Hydrology and Earth System Sciences

Calibration and evaluation of a semi-distributed watershed model of Sub-Saharan Africa using GRACE data

H. Xie[1], L. Longuevergne[2], C. Ringler[1], and B. R. Scanlon[3]

[1]International Food Policy Research Institute, 2033 K Street NW, Washington D.C. 20006, USA
[2]CNRS – UMR 6118, Géosciences Rennes Université Rennes 1, 35042 Rennes, France
[3]Bureau of Economic Geology, Jackson School of Geosciences, University of Texas, Austin, TX 78713-8926, USA

Correspondence to: H. Xie (h.xie@cgiar.org)

Abstract. Irrigation development is rapidly expanding in mostly rainfed Sub-Saharan Africa. This expansion underscores the need for a more comprehensive understanding of water resources beyond surface water. Gravity Recovery and Climate Experiment (GRACE) satellites provide valuable information on spatio-temporal variability in water storage. The objective of this study was to calibrate and evaluate a semi-distributed regional-scale hydrologic model based on the Soil and Water Assessment Tool (SWAT) code for basins in Sub-Saharan Africa using seven-year (July 2002–April 2009) 10-day GRACE data and multi-site river discharge data. The analysis was conducted in a multi-criteria framework. In spite of the uncertainty arising from the trade-off in optimising model parameters with respect to two non-commensurable criteria defined for two fluxes, SWAT was found to perform well in simulating total water storage variability in most areas of Sub-Saharan Africa, which have semi-arid and sub-humid climates, and that among various water storages represented in SWAT, water storage variations in soil, vadose zone and groundwater are dominant. The study also showed that the simulated total water storage variations tend to have less agreement with GRACE data in arid and equatorial humid regions, and model-based partitioning of total water storage variations into different water storage compartments may be highly uncertain. Thus, future work will be needed for model enhancement in these areas with inferior model fit and for uncertainty reduction in component-wise estimation of water storage variations.

1 Introduction

Sub-Saharan Africa (SSA) is used as a collective term that refers to African nations which lie (or partially lie) south of the Sahara. The region makes up about 80 % of the African and 10 % of the global population. Agriculture forms the backbone of the SSA economy; however, SSA countries largely missed the green revolution. The agricultural productivity in SSA countries remains low relative to other parts of the world and the region is still beset with food insecurity. The number of estimated undernourished people in SSA in 2010 reached 239 million (FAO, 2010). SSA is also the only region where childhood malnutrition is projected to increase as a result of rapid population growth, climate change and continued low productivity in agriculture (Rosegrant et al., 2009). Annual population growth in SSA is 2.2 %, much higher than global average of 1.1 % (World Bank, 2009). In addition, SSA is regarded as the region with a particularly low capacity to adapt to climate change (IPCC, 2007).

Sustainable intensification of agriculture, with a focus on irrigation development, is considered a key pillar for increasing agricultural productivity in SSA (Rosegrant et al., 2002; Molden, 2007; Rockström et al., 2007). SSA straddles the Equator and is dominated by tropical and sub-tropical climate. Rainfall in SSA is highly variable both spatially and temporally and constitutes a more critical factor than temperature for agriculture. Limited water availability, particularly during droughts, is a key reason for crop failure, especially considering the fact that SSA agriculture is predominantly rainfed with only 3 % of the cultivated area irrigated (Siebert, 2010; FAO, 2011). Both international development

banks and African governments have pledged to significantly increase irrigation development to address low agricultural productivity, rural poverty and food security challenges in the region.

Significant expansion of irrigated agriculture in SSA, however, will require a more comprehensive understanding of water resources in the region. Mathematical models are important tools for scientific investigation and to support policy decisions. They provide a feasible and economical way to explore key hydrologic processes and to evaluate alternative management options where direct observation and experimentation are not possible, are costly, or both. However, hydrologic modelling is challenging, particularly for regions with limited data. Models are only a rough representation of reality. Model calibration and evaluation using historical monitoring data is critical before the model is used to provide reliable results.

In this study, we present the calibration and evaluation of a regional-scale semi-distributed watershed model using GRACE data. The model was developed to simulate hydrology in Sub-Saharan African countries, and the model development is based on the Soil and Water Assessment Tool (SWAT) code (Arnold et al., 1998). SWAT is a physically-based comprehensive river basin model with a proven track record of successful application globally, including in agricultural water management (e.g., Kim et al., 2008; Xie et al., 2008, 2011; Dhar and Mazumdar, 2009; Oeurng et al., 2011). The size of the study river basins in reported SWAT applications typically range from a few square kilometres to tens of thousands of square kilometres. However, the model also shows potential for watershed studies at very large scales. Related to Africa, SWAT was applied to West Africa and to the entire continent to estimate blue and green water resources by Schuol et al. (2008a, b).

Conventional processes for calibrating and validating hydrologic models generally use stream discharge data. In previous SWAT applications in Africa, the model was calibrated and validated using river discharge time series data on monthly basis (Schuol et al., 2008a, b). In the model calibration and evaluation study reported in this paper, satellite-based observations of total water storage (TWS) variations derived from the Gravity Recovery And Climate Experiment (GRACE) were used to complement discharge data. GRACE is a joint mission launched in 2002 by NASA and the German Space Agency (DLR) to accurately map the Earth's gravity field (Tapley et al., 2004). After corrections for tidal and atmospheric mass variations, the hydrologic cycle is the primary source of variations in the Earth's gravity field on the continents (Schmidt et al., 2008). Variations in TWS (i.e., water storages variations integrated vertically over all water storage layers) can be inferred from the GRACE gravity signal.

Including additional state observations other than river discharge expands the data base for model evaluation and may help generate additional insights into model perfor-mance (Parajka et al., 2006; Fenicia et al., 2008; Konz and Seibert, 2010). In this study, the merits of incorporating GRACE-based hydrologic observations into the calibration and evaluation of the SWAT model of Sub-Saharan Africa (SWAT-SSA) are two-fold. Firstly, the river systems in SSA are poorly monitored and many river basins are ungauged. GRACE-based TWS variations have a global coverage and, thus, offer the opportunity to calibrate and evaluate the model for those areas where river discharge data are not available or sparse. Secondly, river discharge is part of "blue" water, which is the traditional focus of water resources planning and management, but only accounts for a small portion of total water resources. Over the past decade, the definition of agricultural water management has widened to include the entire hydrologic cycle (e.g., Falkenmark and Rockström, 2006). GRACE data provide direct estimates of TWS to help verify the capacity of the SWAT model to simulate spatio-temporal variability across all water balance components.

GRACE has been widely used to monitor changes in water mass redistribution for various basins globally. For example, GRACE has been used to quantify changes in water storage in response to droughts with a specific focus on groundwater systems (Leblanc et al., 2009; Chen et al., 2010). Water storage in East African Great Lakes was estimated using GRACE data as well (Becker et al., 2010). Many studies evaluated groundwater depletion related to irrigation (NW India: Rodell et al., 2009; California, US: Famiglietti et al., 2011) with some studies emphasising ground referencing using well data (Longuevergne et al., 2010; Scanlon et al., 2012). Good correspondence was found between GRACE-based storage estimates and well data within uncertainty envelopes of GRACE-based estimates. In addition to basin scale and global studies of changes in water storage, GRACE is also widely used in modelling studies to condition land surface models (Guntner, 2008) and for data assimilation (Zaitchik et al., 2008). To date most studies that use GRACE data for model conditioning are limited to model validation without significant calibration or model parameter tuning to GRACE data (Niu and Yang, 2006; Ngo-Duc et al., 2007; Syed et al., 2008; Yirdaw et al., 2009; Alkama et al., 2010; Tang et al., 2010; Grippa et al., 2011; Yang et al., 2011). Werth et al. (2009, 2010) may be the first to present calibration analyses for water storage variability in global hydrologic modelling using GRACE data. In their studies, GRACE-based water storage variations were used to calibrate and validate the WaterGAP Global Hydrology Model (WGHM) for 28 major river basins globally. More recently, Milzow et al. (2011) combined GRACE data with altimetry and Synthetic Aperture Radar (SAR) surface soil moisture data to calibrate and validate the SWAT model for the Okavango catchment in Southern Africa. The study presented in this paper is focused on the calibration and evaluation of the SWAT model at a regional scale. The modelled area covers all SSA. Furthermore, GRACE data used in this study have a
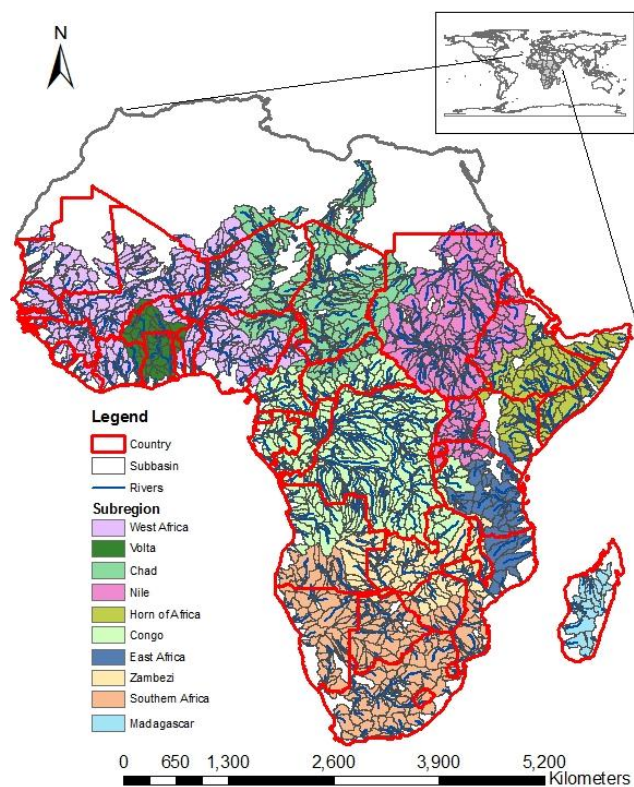
Fig. 1. Study area boundary, sub-region division and watershed delineation in SWAT-SSA model setup.

finer temporal scale (10-day) relative to the typical monthly interval for most GRACE products.

The rest of the paper is organised as follows: the setup of the SWAT model is described in Sect. 2, and the key datasets and steps for calibrating and evaluating the SWAT-SSA models are described in Sect. 3 through 5. Section 6 presents the results of the model calibration and validation. A summary of the major findings from this study and their implications are provided in Sect. 7.

2   SWAT model setup

The model area in this study is ~ 21 million km^2 (Fig. 1). The major datasets used for the setup or initial parameterisation of the SWAT-SSA model are listed in Table 1. The data acquisition and processing strategy in our study are similar to those described in Schuol et al. (2008a, b), but updated data or alternative options were selected in most cases.

