2009: 1896 – 2006 Sahelian annual rainfall variability and runoff increase of Sahelian Rivers. *Comput. Rendus Geosci.*, **341**, 538–546, doi:10.1016/j.crte.2009.05.002.

Mamadou, O., Galle, S., Cohard, J.-M., Peugeot, C., Kounouhewa, B., et Zannou, A., 2016: Dynamics of water vapor and energy exchanges above two contrasting ecosystems in sudanian climate, Northern Benin (West Africa). *J. Geophys. Res. Atmos.*, **121** (19), 269-286, doi:10.1002/2016JD024749

Miralles, D. G., T. R. H. Holmes, R. A. M. De Jeu, J. H. Gash, A. G. C. A. Meesters, and A. J. Dolman, 2011: Global land-surface evaporation estimated from satellite-based observations. *Hydrol. Earth Syst. Sci.*, **15**, 453–469, doi:10.5194/hess-15-453-2011.

Mu, Q., M. Zhao, and S. W. Running, 2011: Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Remote Sens. Environ.*, **115**, 1781–1800, doi:10.1016/j.rse.2011.02.019.

Ndehedehe, C., J. Awange, N. Agutu, M. Kuhn, and B. Heck, 2016: Understanding changes in

terrestrial water storage over West Africa between 2002 and 2014. *Adv. Water Resour.*, **88**, 211–230, doi:10.1016/j.advwatres.2015.12.009.

Niu, G.-Y., Z.-L. Yang, R. E. Dickinson, and L. E. Gulden, 2005: A simple TOPMODEL-based runoff parameterization (SIMTOP) for use in global climate models. *J. Geophys. Res.*, **110**, D21106, doi:10.1029/2005JD006111.

——, ——, ——, ——, and H. Su, 2007: Development of a simple groundwater model for use in climate models and evaluation with Gravity Recovery and Climate Experiment data. *J. Geophys. Res.*, **112**, D07103, doi:10.1029/2006JD007522.

Oleson, K. W., and Coauthors, 2010: Technical Description of version 4.0 of the Community Land Model (CLM).

Panthou, G., T. Vischel, and T. Lebel, 2014: Recent trends in the regime of extreme rainfall in the Central Sahel. *Int. J. Climatol.*, **34**, 3998–4006, doi:10.1002/joc.3984.

Pavelic, P., V. Smakhtin, G. Favreau, and K. G. Villholth, 2012: Water-balance approach for assessing potential for smallholder groundwater irrigation in Sub-Saharan Africa. *Water SA, Int. Conf. Groundw. Spec. Ed.*, **38**, 399–406, doi:10.4314/wsa.v38i3.5.

Pokhrel, N. Y., S. Koirala, P. J.-F. Yeh, N. Hanasaki, L. Longuevvergne, S. Kanae, and T. Oki, 2015: Incorporation of groundwater pumping in a global Land Surface Model with the representation of human impacts. *Water Resour. Res.*, **51**, 78–96, doi:10.1002/2014WR015602.

Rasul, G., and B. Sharma, 2016: The nexus approach to water–energy–food security: an option for adaptation to climate change. *Clim. Policy*, **16**, 682–702, doi:10.1080/14693062.2015.1029865.

Redelsperger, J.-L., C. D. Thorncroft, A. Diedhiou, T. Lebel, D. J. Parker, and J. Polcher, 2006:

African Monsoon Multidisciplinary Analysis: An International Research Project and Field Campaign. *Bull. Am. Meteorol. Soc.*, **87**, 1739–1746, doi:10.1175/BAMS-87-12-1739.

Richard, A., S. Galle, M. Descloitres, J. Cohard, J. Vandervaere, L. Séguis, and C. Peugeot, 2013: Interplay of riparian forest and groundwater in the hillslope hydrology of Sudanian West Africa (northern Benin). *Hydrol. Earth Syst. Sci.*, **17**, 5079–5096, doi:10.5194/hess-17-5079-2013.

Ringler, C., and R. Lawford, 2013: The nexus across water, energy, land and food (WELF): potential for improved resource use efficiency? *Curr. Opin. Environ. Sustain.*, **5**, 617–624, doi:10.1016/J.COSUST.2013.11.002.

Séguis, L., and Coauthors, 2011: Origins of streamflow in a crystalline basement catchment in a sub-humid Sudanian zone: The Donga basin (Benin, West Africa). Inter-annual variability of water budget. *J. Hydrol.*, **402**, 1–13, doi:10.1016/j.jhydrol.2011.01.054.

Sellers, P. J., Y. Mintz, Y. C. Sud, and A. Dalcher, 1986: A simple biosphere model (SiB) for use within general circulation models. *J. Atmos. Sci.*, **43**, 505–531, doi:10.1175/1520-0469(1986)043<0505:ASBMFU>2.0.CO;2.

——, and Coauthors, 1996: A revised land surface parameterization (SiB2) for atmospheric GCMs. Part I: Model formulation. *J. Clim.*, **9**, 676–705, doi:10.1175/1520-0442(1996)009<0676:ARLSPF>2.0.CO;2.

Stieglitz, M., D. Rind, J. Famiglietti, and C. Rosenzweig, 1997: An efficient approach to modeling the topographic control of surface hydrology for regional and global climate modeling. *J. Clim.*, **10**, 118–137, doi:10.1175/1520-0442(1997)010<0118:AEATMT>2.0.CO;2.

Swenson, S. C., and D. M. Lawrence, 2015: A GRACE-based assessment of interannual groundwater dynamics in the Community Land Model, *Water Resour. Res.*, **51**, 8817 8833, doi:10.1002/2015WR017582.

