Advances in Geosciences

Calibration and uncertainty issues of a hydrological model (SWAT) applied to West Africa

J. Schuol and K. C. Abbaspour

Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland

Abstract. Distributed hydrological models like SWAT (Soil and Water Assessment Tool) are often highly over-parameterized, making parameter specification and parameter estimation inevitable steps in model calibration. Manual calibration is almost infeasible due to the complexity of large-scale models with many objectives. Therefore we used a multi-site semi-automated inverse modelling routine (SUFI-2) for calibration and uncertainty analysis. Nevertheless, the question of when a model is sufficiently calibrated remains open, and requires a project dependent definition. Due to the non-uniqueness of effective parameter sets, parameter calibration and prediction uncertainty of a model are intimately related.

We address some calibration and uncertainty issues using SWAT to model a four million km^2 area in West Africa, including mainly the basins of the river Niger, Volta and Senegal. This model is a case study in a larger project with the goal of quantifying the amount of global country-based available freshwater. Annual and monthly simulations with the "calibrated" model for West Africa show promising results in respect of the freshwater quantification but also point out the importance of evaluating the conceptual model uncertainty as well as the parameter uncertainty.

1 Introduction

This study is part of the project "GIS-based hydrological modelling of global freshwater availability" with the main objective of quantifying the country-based freshwater availability at a sub-country level, a number widely sought after in many global studies. The available estimates of freshwater are imprecise and do not quantify the temporal and spatial distributions of the available water, which in some cases are

Correspondence to: J. Schuol
(juergen.schuol@eawag.ch)

more important than the available water figure itself. While former assessments of the country-based freshwater availability were mainly done on the basis of data generalization of the world hydrological network (e.g. Shiklomanov, 2000), there also exist a couple of global hydrological models with a spatial resolution of 0.5°. Probably the most sophisticated of these models is WaterGAP 2 (Alcamo et al., 2003; Döll et al., 2003), which calculates surface runoff and groundwater recharge. It is tuned against annual discharges at 724 gauging stations by adjusting the runoff coefficient and, in case this was not sufficient, by applying up to two correction factors.

Our goal is to model monthly values of river discharges (blue water) as well as the soil water (green water), which is the main source of rain-fed agriculture. Furthermore, it is essential to clearly show the uncertainty of the model results – a task that hasn't been carried out, hence, making uncertainty and risk analysis very difficult. To accomplish the objectives, it is essential to have a calibrated hydrologic model.

The focus of this paper is, therefore, on the calibration of the integrated, continuous-time, large-scale daily water balance model SWAT. The model is spatially distributed and accounts for differences in soils, land use, topography and climate. Furthermore, SWAT provides the possibility to extend the hydrological model with sediment and nutrient sub-models – an issue that will become the subject of a follow-up project. More detailed descriptions of the model can be found in Arnold et al. (1998) and Arnold et al. (1999).