The drainage topology of the study region is represented in SWAT by partitioning the river basins into subbasins and defining the corresponding drainage network of the river system with one river channel segment in each subbasin. Elevation data used in this step of watershed delineation were clipped from the HydroSHEDS database (Lehner et al., 2008). HydroSHEDS is a derivative mapping product

Table 1. The datasets for SWAT model setup.

Category	Source
Elevation	HydroSHEDS
Soil	Harmonized world soil database (HWSD)
Land cover	Global land cover (GLC) 2000
Lakes & reservoirs	Global lake and wetland database (GLWD)
Climate	Precipitation: Global Precipitation Climate Project (GPCP) Temperature and relative humidity: Goddard Earth Observing System model version 4 and version 5 (GEOS-4 & GEOS-5) Solar radiation: Release 3 of the NASA/GEWEX Surface Radiation Budget (GEWEX SRB 3.0) project and NASA's Fast Longwave And SHortwave Radiative Fluxes (FLASHFlux) project

from NASA's 3 arc-second (approximately 90 m in equatorial area) SRTM (Shuttle Radar Topography Mission) elevation data and is the best currently available (with highest resolution) hydrologically conditioned digital elevation dataset for SSA. Based on topographic analysis of HydroSHEDS elevation data, SSA was divided into 1488 subbasins (Fig. 1). Furthermore, SWAT is a semi-distributed watershed model with potentially different parameter values for different subbasins/reaches. To reflect the spatial variability of SWAT parameters in calibration and considering computational efficiency, the SSA was divided into 10 sub-continental regions; one SWAT model was set up and calibrated separately for each sub-continental region (Fig. 1; Table 2).

Within a subbasin, SWAT allows multiple hydrologic response units (HRUs) to be defined that reflect spatial variability in soil and land cover distributions. However, due to computational limitations, only one HRU with the dominant land cover and soil was created for each subbasin (Winchell et al., 2007). The soil data were obtained from the Harmonized World Soil Database (HWSD, v. 1.1, FAO/IIASA/ISRIC/ISSCAS/JRC, 2009) and the land cover data were obtained from the Global Land Cover 2000 database (European Commission, Joint Research Centre, 2003, http://bioval.jrc.ec.europa.eu/products/glc2000/glc2000.php). The HWSD contains updated soil data for eastern, central and southern African countries relative to the FAO/UNESCO Soil Map of the World. The soil attribute data in HWSD meet most requirements for SWAT model parameterisation; however, two important parameters that describe the hydrologic properties of soils

Table 2. Sub-regions in SWAT-SSA model setup and calibration.

Name	Area ($\times 10^3$ km^2)	Name	Area ($\times 10^3$ km^2)
West Africa	3550	Congo	4474
Volta	534	Eastern Africa	808
Chad	2576	Southern Africa	2928
Nile	2841	Zambezi	1365
Horn of Africa	1477	Madagascar	309
Total: 20 862			

(available water capacity and saturated hydraulic conductivity) are missing and were estimated using pedotransfer functions (Saxton et al., 1986; Schaap et al., 2001).

Climate forcing data for SWAT include 1 degree daily (1DD) precipitation, temperature, solar radiation and relative humidity data and were obtained from the NASA Langley Research Center POWER Project. A GIS layer of polygon-based SWAT subbasin boundaries was overlaid with a GIS layer of climate gridded data to calculate fractions of area covered by different climate data grid cells for each subbasin and to compute area-weighted values of climatic variables as basin-wide estimates of these variables. The original source of the precipitation data is the Global Precipitation Climate Project (GPCP, http://precip.gsfc.nasa.gov). The 1-DD GPCP dataset combined observations from multiple sensors (Huffman et al., 2001). The data for other climate variables were compiled from various NASA's projects (Table 1) by the POWER project and were primarily used for estimation of potential evapotranspiration (PET). SWAT includes three different methods for estimating PET (Neitsch et al., 2005) with varying data requirements and the Priestley-Taylor method (Priestley and Taylor, 1972) was selected because it is considered more accurate than the Hargreaves method (Hargreaves and Samani, 1985), which is temperature-based, and reliable estimates of wind speed required for the Penman-Monteith method (Monteith, 1965) were not available at the time of this study.

SWAT also provides two options for simulating flow routing in river channels. The variable storage method was used to route water in river channels because pilot simulations suggested that this is more robust than the Muskingum option in this study. Anthropogenic impacts on water resources were considered to be negligible in SSA. Agriculture is the dominant water use sector. However, current agriculture in SSA is mainly rainfed; the area of SSA equipped for irrigation only accounts for 3 % of the total cultivated area (Siebert, 2010; FAO, 2011). Therefore, existing irrigation was not simulated in this study.

SSA has a number of large fresh water bodies, such as Lake Victoria, the world's second largest fresh water lake in terms of surface area (239 000 km^2), and Lake Volta, the world's largest reservoir in terms of surface area (8502 km^2).
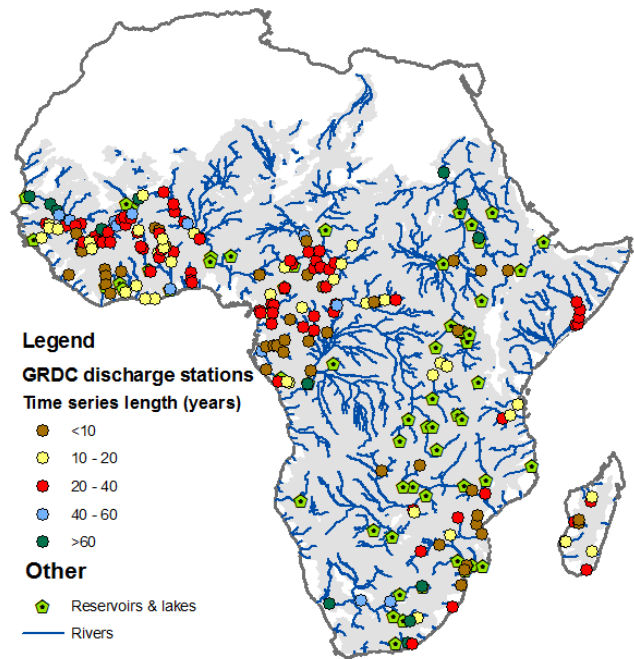

Fig. 2. Global Runoff Data Centre (GRDC) stations and reservoirs/lakes included in the SWAT-SSA model setup and calibration (lengths of GRDC discharge data series are marked with missing values excluded).

The major lakes and reservoirs in SSA were defined in our SWAT-SSA model (Fig. 2). Locations and storage capacities of these water impoundments were obtained from the Global Lakes and Wetlands Database (GLWD) (Lehner and Döll, 2004). Due to signal leakage, mass variations in these lakes and reservoirs may have a significant contribution to GRACE TWS observations (e.g., Becker et al., 2010), even if their size is much less than the GRACE footprint ($\sim$ 450 km, i.e., 200 000 km^2 in area). We compared the simulated water level change data and water level change data obtained with satellite altimetry (Crétaux et al., 2011, see http://www.legos.obs-mip.fr/en/soa/hydrologie/hydroweb/) and found that it is difficult to adequately simulate water storage variations in these lakes and reservoirs because of lack of detailed bathymetry and reservoir operation data. In this study, an alternative modelling strategy was taken, i.e., lake and reservoir mass correction was applied to GRACE TWS data according to height and volume variations of the 22 largest lakes and reservoirs in SSA (Table 3) from satellite altimetry data analysis for a fair comparison between GRACE and hydrologic model. Accordingly, simulated water mass variations in lakes and reservoirs were excluded from the model-based TWS variation calculation.

Table 3. Lakes/reservoirs to which mass corrections in GRACE data processing were applied.

1. Albert	6. Kainji	11. Malawi	17. Tana	22. Volta
2. Bangweulu	7. Kariba	12. Mweru	18. Tanganyika	
3. Buyo	8. KhashmGirba	13. Nasser	19. Chad	
4. CahoraBassa	9. Kivu	15. Roseires	20. Turkana	
5. Edward	10. Kyoga	16. Rukwa	21. Victoria	

Table 4. SWAT hydrologic calibration parameters.

Parameter	Level	Possible range	
$CN2$:	SCS curve number	HRU/subbasin	$35 \sim 90$
$ESCO$:	Soil evaporation compensation factor	Basin	$0 \sim 1$
GW_DELAY:	Groundwater delay coefficient [days]	HRU/subbasin	$0 \sim 100$
GW_REVAP:	Groundwater revap coefficient	HRU/subbasin	$0.02 \sim 0.2$
$ALPHA\_BF$:	Baseflow alpha factor [days]	HRU/subbasin	$0 \sim 1$
$REVAPMN$:	Threshold depth to water in the shallow aquifer for "revap" to occur [mm]	HRU/subbasin	$0 \sim 500$
$GWQMN$:	Threshold depth to water in the shallow aquifer required for groundwater flow to occur [mm H_2O]	HRU/subbasin	$0 \sim 1000$
$SURLAG$:	Surface runoff lag coefficient	Subbasin	$1 \sim 10$
SOL_AWC_X[a]:	Calibration factor for soil water available capacity	Soil layer/subbasin	$0.5 \sim 2$
SOL_D_X[a]:	Calibration factor for depth from soil surface to bottom of layer	Soil layer/sbubasin	$1 \sim 2$
SOL_K_X[a]:	Calibration factor for saturated hydraulic conductivity	Soil layer/sbubasin	$0.5 \sim 1.5$

[a] These factors are defined for model calibration purposes only. Actual values of these parameters used in simulation are equal to their default values multiplied by the calibration factors.

3 GRACE data

GRACE data used in this study were obtained from CNES-GRGS (Centre National d'Etudes Spatiales-Groupe de Recherches de Géodésie Spatiale), RL2 product (Bruinsma et al., 2010). The data are provided as spherical harmonics as 10 day means. The Stokes coefficients are truncated at degree 50 to remove high frequency noise. No further filtering is required for these solutions. Stokes coefficients were recombined following Wahr et al. (1998) and projected on a 0.5°latitude/longitude grid. In terms of time frame of the data, we used 232 10-d periods from 29 July 2002 through 22 April 2009 (with missing values from 26 November 2002 to 23 February 2003, 25 May 2003 to 3 July 2003, and 20 January 2004 to 29 January 2004).

The GRACE CSR RL04 product (Center for Space Research, University of Texas at Austin, monthly timescales, Bettadpur, 2007) was also used to estimate GRACE errors at a 10-day timescale. CSR data were destriped according to Swenson and Wahr (2006). For error calculation, both GRGS and CSR GRACE products were truncated at degree 30 and smoothed using a 300-km Gaussian smoother to evaluate large-scale errors. Error is computed at a monthly time step as the difference between CSR and GRGS data and resampled as 10-day errors.