Sylla, M. B., J. S. Pal, G. L. Wang, and P. J. Lawrence, 2016: Impact of land cover characterization on regional climate modeling over West Africa. *Clim. Dyn.*, **46**, 637–650, doi:10.1007/s00382-015-2603-4.

Tadesse, T., and Coauthors, 2008: The Need for Integration of Drought Monitoring Tools for Proactive Food Security Management in Sub-Saharan Africa food security management in sub-Saharan Africa. *Nat. Resour. Forum*, **32**, 265–279, doi: 10.1111/j.1477-8947.2008.00211.x.

Takata, K., S. Emori, and T. Watanabe, 2003: Development of the minimal advanced treatments of surface interaction and runoff. *Glob. Planet. Change*, **38**, 209–222, doi:10.1016/S0921-8181(03)00030-4.

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.*, ***106***(D7), 7183-7192, doi:10.1029/2000JD900719

Taylor, R., 2014: When wells run dry. *Nature* 516(7530): 179-180, doi: 10.1038/516179a.

Tapley, B. D., S. Bettadpur, M. Watkins, and C. Reigber, 2004: The gravity recovery and climate experiment: Mission overview and early results. *Geophys. Res. Lett.*, **31**, 1–4, doi:10.1029/2004GL019920.

Thibaut, A., K. Tchuente, J. Roujean, and S. Faroux, 2010: Remote Sensing of Environment ECOCLIMAP-II: An ecosystem classification and land surface parameters database of Western Africa at 1 km resolution for the African Monsoon Multidisciplinary Analysis (AMMA) project. *Remote Sens. Environ.*, **114**, 961–976, doi:10.1016/j.rse.2009.12.008.

United Nations, Department of Economic and Social Affairs, Population Division, 2015: World Population Prospects: *The 2015 Revision, Key Findings and Advance Tables*. Working Paper No. ESA/P/WP.241.

Vischel, T., G. Quantin, T. Lebel, J. Viarre, M. Gosset, F. Cazenave, and G. Panthou, 2011: Generation of high-resolution rain fields in West Africa: Evaluation of dynamic interpolation methods. *J. Hydrometeor.*, **12**, 1465–1482, doi:10.1175/JHM-D-10-05015.1.

Vörösmarty, C. J., P. Green, J. Salisbury, and R. Lammers, 2000: Global Water Resources: Vulnerability from Climate Change and Population Growth. *Science.*, **289**, 284–288, doi:10.1126/science.289.5477.284.

Vouillamoz, J. M., F. M. A. Lawson, N. Yalo, and M. Descloitres, 2015: Groundwater in hard rocks of Benin: Regional storage and buffer capacity in the face of change. *J. Hydrol.*, **520**, 379–386, doi:10.1016/j.jhydrol.2014.11.024.

Wada, Y., and M. F. P. Bierkens, 2014: Sustainability of global water use: past reconstruction and future projections. *Environ. Res. Lett.*, **9**, 104003, doi:10.1088/1748-9326/9/10/104003.

——, L. P. H. Van Beek, C. M. Van Kempen, J. W. T. M. Reckman, S. Vasak, and M. F. P. Bierkens, 2010: Global depletion of groundwater resources. *Geophys. Res. Lett.*, **37**, 1–5, doi:10.1029/2010GL044571.

Watanabe, M., and Coauthors, 2010: Improved climate simulation by MIROC5: Mean states, variability, and climate sensitivity. *J. Clim.*, **23**, 6312–6335, doi:10.1175/2010JCLI3679.1.

Wey, H. W., M. H. Lo, S. Y. Lee, J. Y. Yu, and H. H. Hsu, 2015: Potential impacts of wintertime soil moisture anomalies from agricultural irrigation at low latitudes on regional and global climates. *Geophys. Res. Lett.*, **42**, 8605–8614, doi:10.1002/2015GL065883.

Xue, Y., A. Boone, and C. M. Taylor, 2012: Review of Recent Developments and the Future

Prospective in West African Atmosphere / Land Interaction Studies. *Int. J. Geophys.*, 1–12, doi:10.1155/2012/748921.

Yeh, P. J. F., and E. A. B. Eltahir, 2005: Representation of water table dynamics in a land surface scheme. Part I: Model development. *J. Clim.*, **18**, 1861–1880, doi:10.1175/JCLI3330.1.

Yeh, P. J. F., and E. A. B. Eltahir, 2005: Representation of water table dynamics in a land surface scheme. Part II: Subgrid Variability. *J. Clim.*, **18**, 1881–1901, doi: 10.1175/JCLI3331.1

Zeng, X., and M. Decker, 2009: Improving the Numerical Solution of Soil Moisture – Based Richards Equation for Land Models with a Deep or Shallow Water Table. *J. Hydrometeorol.*, **10**, 308–319, doi:10.1175/2008JHM1011.1.

Zeng, Y., Z. Xie, Y. Yu, S. Liu, L. Wang, J. Zou, P. Qin, and B. Jia, 2016: Effects of anthropogenic water regulation and groundwater lateral flow on land processes. *J. Adv. Model. Earth Syst.*, **8**, 1106–1131, doi:10.1002/ 2016MS000646.Zeng, Y., Z. Xie, and J. Zou, 2017: Hydrologic and climatic responses to global anthropogenic groundwater extraction. *J. Clim.*, **30**, 71–90, doi:10.1175/JCLI-D-16-0209.1.