It is well-known that calibration of distributed watershed models is a very difficult task that doesn't lead to a unique solution. Over the years different approaches have been proposed (Duan et al., 2003), and one main trend is to perform multi-objective, global, automated inverse modeling (e.g., Duan et al., 1992; Beven and Binley, 1992). SUFI-2 (Sequential Uncertainty Fitting, ver. 2, Abbaspour et al., 2004), the procedure we apply in this study, belongs to this group and combines parameter calibration and uncertainty prediction.
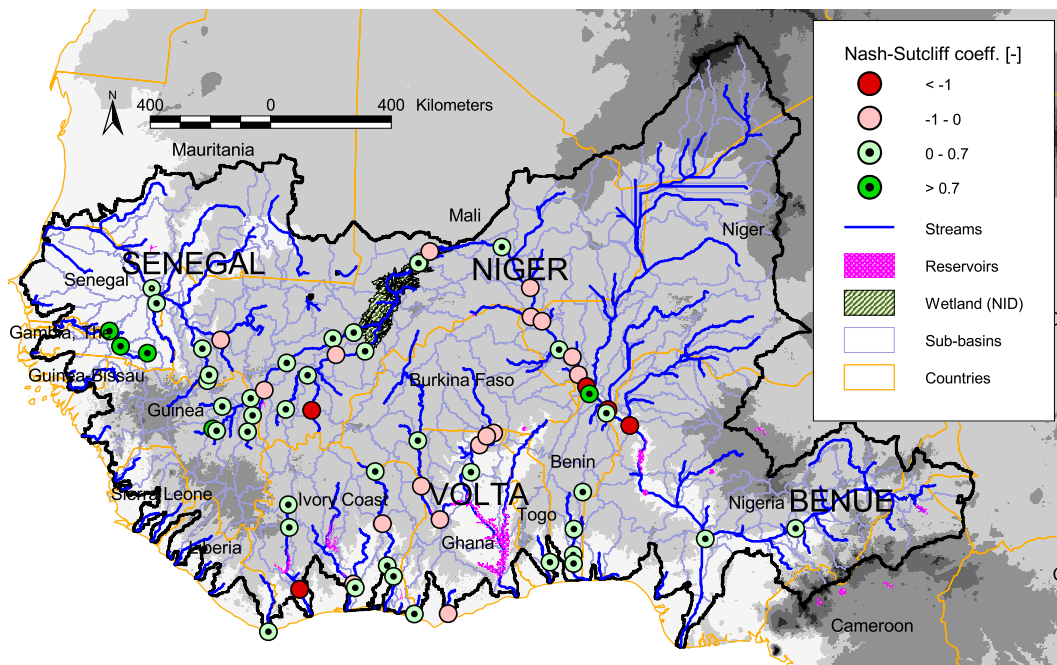
Fig. 1. Nash-Sutcliff coefficient of the monthly runoff calibration results at all 64 stations.

A four million km^2 area in West Africa was selected as a case study for the analysis of a large-scale hydrological SWAT model. The setup of this model will be described in Sect. 2. Model calibration is an inevitable step in the development of any complex hydrological model and requires efficient methods for parameterisation and parameter estimation. In Sect. 3 we present such a procedure together with first results for the West African model. The limited database and the non-uniqueness of effective model parameters result in substantial model uncertainty. In addition, we have to deal with conceptual model uncertainty due to simplified and/or not included real world processes. Some approaches to quantify and minimize these uncertainties are presented but while the calibration is an ongoing process, further in-depth research is still necessary.

2 Setup of the West Africa model

The selected watershed in West Africa (approx. one-seventh of Africa) includes the basins of the river Niger, Volta and Senegal (Fig. 1). The area covers climatic zones from hyper-arid to humid and accordingly, though savannah dominates, the land use varies from barren/sparsely vegetated deserts to evergreen rainforests. This area was selected because of its severe water problems but also for scientific reasons, as it is quite challenging due to the comparably small database, which is typical for most areas outside of Europe and North-America. In order to setup a hydrological SWAT model, the following basic data are required: topography, soil, land use

and climatic data. The digital maps and databases for West Africa are all extracted from a global database. The model parameterisation was derived using the ArcView GIS inter-face for SWAT (Di Luzio et al., 2001), which provides a graphical support for the disaggregation scheme and thus facilitates the data handling. First, the whole watershed was subdivided into 292 sub-basins (Fig. 1), with a minimum drainage area of 10 000 km^2. In the next step, land uses and soils were characterized and overlaid in each sub-basin. Finally, the weather input files were created. The climatic data is one of the most fundamental model inputs of SWAT but in many areas of the world – and also in West Africa – the measuring station network is not very dense and the time periods with measured data are short and/or have many missing and sometimes even erroneous data. Therefore, based on SIMMETEO (Geng et al., 1986), we developed and applied a daily weather generator algorithm (Schuol and Abbaspour, 2006[1]) that uses the currently available 0.5° monthly weather statistics from the Climatic Research Unit (Mitchell and Jones, 2005; New et al., 2000). Using this procedure we obtained time series of daily precipitation as well as minimum and maximum temperature for each subbasin and assigned them to the subbasin centroid. This step completed the SWAT setup, and we simulated a 25-year period from 1971 to 1995.