In the mass correction, the impact of 22 lakes and reservoirs were first forward modelled at GRACE GRGS resolution, prior to subtraction from GRACE. Lake volume variations were distributed on a grid and projected on spherical harmonics. They were then recombined up to degree 50 on a 0.5 degree grid.

4 Total water storage variation calculation in SWAT

SWAT was developed to provide continuous simulations of the basin hydrology at a daily timescale. For each day of the simulation, the model first computes the water yields on land and then routes the water through the defined river channel network. In the land phase simulation, SWAT uses the Soil Conservation Service (SCS) curve number method (SCS 1972) to estimate the volume of overland flow and storage routing techniques to simulate percolation and lateral movement of water through the soil profile. The water leaving the base of the soil profile does not enter aquifers immediately, but is time lagged based on transport through the vadose zone. The vadose zone is defined in SWAT as the unsaturated zone beneath the base of the soil profile and above the groundwater table. An exponential decay weighting function, proposed by Venetis (1969), was used to account for the time delay of water drainage in the vadose zone and to
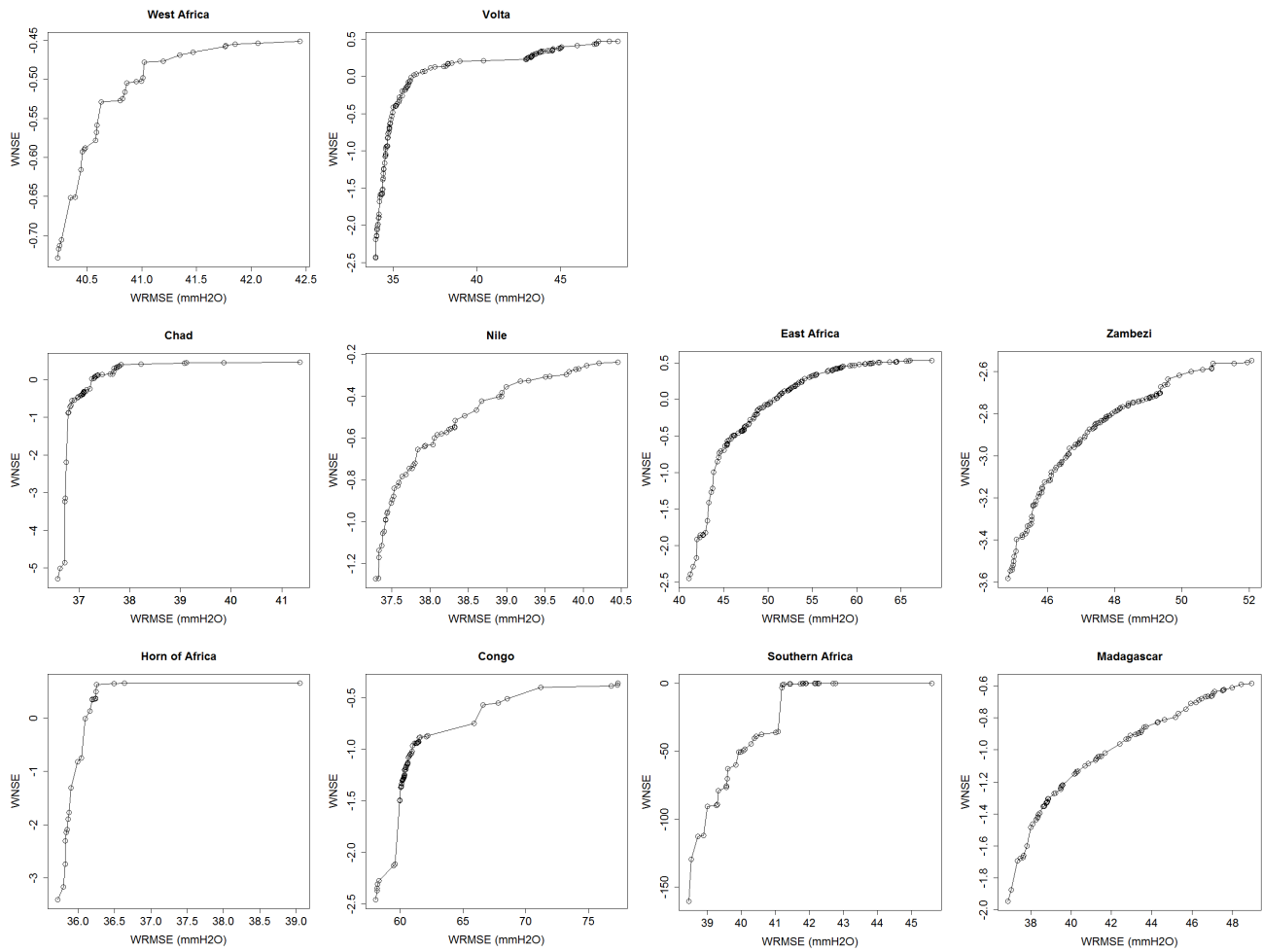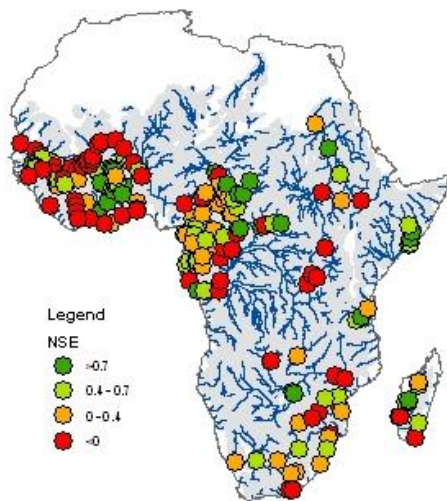
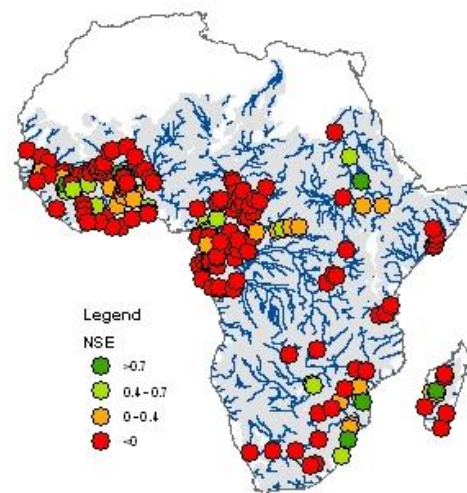**Fig. 3.** The Pareto fronts in objective space.

Fig. 4. Model fit in river discharge simulations.

Table 5. Nash-Sutcliffe efficiency coefficients for calibrated SWAT models for simulation of variations in TWS.

	Calibration			Validation		
	Mean	Max.	Min.	Mean	Max.	Min.
West Africa	0.91	0.92	0.89	0.91	0.92	0.90
Volta	0.86	0.90	0.75	0.91	0.94	0.82
Chad	0.81	0.83	0.79	0.66	0.73	0.55
Nile	0.86	0.87	0.84	0.75	0.78	0.72
Horn of Africa	0.41	0.45	0.40	0.31	0.34	0.25
Congo	0.19	0.34	−0.44	0.12	0.30	−0.80
Eastern Africa	0.85	0.91	0.71	0.80	0.87	0.67
Zambezi	0.91	0.92	0.90	0.91	0.93	0.89
South Africa	0.46	0.54	0.40	0.80	0.83	0.75
Madagascar	0.81	0.85	0.76	0.82	0.84	0.79

predict effective recharge into shallow aquifers (Sangrey et al., 1984). Variation in groundwater flow to rivers is linearly related to changes in water-table elevation.

SWAT simulates several water storage components that make up total water storage to compare with GRACE TWS. These storages include:

1. Overland water storage (V_1), including river channels, bank storage and canopy storage. Due to the mass correction in GRACE data processing, the water storage variations in lakes/reservoirs were not taken into account.

2. Storages ($V_2 + V_3$) for lagged surface runoff and lateral flow. The two storages are defined in SWAT for estimating the amount of overland and lateral flow reaching river channels on a daily time step. SWAT allows for delayed release of overland flow and lateral flow yielded in river basins with time of concentration greater than one day.

3. Soil profile (V_4).

4. Vadose zone (V_5). Water storage in the vadose zone is typically not considered as a storage in SWAT water balance analysis because the Venetis' exponential decay weighting function (1969) does not alter the quantity of water from soil into aquifers. However, the time delay for water to move through the vadose zone results in variations in water storage and needs to be addressed in TWS variation calculations (Milzow et al., 2011).

5. Groundwater (V_6). SWAT simulates an unconfined shallow aquifer and a confined deep aquifer in each subbasin. Water storage in shallow aquifers may contribute to flow in the main river channels or be re-evaporated into the soil. By contrast, there is no simulated outlet for water in deep aquifers except pumpage. "Water that enters the deep aquifer is assumed to contribute to streamflow somewhere outside of the watershed" (Neitsch et al., 2005). While this assumption may hold in studies for small river basins, it is no longer valid at a continental scale. Due to the accumulation of water percolated from shallow aquifer, an upward trend in water storage in deep aquifers would be found, which is unrealistic. To circumvent this problem, the deep aquifer was excluded from the simulations and the calculation of TWS by setting the percolation rate to the deep aquifer to zero.

For each subbasin, the model-based TWS for each 10-day period was calculated as:

$$\text{TWS}_t = V_{1,t} + V_{2,t} + V_{3,t} + V_{4,t} + V_{5,t} + V_{6,t}$$

where t is the index for the 10-day period. The series of SWAT subbasin-wide TWS anomalies (TWSV_t) was computed by differencing the TWS for each 10-day period TWS_t and the mean of the TWS over the entire GRACE data period:

$$\text{TWSV}_t = \text{TWS}_t - \overline{\text{TWS}}$$

where $\overline{\text{TWS}}$ is the mean of the TWS_t over the GRACE data period, or was calculated by taking the average of 10-d TWS_t's during July 2002–April 2009.