## Captions for Table and Figures

Table 1. Annual water budget terms in the Donga basin during the period 2005–2008 (precipitation, surface runoff, baseflow, total runoff, groundwater recharge, and ET, all with units of mm/yr) in the three LSMs ("MTS" denotes the MATSIRO model) and according to observations (observed streamflow and ET estimates from GLEAM ("GLM") and MODIS ("Mod 16") satellite data). Values in italic represent the percentage of each water flux with respect to precipitation, which was the same for all three models.

FIG. 1. Map of Donga Basin showing elevation in meters (Aster DEM), and the location of the main streams, the stream gauge station, and the available observation wells. The longitudes and latitudes are given in decimal degrees.

FIG. 2. Time series of the fluxes related to the total water budget (mm/d), on average over the Donga basin: (a) input precipitation; (b) simulated and observed ET; (c) simulated and observed total runoff; (d) simulated and observed total water storage changes (derived from P-E-Q, and E is the average of MODIS and GLEAM products).

FIG. 3. Simulated and observed monthly WTD from 2005 to 2008: CLM4 (blue line), CLSM (green line), MATSIRO (red line), average of observations from 24 wells (black line), grey envelope indicates the well's spatial standard deviations for each month, along with precipitation (mm; blue bars).

FIG. 4. Seasonal ΔGWS (in mm) from CLM4, CLSM, and MATSIRO against observations from 2005 to 2008. Vertical blue bars represent monthly average rainfall (mm/d) from 2005 to 2008.

FIG. 5. Taylor diagram presenting a comparison between observed (black circle) and simulated ΔGWS (circle), ET (square), and R (diamond) from 2005 to 2008. Blue is for CLM, green is for CLSM, and red is for MATSIRO. Solid black lines represent the radial distance from the origin proportional to the normalized standard deviation.

FIG. 6. Time series of fluxes related to the groundwater budget ( mm/d), on average over the Donga basin: (a) simulated and observed ΔGWS ; (b) simulated surface runoff; (c) simulated groundwater recharge from CLM4; (d) simulated groundwater recharge from CLSM and MATSIRO; (e) simulated baseflow.

FIG. 7. Monthly effective groundwater rating curves (a) WTD and (b) ΔGWS versus monthly baseflow simulations during 2005–2008 in the Donga catchment, as obtained from three LSMs. Blue lines are the baseflow simulated by the SIMTOP scheme in CLM4 (without the saturation excess from the soil layers).

| | | 2005 | | 2006 | | 2007 | | 2008 | | Average | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **P** | | 1130 | | 1062 | | 1253 | | 1343 | | 1197 | |
| **Qs** | **CLM** | 219 | ***19*** | 191 | ***18*** | 230 | ***18*** | 280 | ***21*** | 230 | ***19*** |
| | **CLSM** | 452 | ***40*** | 407 | ***38*** | 524 | ***42*** | 607 | ***45*** | 498 | ***42*** |
| | **MTS** | 145 | ***13*** | 116 | ***11*** | 131 | ***10*** | 194 | ***14*** | 147 | ***12*** |
| **Qsb** | **CLM** | 62 | ***5*** | 35 | ***3*** | 119 | ***10*** | 216 | ***16*** | 108 | ***9*** |
| | **CLSM** | 0 | ***0*** | 0 | ***0*** | 3 | ***0*** | 12 | ***1*** | 4 | ***0*** |
| | **MTS** | 59 | ***5*** | 46 | ***4*** | 61 | ***5*** | 107 | ***8*** | 68 | ***6*** |
| **Q (Qs+Qsb)** | **CLM** | 280 | ***25*** | 226 | ***21*** | 350 | ***28*** | 497 | ***37*** | 338 | ***28*** |
| | **CLSM** | 453 | ***40*** | 408 | ***38*** | 528 | ***42*** | 619 | ***46*** | 502 | ***42*** |
| | **MTS** | 204 | ***18*** | 163 | ***15*** | 192 | ***15*** | 301 | ***22*** | 215 | ***18*** |
| | **Obs** | 129 | ***11*** | 134 | ***13*** | 194 | ***15*** | 307 | ***23*** | 191 | ***16*** |
| **R** | **CLM** | -56 | ***-5*** | -37 | ***-3*** | 141 | ***11*** | 160 | ***12*** | 52 | ***4*** |
| | **CLSM** | 0 | ***0*** | 1 | ***0*** | 35 | ***3*** | 27 | ***2*** | 16 | ***1*** |
| | **MTS** | 59 | ***5*** | 40 | ***4*** | 113 | ***9*** | 155 | ***12*** | 92 | ***8*** |
| **ET** | **CLM** | 866 | ***77*** | 840 | ***79*** | 881 | ***70*** | 848 | ***63*** | 859 | ***72*** |
| | **CLSM** | 677 | ***60*** | 641 | ***60*** | 675 | ***54*** | 725 | ***54*** | 679 | ***57*** |
| | **MTS** | 936 | ***83*** | 906 | ***85*** | 966 | ***77*** | 958 | ***71*** | 942 | ***79*** |
| | **Mod16** | 862 | ***76*** | 854 | ***80*** | 842 | ***67*** | 839 | ***62*** | 849 | ***71*** |
| | **GLM** | 817 | ***72*** | 734 | ***69*** | 780 | ***62*** | 827 | ***62*** | 790 | ***66*** |

Table 1. Annual water budget terms in the Donga basin during the period 2005–2008 (precipitation, surface runoff, baseflow, total runoff, groundwater recharge, and ET, all with units of mm/yr) in the three LSMs (“MTS” denotes the MATSIRO model) and according to observations (observed streamflow and ET estimates from GLEAM (“GLM”) and MODIS (“Mod 16”) satellite data). Values in italic represent the percentage of each water flux with respect to precipitation, which was the same for all three models.