---

[1]Schuol, J. and Abbaspour, K. C.: A daily weather generator for predicting rainfall and maximum-minimum temperature using monthly statistics based on a half-degree climate grid, Ecol. Model., submitted, 2006.
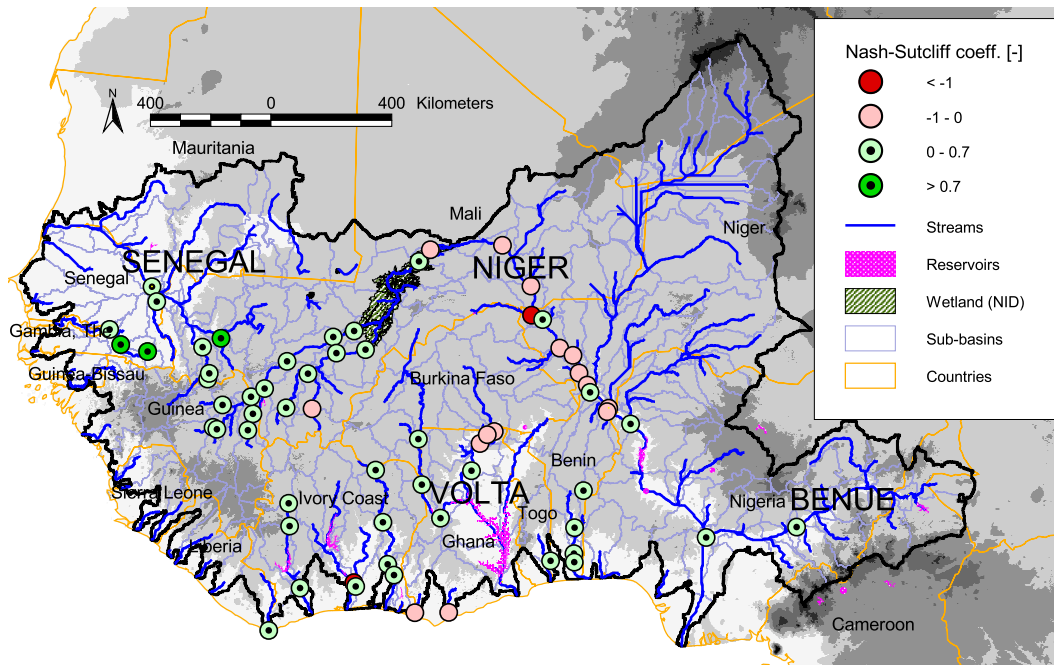
Fig. 2. Nash-Sutcliff coefficient of the monthly runoff validation results at all 64 stations.

3 Calibration procedure

Measured river discharges at 64 stations in West Africa were used for annual and monthly calibration purposes. At many stations the available data doesn't cover the whole simulation period of 25 years but the available years/months were always split into equal time periods for calibration (more recent data) and validation (prior data). An initial annual calibration was followed by a monthly calibration.

SWAT is a distributed hydrological model and consequently there are potentially many (thousands) parameters. As it is impossible to calibrate all of them, a reduction of the number of parameters to estimate – the so-called parameter specification – is inevitable. Based on an initial absolute sensitivity analysis (sequentially varying one parameter while keeping all others constant) and relative sensitivity analyses (varying all parameters simultaneously), we decided which parameters to include in the calibration.