5 Calibration approach

Calibration and evaluation of the SWAT-SSA model in this study was carried out using a multi-criteria framework, similar to the studies by Werth and Güntner (2010). The multi-criteria approach extends the traditional calibration approach by casting the calibration into a multi-objective optimisation problem, and for independent data, allows evaluation of model performance against more than one objective to improve model robustness and predictability (Gupta et al., 1998). The solution to the multi-criteria optimisation programme consists of the non-dominated calibration parameter sets, which are optimal in a Pareto efficiency sense. The trade-off between model fits evaluated by different criteria
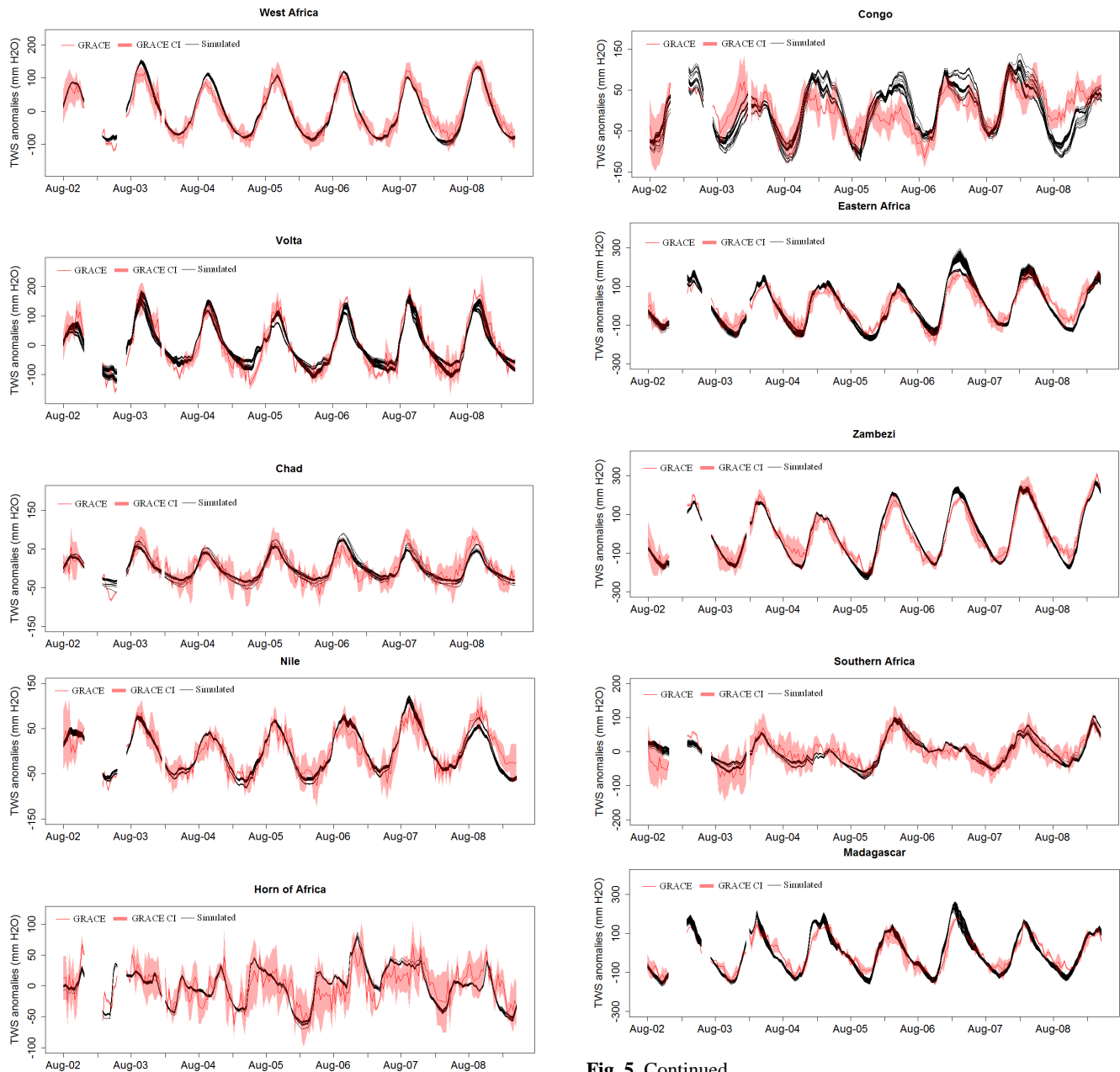
Fig. 5. Observed and zonally averaged simulated TWS variations for ten sub-continental regions (TWS-total water storage; CI-confidence).

Fig. 5. Continued.

reflects the minimum parameter uncertainty (Vrugt et al., 2003) caused by errors in the input and measured data as well as by the model structure.

For the calibration of the SWAT-SSA model, two objective functions were defined. Their definitions and calculations are explained in detail below. The multi-objective optimisation problems defined in the multi-criteria calibration of the SWAT-SSA models were solved using the Non-dominated Sorting Genetic Algorithm II (NSGA-II, Deb et al., 2002),

a population-based heuristic evolutionary optimisation technique with a proven track record of success in solving many large-scale optimisation problems. The population sizes chosen in the optimisations varied from 150 to 300, and the optimisations lasted for $50 \sim 100$ generations until no significant improvements in the solution were observed.

5.1 Comparison of model-based and GRACE-derived TWS variations

As GRACE provides a filtered image of reality, the modelled storage variations from SWAT were first converted to GRACE resolution to provide storage values at the same
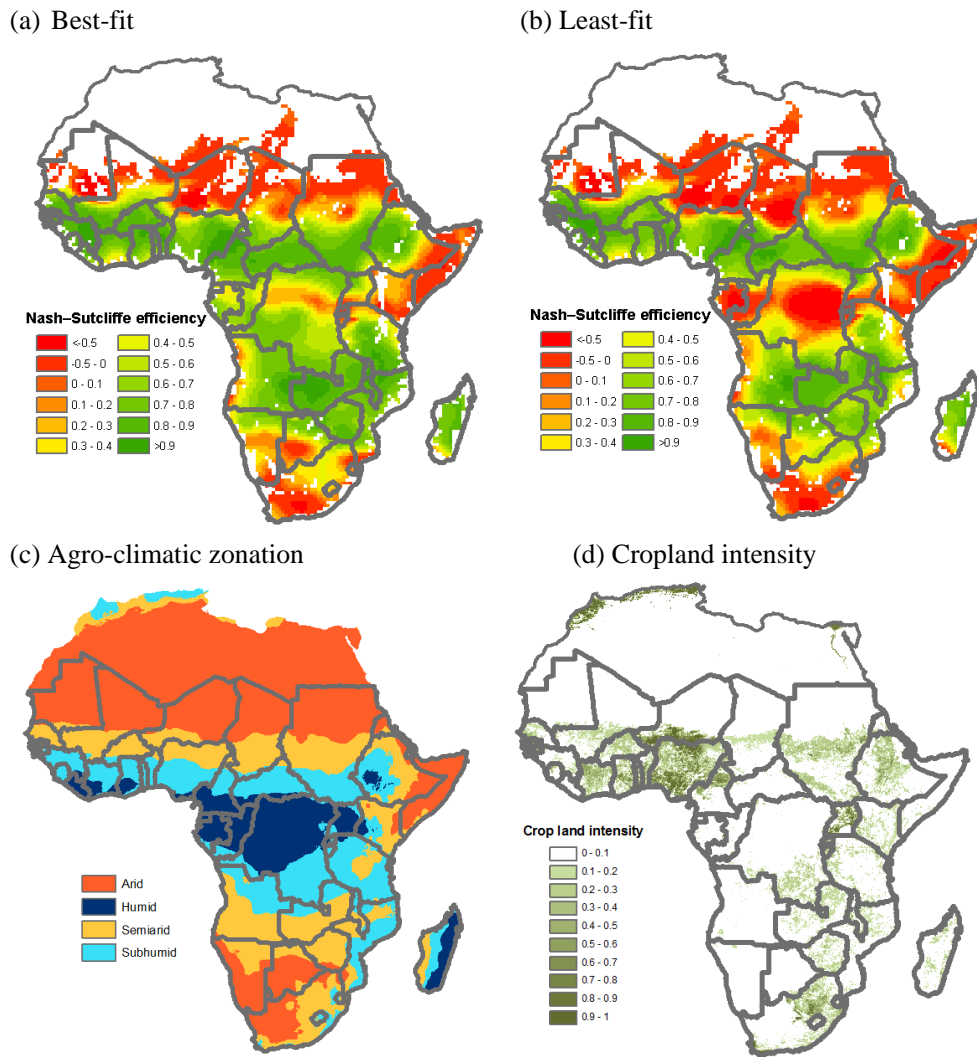
(a) Best-fit

(b) Least-fit

(c) Agro-climatic zonation

(d) Cropland intensity

Fig. 6. Spatial patterns of SWAT model fits in simulations of total water storage variations in Sub-Saharan Africa (country boundaries are shown).

spatial scales for comparison. This mathematical process involves projecting SWAT modelled spatial fields to Spherical Harmonics (SH) up to degree 50 (in this study, SH transformation was conducted using SHTOOLS, http://www.ipgp.fr/~wieczor/SHTOOLS/SHTOOLS.html) in which the SWAT-based basin-wide TWS variations for each 10-day period were first disaggregated into a 0.5 by 0.5 degree grid prior to the transformation. In order to allow for a comparison between GRACE- and SWAT-based TWS variations for sub-continental regions, simulated variations in TWS by the Noah land surface model (Ek et al., 2003) in NASA's Global Land Data Assimilation System (GLDAS) (Rodell et al., 2004) were used as a priori information to fill areas outside of the SSA sub-region of interest in the SH transformation.

Agreement between GRACE-derived and model-based TWS variations was evaluated using a weighted total square error (WTSE) function:

$$\text{WTSE} = \sum_{t=1}^{T} \sum_{i=1}^{I} \sum_{j}^{J} I_s \times w_{i,j,t}$$
$$\times (\text{TWSV}_{i,j,t,\text{SWAT}} - \text{TWSV}_{i,j,t,\text{GRACE}})^2 \quad (1)$$

where $\text{TWSV}_{i,j,t,\text{SWAT}}$ and $\text{TWSV}_{i,j,t,\text{GRACE}}$ are SWAT- and GRACE-based TWS variations for 10-day period t and grid cell (i, j), respectively. I_s is an indicator function. $I_s = 1$ if the grid cell is located within the study region; otherwise, $I_s = 0$, $w_{i,j,t}$ is the weight, an inverse of standard error of GRACE-based TWS variations $\text{TWSV}_{i,j,t,\text{GRACE}}$.

Finally, following the convention in hydrologic model calibration, available GRACE data were divided into two groups: the first 112 10-day periods (29 July 2002– December 2005) were used for calibration and the data for remaining 120 10-day periods were reserved for validation.

Table 6. Temporal variability of zonally averaged component-wise water storage variations (%).