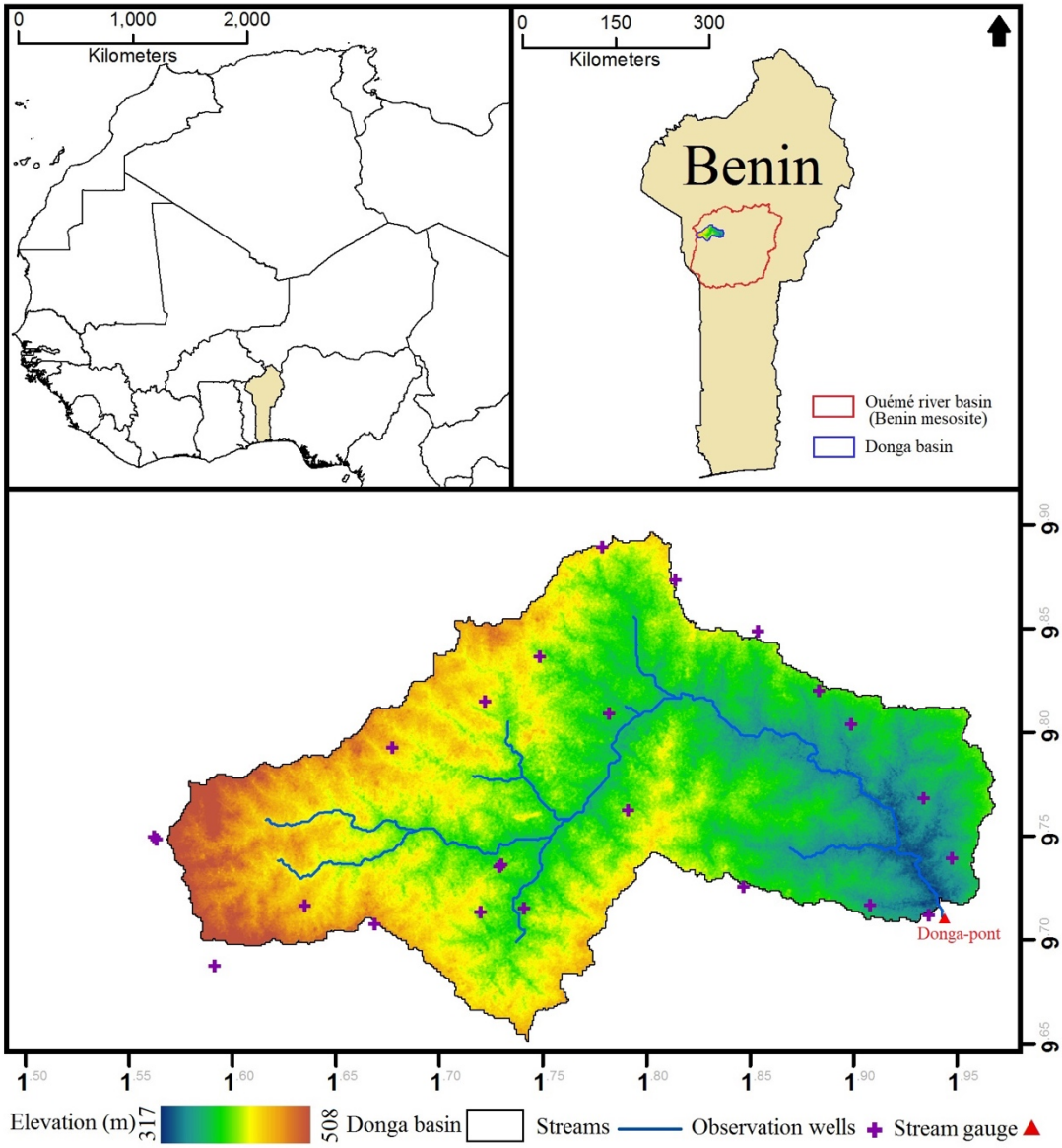


FIG. 1. Map of Donga Basin showing elevation in meters (Aster DEM), and the location of the main streams, the stream gauge station, and the available observation wells. The longitudes and latitudes are given in decimal degrees.