Traditional manual parameter estimation is almost infeasible due to the complexity of large-scale models and a large number of parameters. We used the SUFI-2 program for a combined optimization-uncertainty analysis. SUFI-2 is a multi-site, semi-automated global search procedure. The objective function was formulated as the Nash-Sutcliff (*NS*) coefficient between the measured and simulated discharges. In SUFI-2, parameter uncertainty is depicted as uniform distributions. This uncertainty is propagated through a Latin hypercube sampling scheme, and the uncertainty (referred to as the 95% prediction uncertainty, 95PPU) is calculated at the 2.5% and 97.5% levels for each simulated variable. A

procedure – systems analysis interface (SWAT-SAI) developed by Yang et al. (2005)[2] was used to assign and update the SWAT parameter files. Thus the model could be run n times and results could be analyzed automatically for all n parameter combinations. Starting from initially large parameter uncertainties, SUFI-2 is iterated a few times until an optimum solution is reached. Each iteration results in a narrower parameter uncertainty. Two stopping rules quantifying the uncertainty are defined: (1) bracketing "most" of the measured data within the 95PPU band, and (2) obtaining a "small" ratio of the average distance between the 2.5th and 97.5th prediction percentiles and the standard deviation of the measured data (*R-factor*). These two measures quantify the model uncertainty. The ideal situation would be to account for 100% of the observed data in the 95PPU while at the same time have an *R-factor* close to zero. But this is seldom the case because of measurement errors, conceptual model uncertainty, and non-uniqueness issues. To answer the question of "when a model is calibrated" remains a problem dependent and subjective issue. The values of the % bracketed data, *R-factor*, as well as the R^2 and *NS* between the observation and the best simulation (i.e., the simulation with the smallest value of the objective function) determine the strength of a calibrated model.

---

[2]Yang, J., Abbaspour, K. C., and Reichert, P.: Interfacing SWAT with Systems Analysis Tools: A Generic Platform, Environ. Modell. Softw., submitted, 2005.

Table 1. The ten SWAT model parameters included in the calibration procedure and their final calibrated ranges.

Parameter name	Definition	Final parameter range
CN2_GRAS	Curve number for grassland [–]	39–64/49–74*
CN2_SAVA	Curve number for savannah [–]	30–50/42–62/47–67*
SOL_AWC_SCL	Soil available water storage capacity for the soil texture type "sandy-clay-loam" [mm H_2O/mm soil]	0.145–0.175
ESCO	Soil evaporation compensation factor [–]	0.12–0.50
SURLAG	Surface runoff lag coefficient [days]	1.5–5.5
GWQMN	Threshold depth of water in the shallow aquifer required for return flow to occur [mm H_2O]	50–100
REVAPMN	Threshold depth of water in the shallow aquifer for "revap" or percolation to the deep aquifer to occur [mm H_2O]	1–40
RCHRG_DP	Deep aquifer percolation fraction [–]	0.56–0.70
MSK_CO1	Calibration coefficient that controls impact of the storage time constant for normal flow [–]	2–9
MSK_CO2	Calibration coefficient that controls impact of the storage time constant for low flow [–]	1–6

* Ranges vary depending on the hydrologic soil group within one land use type

4 Results and discussion

The results for the West Africa model presented in the following are preliminary and rather a basis for discussion of further improvements. Ten of the most sensitive parameters were included in the calibration procedure (Table 1). Figure 1 shows the *NS* of the monthly runoff calibration results at 64 stations, while the validation results are presented in Fig. 2. A closer look reveals a clear pattern where most of the stations in the west have a positive *NS* and many of them also possess a *NS* higher 0.7 but stations further downstream the River Niger are insufficiently simulated. The stations with a positive *NS* in the East are all at tributaries to the River Niger having comparatively small watersheds. Figure 3 illustrates the wide range of quality in the model fit that ranges from very good fits (e.g. Gouloumbou on River Gambia: $NS_{calib.}=0.82$ and $NS_{valid.}=0.54$) to very poor fits (e.g. Malanville on River Niger: $NS_{calib.}=-1.16$ and $NS_{valid.}=-0.63$).