		West Africa	Volta	Chad	Nile	Horn of Africa	Congo	Eastern Africa	Zambezi	South Africa	Madagascar
Soil	Mean	55.2	11.5	31.5	37.9	44.8	10.0	28.8	18.9	20.8	9.1
	Max	61.1	16.2	58.5	48.9	48.8	31.0	38.9	21.4	38.5	13.8
	Min	50.3	8.5	20.8	31.4	37.0	5.3	17.1	17.4	8.6	6.8
Vadose	Mean	5.2	46.7	26.9	4.5	9.8	37.1	19.4	16.8	13.9	42.6
	Max	9.5	55.0	34.2	11.6	13.1	51.5	31.0	19.8	40.8	59.2
	Min	2.7	8.1	3.0	0.0	6.0	0.24	0.0	12.5	3.1	21.6
Groundwater	Mean	1.6	15.3	6.2	13.0	11.3	8.4	5.1	8.4	25.3	2.9
	Max	3.6	56.4	24.8	24.5	16.2	59.1	25.5	16.2	41.5	8.6
	Min	0.42	5.3	2.0	6.7	8.1	0.01	2.7	6.6	11.5	0.08
Overland water	Mean	0.03	0.04	0.01	0.4	0.007	0.10	0.004	0.010	0.03	0.0004
	Max	0.05	0.10	0.02	0.58	0.009	0.30	0.010	0.020	0.06	0.0010
	Min	0.02	0.01	0.01	0.25	0.006	0.01	0.002	0.005	0.01	0.0002
Surface water lag	Mean	0.003	0.20	0.07	0.03	0.01	0.06	0.01	0.22	0.31	0.02
	Max	0.005	1.23	0.52	0.14	0.01	0.32	0.02	0.59	0.83	0.03
	Min	0.002	0.001	0.003	0.005	0.01	0.03	0.002	0.04	0.02	0.01
Lateral flow lag	Mean	0.0007	0.0001	0.00006	0.004	0.0014	0.010	0.005	0.00042	0.0009	0.011
	Max	0.0010	0.0002	0.0001	0.04	0.0015	0.020	0.030	0.0010	0.0015	0.030
	Min	0.0004	0.00006	0.00004	0.001	0.0013	0.001	0.001	0.00036	0.0005	0.006

5.2 Criterion/objective function for evaluating goodness of model fit in runoff field simulation

Observed monthly river discharge data from 187 discharge stations in SSA (Fig. 2) were used for this calibration study. The data were obtained from Global Runoff Data Centre (GRDC), a primary source of information for global river discharge to support large-scale hydrologic studies (date of data retrieval: 30 September 2009). The starting and ending dates of the discharge series for these stations vary by station. The earliest discharge data date back to 1900 and were obtained from the station near Khartoum on the Blue Nile River. The most recent data are from 2001 and several stations on the Orange River, the Great Fish River and the Limpopo River in South Africa. For the majority of stations in SSA, river discharge data are available up to 1980s and early 1990s. The different time frames among the GRDC river discharge data, GRACE data (2002–2009), and GPCP 1-DD precipitation data (1997–2009) pose difficulty for model calibration. In this study, we focused on evaluating the performance of SWAT for modelling TWS variability: SWAT was run for 2002–2009 (with five additional years 1997–2001 as the spin-up period) and, following the approach by Werth and Güntner (2010), simulated and observed monthly river discharge rates in two time frames were compared on a multi-year average basis. The fit of the SWAT model at each GRDC station was measured using the Nash-Sutcliffe Efficiency (NSE) coefficient (Nash and Sutcliffe, 1970), which

is defined as

$$\text{NSE} = 1 - \frac{\sum_{t=1}^{12} (Q_{t,\text{obs}} - Q_{t,\text{sim}})^2}{\sum_{t=1}^{12} (Q_{t,\text{obs}} - \bar{Q}_{t,\text{obs}})^2} \tag{2}$$

where $Q_{t,\text{obs}}$ and $Q_{t,\text{sim}}$ are the multi-year averaged monthly discharges calculated using the simulated and available observed discharges ($\text{m}^3\,\text{s}^{-1}$), respectively. $\bar{Q}_{t,\text{obs}}$ is the mean of $Q_{t,\text{obs}}$ ($\text{m}^3\,\text{s}^{-1}$). The NSE coefficient can range from $-\infty$ to 1, where 1 indicates a perfect model fit.

The GRDC station network is relatively dense in West Africa, but limited in other regions (Fig. 2). This highlights the benefit of applying GRACE data to support hydrologic simulations in SSA. The NSE values for all GRDC stations in a sub-regional model were weighted by length of observed monthly river discharge series:

$$\text{WNSE} = \sum_i w_i\, \text{NSE}_i \tag{3}$$

where NSE_i is NSE coefficient at GRDC station i, and w_i is the weighting factor proportional to the length of the monthly river discharge data time series at that station ($\sum_i w_i = 1$).

The weighted NSE (WNSE) serves as the criterion for evaluating performance of SWAT in simulating river discharge.
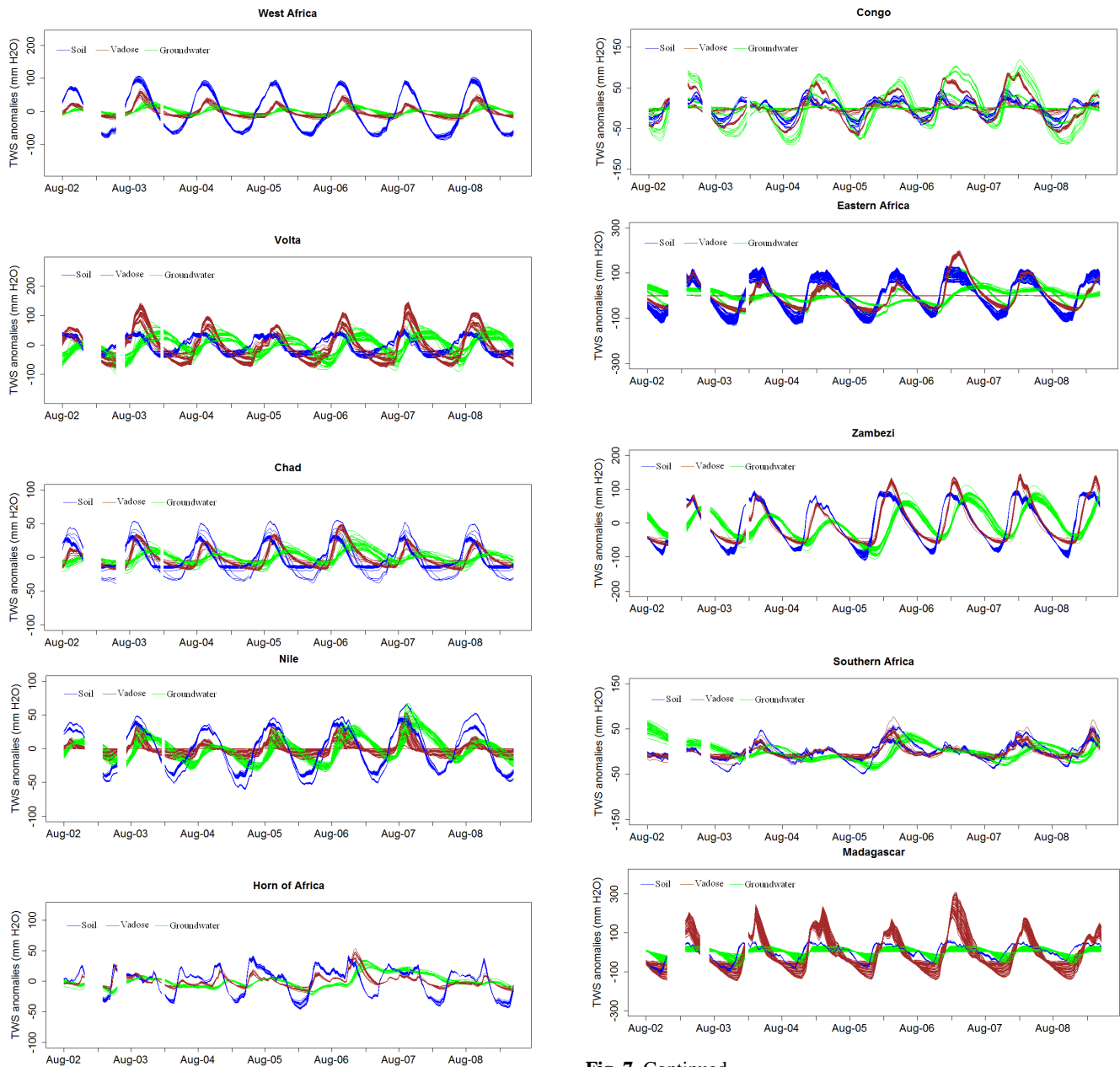
Fig. 7. Zonally averaged water variations in soil, vadose zone and groundwater storages (TWS-total water storage).

Fig. 7. Continued.

5.3 Calibration parameters

The hydrologic processes and watershed properties in SWAT are characterised by many parameters. A list of SWAT parameters selected for calibration, together with their lower and upper bounds of adjustable ranges, are shown in Table 4. This list was determined from literature review, numerical sensitivity analysis (Morris, 1991), and according to results from several test runs of the calibration programmes.