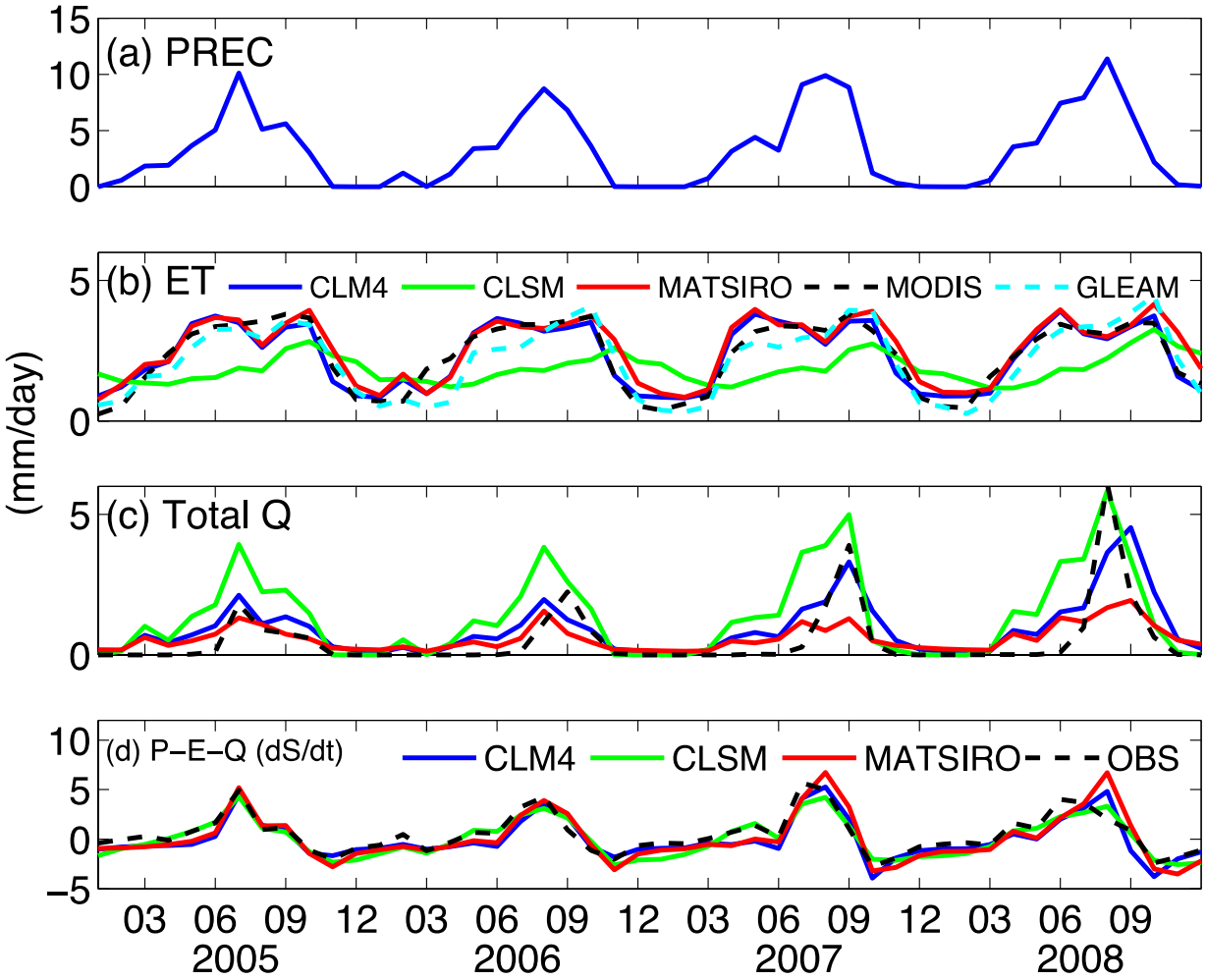


FIG. 2. Time series of the fluxes related to the total water budget (in mm/d), on average over the Donga basin: (a) input precipitation; (b) simulated and observed ET; (c) simulated and observed total runoff; (d) simulated and observed total water storage changes (derived from P-E-Q, and E is the average of MODIS and GLEAM products).