4.1 Parameter uncertainty

The 95PPU as calculated represents a combined model prediction uncertainty including parameter uncertainty resulting from the non-uniqueness of effective model parameters, conceptual model uncertainties, and input (i.e., rainfall) uncertainties. In SUFI-2, the combined effect of all uncertainties is depicted by the final estimates of parameter uncertainties. The idea is that given all the uncertainties, a more precise estimation of the parameter ranges cannot be made. Figure 4 illustrates the 95PPU intervals of the last iteration for an extract of the calibration and validation period at the station Gouloumbou (River Gambia). In the initial iteration, 80% of the observed monthly runoff values at all the 64 stations were

within the 95PPU, but the *R-factor* was quite large (2.92) indicating large model uncertainties. In subsequent iterations for the calibration period the *R-factor* decreased to 0.80 (validation: 0.75) but the 95PPU bracketed only 32% of the observed discharge values (30% for validation), i.e., most of the measurements were unaccounted due to too small uncertainties in the parameters. Hence, striking a balance between these two measures provides a final calibration result.

4.2 Conceptual model uncertainty

The large 95PPU band (or large *R-factor*) necessary to bracket 80% of the observed data indicates that the uncertainty in the conceptual model is also very important, and in our case quite large. It seems that not all processes, especially some that are important further downstream of the River Niger, are included