In these SWAT calibration parameters, SCS curve number is a key parameter for surface runoff estimation. It is defined to characterise the potential maximum soil moisture retention capacity. A low value indicates low runoff, but high infiltration potential. Surface runoff lag coefficient (SURLAG) determines how much total available runoff enters into a river reach on a given day and is a sensitivity parameter for simulating river discharge hydrographs. Soil evaporation compensation factor (ESCO) is defined to specify the depth distribution used to meet soil evaporative demand. As the value of ESCO decreases, more water can be evaporated from deeper soil layers. SOL_AWC, SOL_K, and
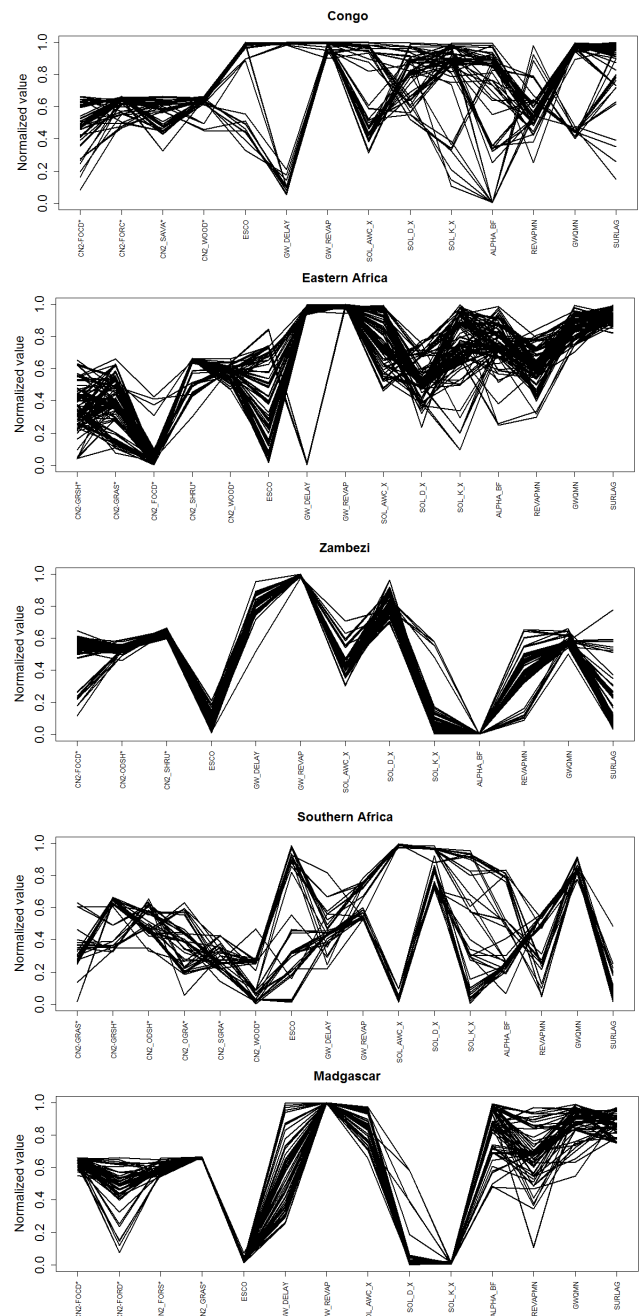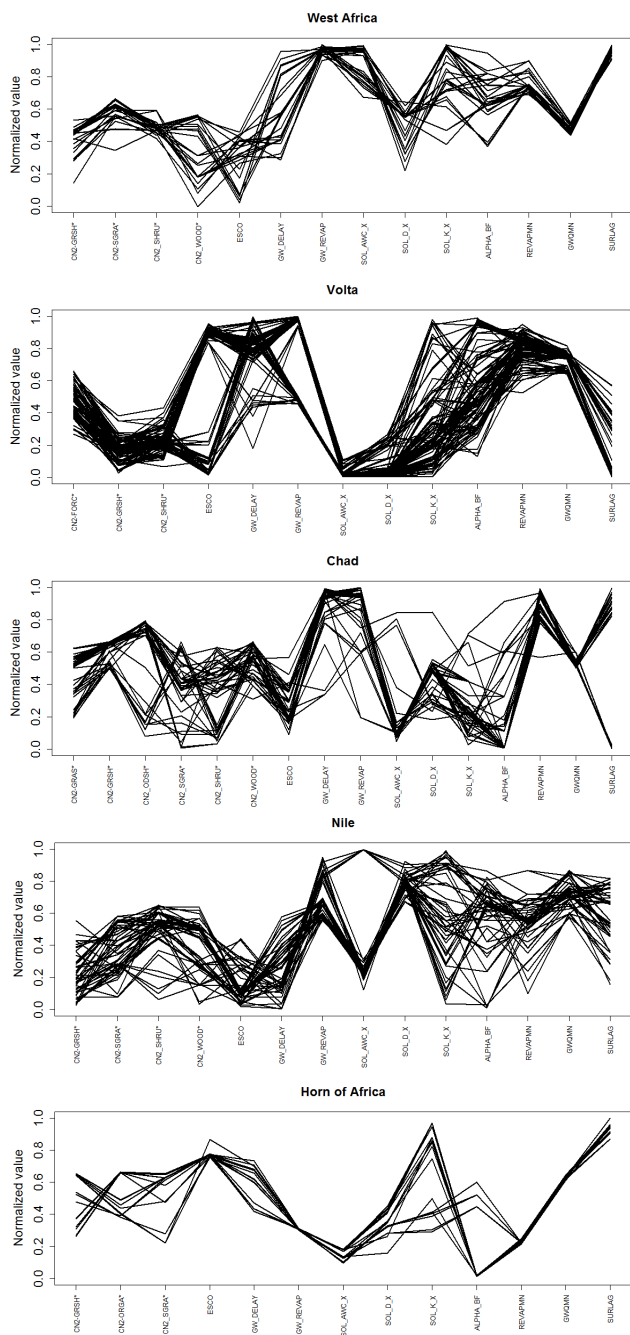
West Africa

Volta

Chad

Nile

Horn of Africa

Congo

Eastern Africa

Zambezi

Southern Africa

Madagascar

Fig. 8. Continued.

Fig. 8. Estimates of SWAT calibration parameters obtained from mutli-criteria calibration (abbreviations of the land cover type: FOCD – Closed deciduous forest; FORC – Closed evergreen lowland forest; FORD – Degraded evergreen lowland forest; FORS – Submontane forest; GRAS – Closed grassland; GRSH – Open grassland with sparse shrubs; ODSH – Open deciduous shrubland; OGRA – Open grassland; SAVA – Mosaic Forest/Savanna; SGRA – Sparse grassland; SHRU – Deciduous shrubland with sparse trees; WOOD – Deciduous woodland).

SOL_D are soil available water capacity, saturated conductivity and the soil layer depth, respectively. All three soil attributes are highly uncertain. Values of the first two parameters were derived using pedotransfer functions; no reliable information about the actual depth of the soil layer in Africa is available from HWSD, only a reference value (in most cases 1 m) was assigned. GW_DELAY (groundwater delay coefficient) characterising the delay time for recharge into the aquifer, is a single controlling parameter for determining

the water storage variation in the vadose zone. The remaining four parameters in the table, GW_REVAP (groundwater revap coefficient), ALPHA_BF (baseflow alpha factor), REVAPMN (threshold depth of water in the shallow aquifer for "revap" to occur) and GWQMN (threshold depth of water in the shallow aquifer required for groundwater flow to occur) control the behaviour of shallow aquifers.

The ten SWAT sub-regional models were calibrated with the parameters shown in Table 4. The model for each sub-region was calibrated independently, but within a region, the same percentage changes were made for those parameters which have spatially varying values (or parameters other than SURLAG and ESCO), except for the SCS curve number (CN2), based on spatial fields of their initial estimates. The CN2 values are correlated with land cover and soil permeability. In this study, within a sub-region the CN2 parameters were grouped by land cover, and the CN2 values for major land covers/uses were considered as independent calibration parameters.

6 Results

The Pareto fronts in the two-dimensional objective space found via multi-criteria calibration for all ten SSA sub-regions are shown in Fig. 3. Paired values of weighted root-mean-square-error (RMSE) and weighted NSE coefficients are plotted on horizontal and vertical axes, respectively (weighted RMSE is a monotonic function of weighted TSE). A model with a perfect fit to GRACE data and river discharge data would have a weighted RMSE of 0 and a weighted NSE coefficient of 1. Thus, the Pareto front curves are convex towards the point (0, 1), reflecting tradeoffs between the ability to fully describe discharge or TWS variations.

With regard to performance of calibrated models in river discharge simulation, the highest values of weighted NSE coefficients obtained vary from -2.55 to 0.66 and are negative for five out of the ten sub-region models (West Africa, Nile, Congo, Zambezi and Madagascar). This measure of goodness of fit statistic is also sensitive to different solutions of parameter sets in Pareto fronts. The deterioration of its value is greater than two in models for all sub-regions other than West Africa, Nile and Zambezi when the parameter set in the Pareto frontier that most closely matches the simulation of GRACE TWS variations was used. The NSE model coefficients for each individual GRDC gauging station are shown in Fig. 4. When the "best-fit" solutions for river discharge simulation were taken, 20 % of the stations have NSEs ≥ 0.7, 43 % ≥ 0.4 and the NSEs for 64 % of the GRDC stations are positive. These percentages decrease from 20 to 6 %, 43 to 17 % and 64 to 30 % if the models are run with the "least-fit" solutions for river discharge simulation.

More satisfactory model fits were achieved in simulation of TWS variations after the calibration of the SWAT model. The ensembles of time series of zonally averaged simulated TWS variations over the 10 sub-regions and associated with Pareto optimal solutions found in multi-criteria calibration are plotted in Fig. 5, together with the time series of zonally averaged GRACE- based mean TWS variations and the related one-sigma (68.7 %) confidence interval (CI). The NSE coefficients for the time series of model-based TWS variations with respect to the GRACE- based mean TWS variations were also calculated and summarised, for calibration and validation periods, respectively (Table 5). Overall, simulated and GRACE-based zonally averaged time series are in good agreement in sub-regions of West Africa, Volta, Chad, Nile, Eastern African, Zambezi and Madagascar during both calibration and validation periods. The means of the NSE coefficients for these sub-regional models range from 0.66 to 0.91. Larger discrepancies were found in the Congo and Horn of Africa. In the simulation of temporal variations of TWS for these two sub-regions, the model still captures the general trends/phase changes of TWS variations, but the mismatch in amplitude is greater. For the Southern African model, the model fit was poorer during the calibration period; however, the model performs much better during the validation period.

The NSE coefficients calculated on a gridded basis are shown in Fig. 6a and b. The "best-fit" and "least-fit" solutions were determined according to model fits with respect to GRACE TWS variations. Figure 6c and d show the agroclimatic zonation (derived from FAO Agroecological Zones crafted by HarvestChoice, Z. Guo, personal communication, 2011), and the density of cropland (fraction of cropland area in 5 min × 5 min grid, Ramankutty et al., 2008) in SSA. Generally, the model performs well in simulating TWS variations in semiarid and sub-humid areas, which encompass most cropland in SSA. The largest discrepancies in simulated TWS variations (NSE coefficients ≤ 0) occurred in arid areas, where water storage amplitude is lower or equivalent to GRACE error (the Sahara, Somalia, western Ethiopia, northwest Kenya, south Namibia and most of Southern Africa) and the equatorial humid area (notably in central Democratic Republic of the Congo).

GRACE TWS data integrate water mass variations from all storage components. Sometimes, interest is focused on estimating water mass variations in certain storage components (e.g., groundwater; Rodell et al., 2009; Tiwari et al., 2009). Temporal variability of zonally aggregated water mass in six water storages parameterised in the SWAT model are characterised by calculating ratios between variances of these storage variables $\sigma^2_{V_i}$ ($i = 1, \cdots, 6$) and variance of model-based total water storage variation $\sigma^2_{V_{\text{Total}}}$ (in unfiltered space), or $\sigma^2_{V_i} / \sigma^2_{V_{\text{Total}}}$. Means and ranges of calculated normalised variances for each storage variable and each sub-region are listed in Table 6 (note that the water mass variations in six storages are not independent; thus, $\sigma^2_{V_{\text{Total}}}$ may not equal $\sum\limits_{i=1}^{6} \sigma^2_{V_i}$). These statistics show that the three water storage components

that have largest temporal variability, thus, contributing most to TWS, are soil, vadose zone and groundwater storage. By contrast, contributions from overland flow, surface runoff and lateral flow lags are trivial. Zonally aggregated time series of soil water, vadose zone water and groundwater storage variations obtained from the calibrated SWAT models are shown in Fig. 7. Systematic phase differences exist among the time series for the three storage variables: in each annual cycle when the rainy season begins the soil moisture is first replenished and peaks, followed by vadose zone water and then groundwater.