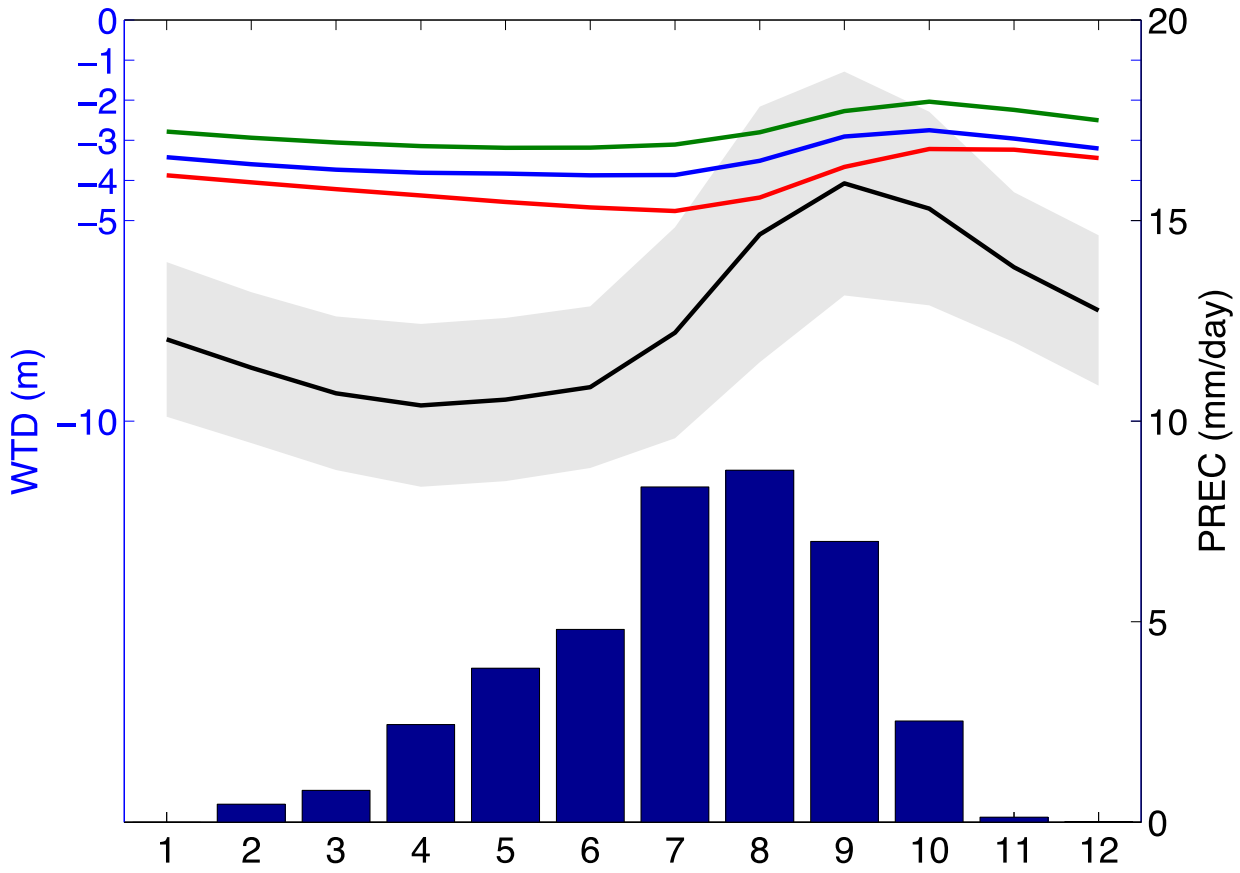


FIG. 3. Simulated and observed monthly WTD from 2005 to 2008: CLM4 (blue line), CLSM (green line), MATSIRO (red line), average of observations from 24 wells ( black line), grey envelope indicates the well's spatial standard deviations for each month, along with precipitation (in mm; blue bars).

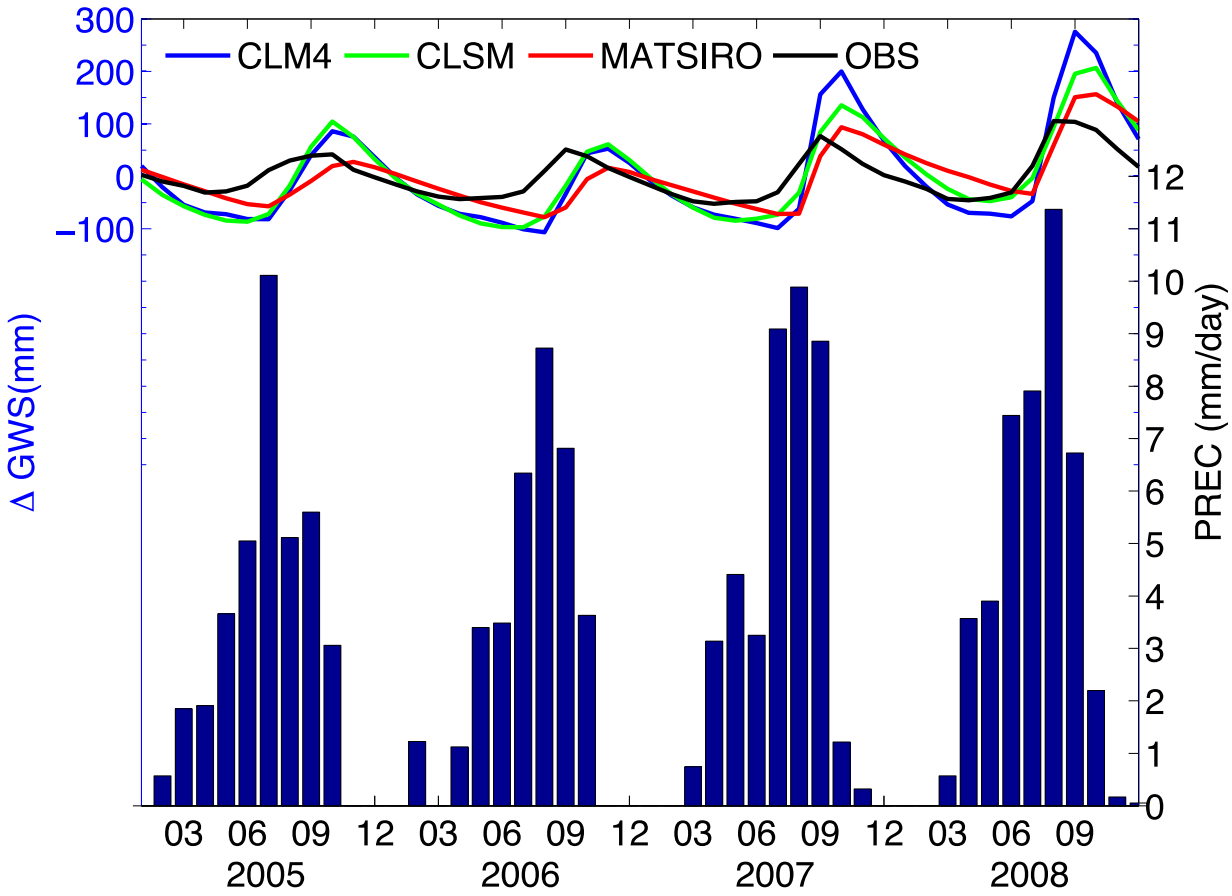


FIG. 4. Seasonal$\Delta$GWS (mm) from CLM4, CLSM, and MATSIRO against observations from 2005 to 2008. Vertical blue bars represent monthly average rainfall (mm/d) from 2005 to 2008.