The statistics in Table 6 and graphs in Fig. 7 also indicate that there could be even larger uncertainties in estimation of component-wise water storages than what is seen in TWS estimation. For example, the model gives divergent estimates for water storage variations in vadose zone and groundwater in Eastern Africa when the model was run with parameter sets across the Pareto frontier. The estimated time series for water storage variations in the vadose zone and groundwater fall into two groups: one group has large variations in the vadose zone water storage, but relatively smaller variations in groundwater storage; in another group, vadose zone water storage variations are almost zero and variations in groundwater storage are much larger. Figure 8 shows the Pareto fronts in parameter space with normalised parameter values in [0, 1] intervals (zero values represent the lower bounds of the adjustable ranges of the parameters and one corresponds to the upper bounds). The disparate estimates for vadose zone and groundwater storage variations can be explained by the dichotomy in the estimated values of GW_DELAY. As the value of GW_DELAY approaches zero, water exiting the bottom of the soil profile can enter aquifers immediately causing no variation in vadose water storage and larger variations in groundwater storage, with the opposite for large values for GW_DELAY. Similar divergent estimates in vadose zone water and groundwater storage variations caused by different GW_DELAY estimates were also found in the Congo model. For Chad, Nile and Southern Africa, there are large uncertainties in estimating soil water storages, which is most likely related to divergent estimates of SOL_AWC_X.

7   Discussion and conclusions

The study presented in this paper concerns calibration/evaluation of a semi-distributed model based on SWAT code (SWAT-SSA) for regional-scale hydrologic simulation in Sub-Saharan African countries. The SWAT-SSA models were calibrated and evaluated in a multi-criteria framework to both river discharge and GRACE TWS data, but with more focus on assessing the model's capacity for simulating TWS variability using GRACE data. In spite of uncertainty arising from the tradeoff in optimising model parameters with respect to two model fitting criteria and in estimation of storage variations contributed by different storage components,

the study showed that the calibrated SWAT-SSA model performs well in simulating TWS variations in semi-arid and semi-humid areas, where agriculture in SSA is concentrated and, therefore, is capable of acting as an effective modelling tool for agricultural water management in SSA.

Any model calibration and validation exercise is subject to certain limitations. A major limitation in this study originated from use of multi-year average monthly river discharge data for a time frame (1900–2001) different from that in which the models were actually run (2002–2009) as a result of limited availability and accessibility to recent stream flow data in Sub-Saharan countries. Therefore, it is difficult to evaluate the model's adequacy in simulating the surface water system. Climate and land use change may alter the flow regimes of rivers (e.g., Amogu et al., 2010) and potentially bias the calibration of river discharge and estimation of the contribution of water mass variations in river systems to TWS variations. Furthermore, an interesting question often raised in model calibration/evaluation is what value is brought to model calibration and evaluation by use of additional dataset(s). The limitation with river discharge data makes it difficult to answer this question. However, the outcome of the SWAT model calibration and validation in this study tends to suggest that use of GRACE data may only provide limited additional constraints to reduce parametric uncertainties of the model because the NSEs shown in Fig. 3 and Table 5 indicate that there might be stronger equifinality (Beven and Binley, 1992) in the TWS simulation than in discharge simulation. On the other hand, it is apparent that GRACE data are of great value in that they provide valuable information and unprecedented opportunity to validate and evaluate the model's capacity to simulate spatio-temporal variability in TWS.

Finally, there is less agreement between model- and GRACE-based TWS variations in arid and equatorial humid areas. A few possible reasons for the larger discrepancies include: firstly, uncertainties associated with GRACE TWS variations themselves. Noticeable disagreement between model- and GRACE-based TWS variations in arid regions were also reported by Ngo-Duc et al. (2007) in a global-scale model validation study, who attributed the disagreements in arid areas to weak GRACE signals and resulting larger errors in GRACE TWS variations. Secondly, errors in climatic forcing data, especially in precipitation data may be important. Due to lack of ground-based observations in precipitation, regional or global scale hydrologic simulations typically rely on use of precipitation estimates from different satellite-based or meteorological reanalysis data products. Uncertainty associated with these precipitation datasets is often a principle source of uncertainty for hydrologic simulation (e.g., Fiedler and Döll, 2007). Thirdly, inadequacy of SWAT parameterisation or algorithms in simulating hydrology in arid and humid areas may also help explain the discrepancy. Future work will be required to identify physical reasons for model misfits and for model enhancement.

References

Alkama, R., Decharme, B., Douville, H., Becker, M., Cazenave, A., Sheffield, J., Voldoire, A., Tyteca, S., and Le Moigne, P.: Global Evaluation of the ISBA-TRIP Continental Hydrological System. Part I: Comparison to GRACE Terrestrial Water Storage Estimates and In Situ River Discharges, J. Hydrometeorol., 11, 583–600, 2010.

Amogu, O., Descroix, L., Yéro, K. S., LeBreton, E., Mamadou, I., Ali, A., Vischel, T., Bader, J. C., Moussa, I. B., Gautier, E., Boubkraoui, S., and Belleudy, P.: Increasing River Flows in the Sahel?, Water, 2, 170–199, 2010.

Arnold, J. G., Srinivasin, R., Muttiah, R. S., and Williams, J. R.: Large Area Hydrologic Modeling and Assessment: Part I. Model Development, Journal of American Water Resources Association, 34, 73–89, 1998.

Becker, M., Llovel, W., Cazenave, A., Güntner, A., and Crétaux, J. F.: Recent hydrological behaviour of the East African Great Lakes region inferred from GRACE, satellite altimetry and rainfall observations, CR Geoscience, 342, 223–233, 2010.

Bettadpur, S.: Level-2 Gravity Field Product User Handbook, GRACE 327–734, The GRACE Project, Center for Space Research, University of Texas at Austin, 2007.

Beven, K. J. and Binley, A. M.: The future of distributed models: model calibration and uncertainty prediction, Hydrol. Process., 6, 279–298, 1992.

Bruinsma, S., Lemoine, J. M., Biancale, R., and Vales, N.: CNES/GRGS 10-day gravity field models (release 2) and their evaluation, Adv. Space Res., 45, 587–601, 2010.

Chen, J. L., Wilson, C. R., Tapley, B. D., Longuevergne, L., Yang, Z. L., and Scanlon, B. R.: Recent La Plata basin drought conditions observed by satellite gravimetry, J. Geophys. Res., 115, D22108, doi:10.1029/2010JD014689, 2010.

Crétaux, J.-F., Jelinski, W., Calmant, S., Kouraev, A., Vuglinski, V., Bergé-Nguyen, M., Gennero, M.-C., Nino, F., Abarca Del Rio, R., Cazenave, A., and Maisongrande, P.: SOLS: A lake database to monitor in the Near Real Time water level and storage variations from remote sensing data, Adv. Space Res., 47, 1497–1507, 2011.

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II, IEEE T. Evolut. Comput., 6, 182–197, 2002.

Dhar, S. and Mazumdar, A.: Hydrological modelling of the Kangsabati river under changed climate scenario: case study in India, Hydrol. Process., 23, 2394–2406, 2009.

Ek, M. B., Mitchell, K. E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G., and Tarpley, J. D.: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, J. Geophys. Res., 108, 8851, doi:10.1029/2002JD003296, 2003.

FAO/IIASA/ISRIC/ISSCAS/JRC: Harmonized World Soil Database (version 1.1), FAO, Rome, Italy and IIASA, Laxenburg, Austria, 2009.

FAO: The State of Food Insecurity in the World 2010: Addressing Food Insecurity in Protracted Crises, Rome, 2010.

FAO (Food and Agriculture Organization of the United Nations): Water Resources Development and Management Service, AQUASTAT database, available at: http://www.fao.org/nr/water/aquastat/dbases/index.stm, last access: July, 2011.

Falkenmark, M. and Rockstrom, J.: The New Blue and Green Water Paradigm: Breaking New Ground for Water Resources Planning and Management, J. Water Res. Pl. ASCE, 132, 129–132, 2006.

Famiglietti, J. S., Lo, M., Ho, S. L., Bethune, J., Anderson, K. J., Syed, T. H., Swenson, S. C., de Linage, C. R., and Rodell, M.: Satellites measure recent rates of groundwater depletion in California's Central Valley, Geophys. Res. Lett., 38, L03403, doi:10.1029/2010GL046442, 2011.

Fenicia, F., McDonnell, J. J., and Savenije, H. H. G.: Learning from model improvement: On the contribution of complementary data to process understanding, Water Resour. Res., 44, W06419, doi:10.1029/2007WR006386, 2008.

Fiedler, K. and Döll, P.: Global modelling of continental water storage changes – sensitivity to different climate datasets, Adv. Geosci., 11, 63–68, doi:10.5194/adgeo-11-63-2007, 2007.

Grippa, M., Kergoat, L., Frappart, F., Araud, Q., Boone, A., de Rosnay, P., Lemoine, J.-M., Gascoin, S., Balsamo, G., Ottlé, C., Decharme, B., Saux-Picart, S., and Ramillien, G.: Land water storage variability over West Africa estimated by Gravity Recovery and Climate Experiment (GRACE) and land surface models, Water Resour. Res., 47, W05549, doi:10.1029/2009WR008856, 2011.

Guntner, A.: Improvement of global hydrological models using GRACE data, Surv. Geophys., 29, 375–397, 2008.

Gupta, H. V., Sorooshian, S., and Yapo, P. O.: Towards Improved Calibration of Hydrologic Models: Multiple and Non-Commensurable Measures of Information, Water Resour. Res., 34, 751–763, doi:10.1029/97WR03495, 1998.

Hargreaves, G. H. and Samani, Z. A.: Reference crop evapotranspiration from temperature, Appl. Eng. Agric., 1, 96–99, 1985.

Huffman, G. J., Adler, R. F., Morrissey, M., Bolvin, D. T., Curtis, S., Joyce, R., McGavock, B., and Susskind, J.: Global Precipitation at One-Degree Daily Resolution from Multi-Satellite Observations, J. Hydrometeor., 2, 36–50, 2001.

IPCC: Climate Change 2007: Impacts, Adaptation, and Vulnerability, IPCC Working Group II, Fourth Assessment Report, edited by: Parry, M. L., Canziani, O. F., Palutikof, J. P., van der Linden, P. J., and Hanson, C. E., Cambridge University Press, 2007.

Kim, N. W., Chung, I. M., Won, Y. S., and Arnold, J. G.: Development and application of the integrated SWAT-MODFLOW MODEL, J. Hydrol., 356, 1–16, 2008.

Konz, M. and Seibert, J.: On the value of glacier mass balances for hydrological model calibration, J. Hydrol., 386, 238–246, 2010.

Leblanc, M. J., Tregoning, P., Ramillien, G., Tweed, S. O., and Fakes, A.: Basin-scale, integrated observations of the early 21st century multiyear drought in southeast Australia, Water Resour. Res., 45, W04408, doi:10.1029/2008WR007333, 2009.

Lehner, B. and Döll, P.: Development and validation of a global database of lakes, reservoirs and wetlands, J. Hydrol., 296, 1–22, 2004.