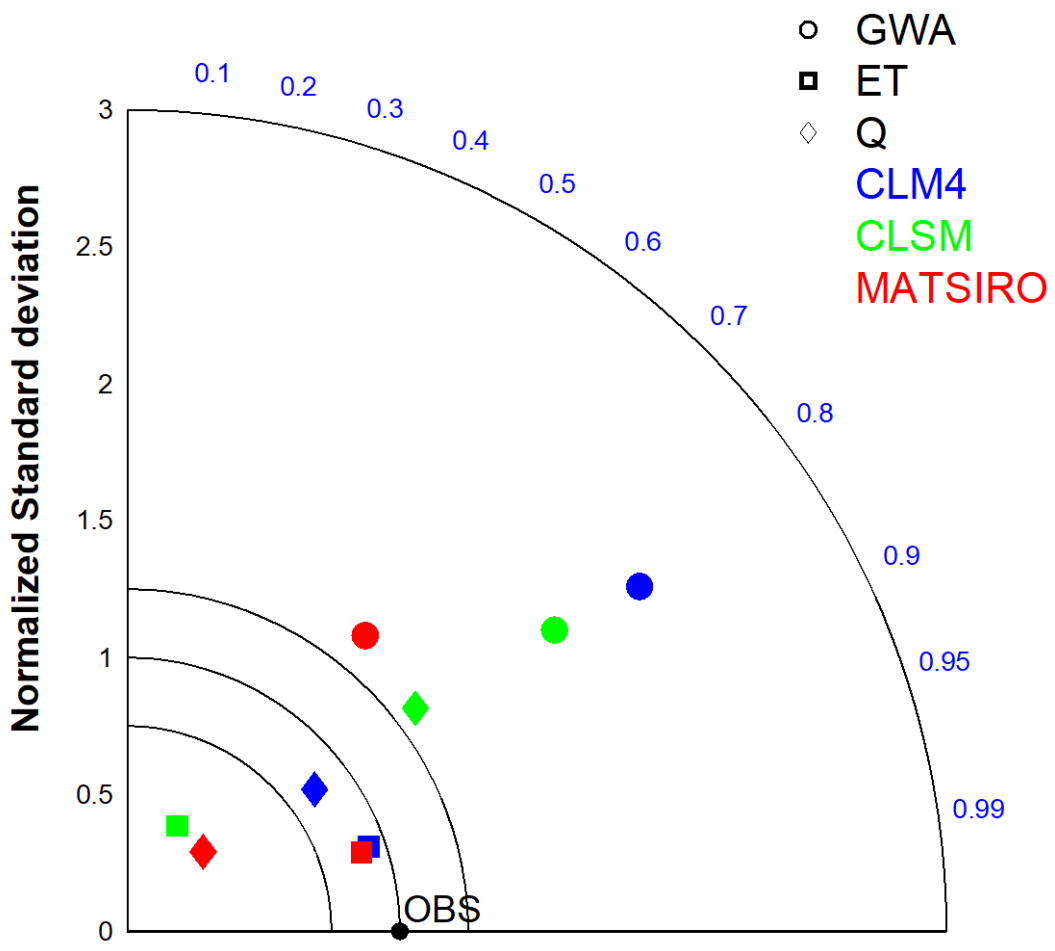


FIG. 5. Taylor diagram presenting a comparison between observed (black circle) and simulated ΔGWS (circle), ET (square), and R (diamond) from 2005 to 2008. Blue is for CLM4, green is for CLSM, and red is for MATSIRO. Solid black lines represent the radial distance from the origin proportional to the normalized standard deviation.