Lehner, B., Verdin, K., and Jarvis, A.: New global hydrography derived from spaceborne elevation data, Eos, Transactions, AGU, 89, 93–94, 2008.

Longuevergne, L., Scanlon, B. R., and Wilson, C. R.: GRACE Hydrological estimates for small basins: Evaluating processing approaches on the High Plains Aquifer, USA, Water Resour. Res., 46, W11517, doi:10.1029/2009WR008564, 2010.

Milzow, C., Krogh, P. E., and Bauer-Gottwein, P.: Combining satellite radar altimetry, SAR surface soil moisture and GRACE total storage changes for hydrological model calibration in a large poorly gauged catchment, Hydrol. Earth Syst. Sci., 15, 1729–1743, doi:10.5194/hess-15-1729-2011, 2011.

Molden, D. (Ed.): Water for food, water for life. A comprehensive assessment of water management in agriculture, London, UK: Earthscan; Columbo, Sri Lanka: International Water Management Institute, 2007.

Monteith, J. L.: Evaporation and the environment, 205–234, in: The state and movement of water in living organisms, 19th Symposia of the Society for Experimental Biology, Cambridge Univ. Press, London, UK, 1965.

Morris, M. D.: Factorial sampling plans for preliminary computational experiments, Technometrics, 33, 161–174, 1991.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I – A discussion of principles, J. Hydrol., 10, 282–290, 1970.

Neitsch, S. L., Arnold, J. G., Kiniry, J. R., and Williams, J. R.: Soil and Water Assessment Tool Theoretical Documentation, Version 2005, Grassland, Soil and Water Research Laboratory, USDA-ARS, 2005.

Ngo-Duc, T., Laval, K., Ramillien, G., Polcher, J., and Cazenave, A.: Validation of the land water storage simulated by Organising Carbon and Hydrology in Dynamic Ecosystems (ORCHIDEE) with Gravity Recovery and Climate Experiment (GRACE) data, Water Resour. Res., 43, W04427, doi:10.1029/2006WR004941, 2007.

Niu, G.-Y. and Yang, Z.-L.: Assessing a land surface model's improvements with GRACE estimates, Geophys. Res. Lett., 33, L07401, doi:10.1029/2005GL025555, 2006.

Oeurng, C., Sauvage, S., and Sánchez-Pérez, J.-M.: Assessment of hydrology, sediment and particulate organic carbon yield in a large agricultural catchment using the SWAT model, J. Hydrol., 401, 145–153, 2011.

Parajka, J., Naeimi, V., Blöschl, G., Wagner, W., Merz, R., and Scipal, K.: Assimilating scatterometer soil moisture data into conceptual hydrologic models at the regional scale, Hydrol. Earth Syst. Sci., 10, 353–368, doi:10.5194/hess-10-353-2006, 2006.

Priestley, C. H. B. and Taylor, R. J.: On the assessment of surface heat flux and evaporation using large-scale parameters, Mon. Weather Rev., 100, 81–82, 1972.

Ramankutty, N., Evan, A. T., Monfreda, C., and Foley, J. A.: Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000, Global Biogeochem. Cy., 22, GB1003, doi:10.1029/2007GB002952, 2008.

Rockström, J., Hatibu, N., Oweis, T., Wani, S., Barron, J., Bruggeman, A., Qiang, Z., Farahani, J., and Karlberg, L.: Managing Water in Rainfed Agriculture, in: Water for Food, Water for Life:

A Comprehensive Assessment of Water Management in Agriculture, edited by: Molden, D., Chapter 9, 315–348, Earthscan, London, 2007.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., and Toll, D.: The global land data assimilation system, B. Am. Meteorol. Soc., 85, 381–394, 2004.

Rodell, M., Velicogna, I., and Famiglietti, J. S.: Satellite estimates of groundwater depletion in India, Nature, 460, 999–1002, 2009.

Rosegrant M. W., Cai, X., and Cline, S. A.: World Water and Food to 2025: Dealing With Scarcity. Washington, International Food Policy Research Institute, Washington D.C., 2002.

Rosegrant, M. W., Fernandez, M., Sinha, A., Alder, J., Ahammad, H., de Fraiture, C., Eickhout, B., Fonseca, J., Huang, J., Koyama, O., Omezzine, A. M., Pingali, P., Ramirez, R., Ringler, C., Robinson, S., Thornton, P., van Vuuren, D., Yana-Shapiro, H., Ebi, K., Kruska, R., Munjal, P., Narrod, C., Ray, S., Sulser, T., Tamagno, C., van Oorschot, M., and Zhu, T.: Looking into the future for agriculture and AKST (Agricultural Knowledge Science and Technology), Chapter 5, in: Agriculture at a Crossroads, edited by: McIntyre, B. D., Herren, H. R., Wakhungu, J., and Watson, R. T., Island Press, Washington DC, 307–376, 2009.

Sangrey, D. A., Harrop-Williams, K. O., and Klaiber, J. A.: Predicting groundwater response to precipitation. J. Geotech. Eng., ASCE, 110, 957–975, 1984.

Saxton, K. E., Rawls, W. L., Rosenberger, J. S., and Papendick, R. I.: Estimating generalized soil-water characteristics from texture, Soil Sci. Soc. Am. J., 50, 1031–1036, 1986.

Scanlon, B. R., Longuevergne, L., and Long, D.: Ground referencing GRACE satellite estimates of groundwater storage changes in the California Central Valley, US, Water Resour. Res., 48, W04520, doi:10.1029/2011WR011121, 2012.

Schaap, M. G., Leij, J. F., and van Genuchten, M. Th.: ROSETTA: a computer program for estimating soilhydraulic parameters with hierarchical pedotransfer functions, J. Hydrol., 251, 163–176, 2001.

Schuol, J., Abbaspour, K. C., Srinivasan, R., and Yang, H.: Estimation of freshwater availability in the West African sub-continent using the SWAT hydrologic model, J. Hydrol., 352, 30–49, 2008a.

Schuol, J., Abbaspour, K. C., Yang, H., Srinivasan, R., and Zehnder, A. J. B.: Modeling blue and green water availability in Africa, Water Resour. Res., 44, 1–18, 2008b.

Schmidt, R., Flechtner, F., Meyer, U., Neumayer, K. H., Dahle, C., König, R., and Kusche, J.: Hydrologic signals observed by the GRACE satellites, Surv. Geophys., 29, 319–334, doi:10.1007/s10712-008-9033-3, 2008.

Siebert, S., Burke, J., Faures, J. M., Frenken, K., Hoogeveen, J., Döll, P., and Portmann, F. T.: Groundwater use for irrigation – a global inventory, Hydrol. Earth Syst. Sci., 14, 1863–1880, doi:10.5194/hess-14-1863-2010, 2010.

Soil Conservation Service: Section 4: Hydrology in National Engineering Handbook, 1972.

Swenson, S., and Wahr, J.: Post-processing removal of correlatederrors in GRACE data, Geophys. Res. Lett., 33, L08402, doi:10.1029/2005GL025285, 2006.

Syed, T. H., Famiglietti, J. S., Rodell, M., Chen, J., and Wilson, C. R.: Analysis of terrestrialwater storage changes from

Hydrol. Earth Syst. Sci., 16, 3083–3099, 2012

www.hydrol-earth-syst-sci.net/16/3083/2012/

GRACE and GLDAS, Water Resour. Res., 44, W02433, doi:10.1029/2006WR005779, 2008.

Tang, Q., Gao, H., Yeh, P., Oki, T., Su, F., and Lettenmaier, D. P.: Dynamics of terrestrial water storage change from satellite and surface observations and modelling, J. Hydrometeorol., 11, 156–170, doi:10.1175/2009JHM1152.1, 2010.

Tapley, B. D., Bettadpur, S., Ries, J. C., Thompson, P. F., and Watkins, M. M.: GRACE measurements of mass variability in the Earth system, Science, 305, 593–505, doi:10.1126/science.1099192, 2004.

Tiwari, V. M., Wahr, J., and Swenson, S.: Dwindling groundwater resources in northern India, from satellite gravity observations, Geophys. Res. Lett., 36, L18401, doi:10.1029/2009GL039401, 2009.

Venetis, C.: A study of the recession of unconfined aquifers, Bull. Int. Assoc. Sci. Hydrol., 14, 119–125, 1969.

Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W., and Sorooshian, S.: Effective and efficient algorithm for multiobjective optimisation of hydrologic models, Water Resour. Res., 39, 1214, doi:10.1029/2002WR001746, 2003.

Wahr, J., Molenaar, M., and Bryan, F.: Time variability of the earth's gravity field: Hydrologic and oceanic effects and their possible detection using GRACE, J. Geophys. Res., 103, 30205–30230, 1998.

Werth, S. and Güntner, A.: Calibration analysis for water storage variability of the global hydrological model WGHM, Hydrol. Earth Syst. Sci., 14, 59–78, doi:10.5194/hess-14-59-2010, 2010.

Werth, S., Güntner, A., Petrovic, S., and Schmidt, R.: Integration of GRACE mass variations into a global hydrological model, Earth Planet. Sci. Lett., 277, 166–173, 2009.

Winchell, M., Srinivasan, R., di Luzio, M., and Arnold, J. G.: Arc-SWAT interface for SWAT 2005, User'sGuide, Blackland Research Center, Texas Agricultural Experiment Station, Temple, 2007.

World Bank: available at: http://data.worldbank.org/indicator/SP. POP.GROW (last access: September 2011), 2009.

Xie, H., Eheart, J. W., and An, H.: Hydrologic and economic implications of climate change for typical river basins of the agricultural Midwestern United States, J. Water Res. Pl.-ASCE, 134, 205–213, 2008.

Xie, H., Nkonya, E., and Wielgosz, B.: Assessing the risks of soil erosion and small reservoir siltation in a tropical river basin in Mali using the SWAT model under limited data condition, Appl. Eng. Agr., 27, 895–904, 2011.

Yang, Z. L., Niu, G. Y., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Longuevergne, L., Manning, K., Niyogi, D., and Tewari, M.: The community Noah land surface model with multiparameterisation options (Noah-MP): 2. Evaluation over global river basins, J. Geophys. Res., 116, D12110, doi:10.1029/2010JD015140, 2011.

Yirdaw, S. Z., Snelgrove, K. R., Seglenieks, F. R., Agboma, C. O., and Soulis, E. D.: Assessment of the WATCLASS hydrological model result of the Mackenzie River basin using the GRACE satellite total water storage measurement, Hydrol. Process., 23, 3391–3400, doi:10.1002/hyp.7450, 2009.

Zaitchik, B. F., Rodell, M., and Reichle, R. H.: Assimilation of GRACE terrestrial water storage data into a Land Surface Model: Results for the Mississippi River basin, J. Hydrometeorol., 9, 535–548, 2008.