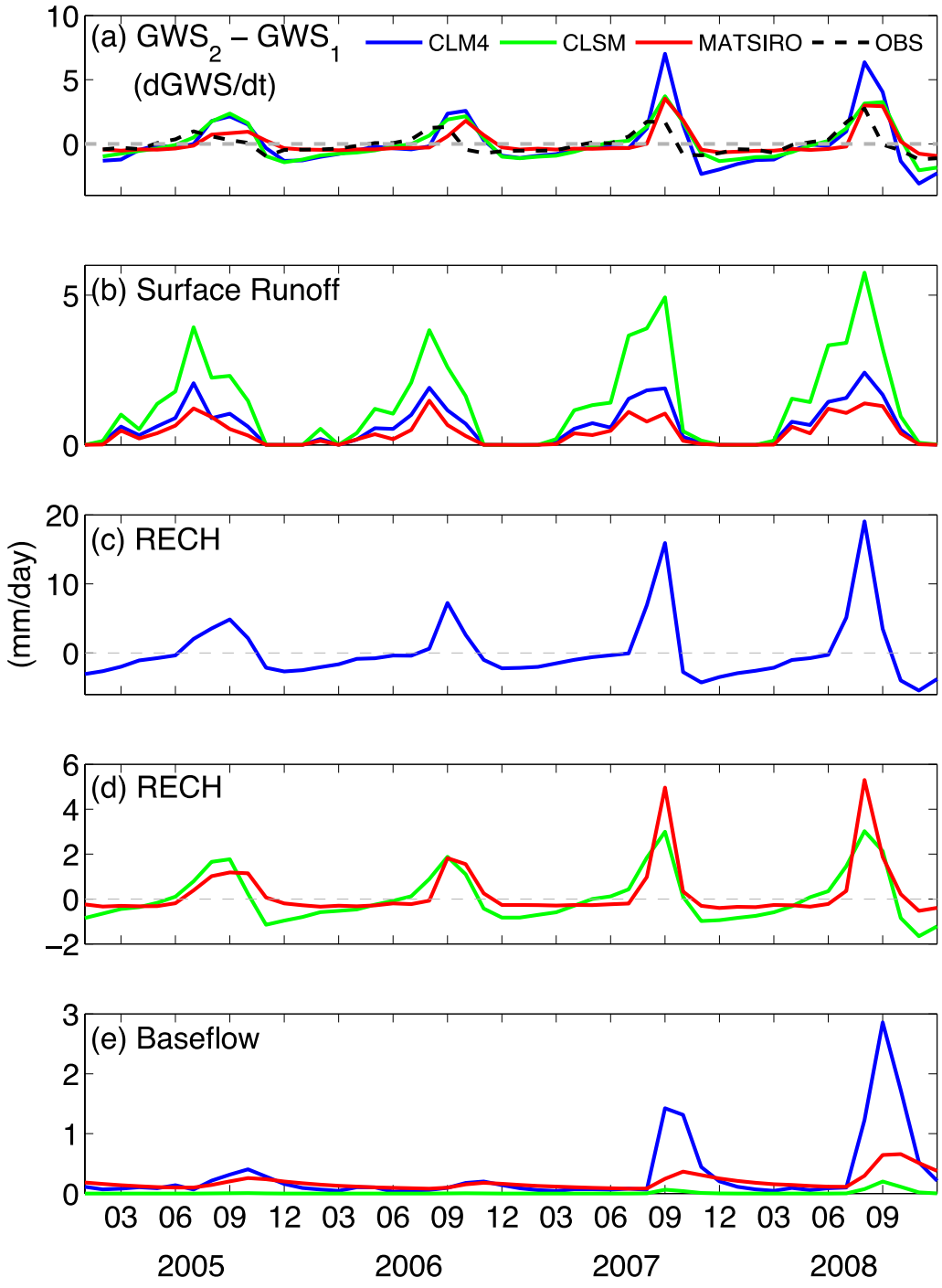


FIG. 6. Time series of fluxes related to the groundwater budget (mm/d), on average over the Donga basin: (a) simulated and observed ΔGWS; (b) simulated surface runoff; (c) simulated groundwater recharge from CLM4; (d) simulated groundwater recharge from CLSM and MATSIRO; (e) simulated baseflow.

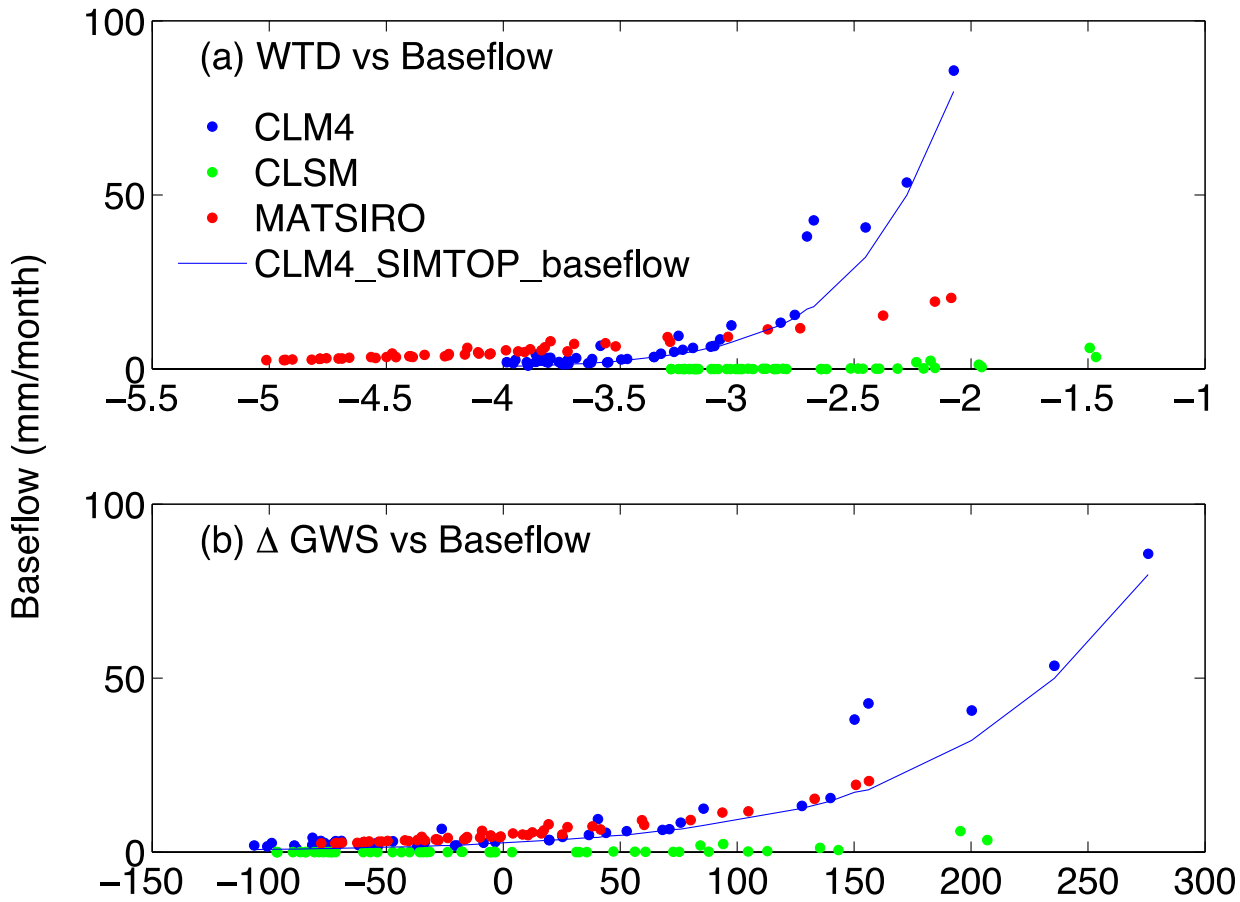


FIG.7. Monthly effective groundwater rating curves (a) WTD and (b) ΔGWS versus monthly baseflow simulations during 2005–2008 in the Donga catchment, as obtained from three LSMs. Blue lines are the baseflow simulated by the SIMTOP scheme in CLM4 (without the saturation excess from the soil layers).