in the model. We believe that these processes are mainly related to the existing large reservoirs (Fig. 1) regulating the runoff of the river Niger and also the large Niger Inland Delta (NID, Fig. 1) delaying the runoff and significantly contributing to higher evaporation losses. To show the potential runoff regulating effect of reservoirs on the downstream hydrology we use an example from Aswan dam in Egypt (Fig. 5a). In our modelling area, while there are quite a few large dams, a long data series, including measurements before and after the closure, is not available. Nevertheless, the effects are comparable to the Aswan dam even though not always so striking. The influence of the NID on the hydrological regime of the River Niger might be of an even higher importance. Figure 5b shows the measured runoff at the gauging stations up- and downstream of the approximately 500-km long NID for the three years 1987, 1990 and 1992. The peak discharge decreases within the wetland
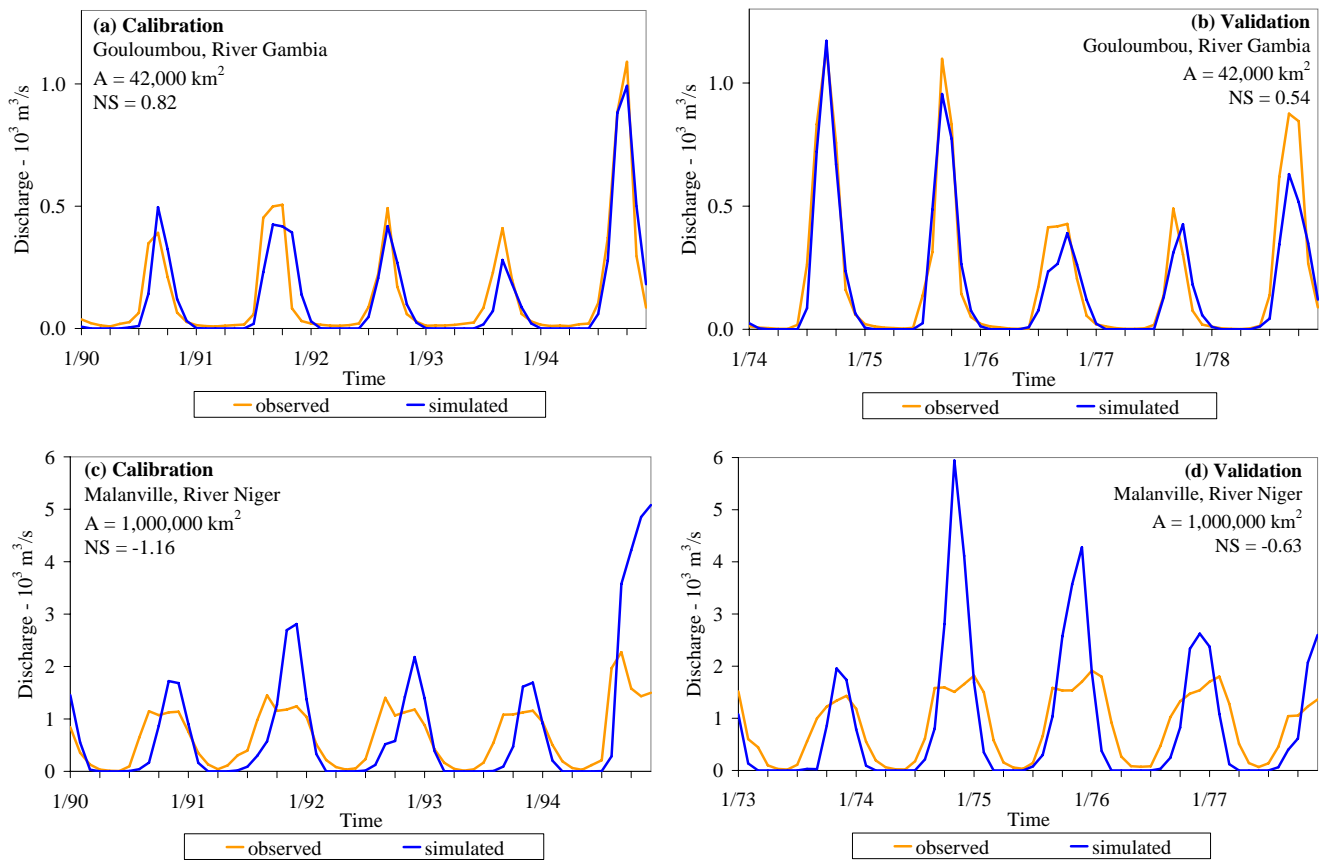
Fig. 3. Extract of the monthly calibration and validation results for runoff at two of the 64 stations: **(a)** and **(b)** Gouloumbou at the River Gambia, **(c)** and **(d)** Malanville at the River Niger.
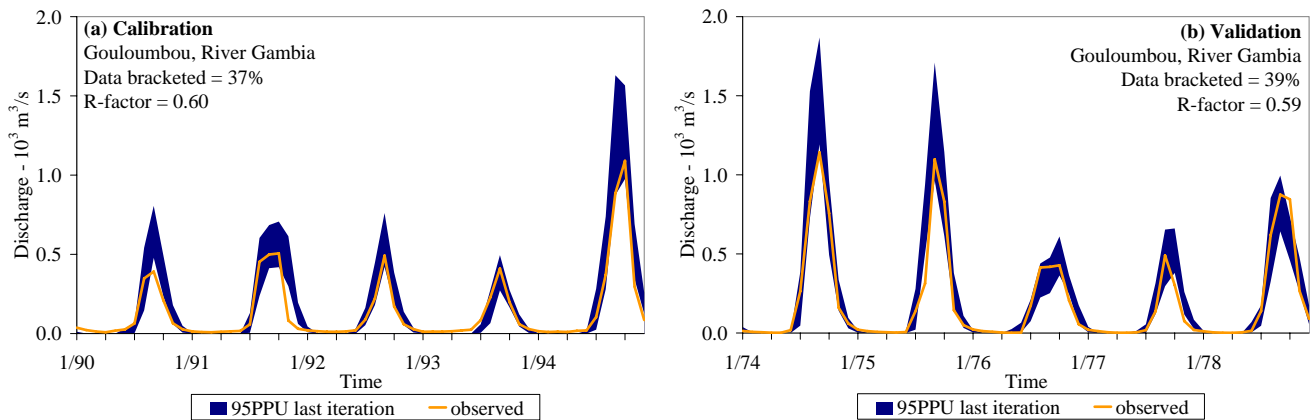

Fig. 4. Monthly calibration **(a)** and validation **(b)** results for Gouloumbou (River Gambia) showing the 95% prediction uncertainty intervals along with the measured discharge.

by about 50 percent and is delayed by about two months. In addition all types of water use and especially the irrigation

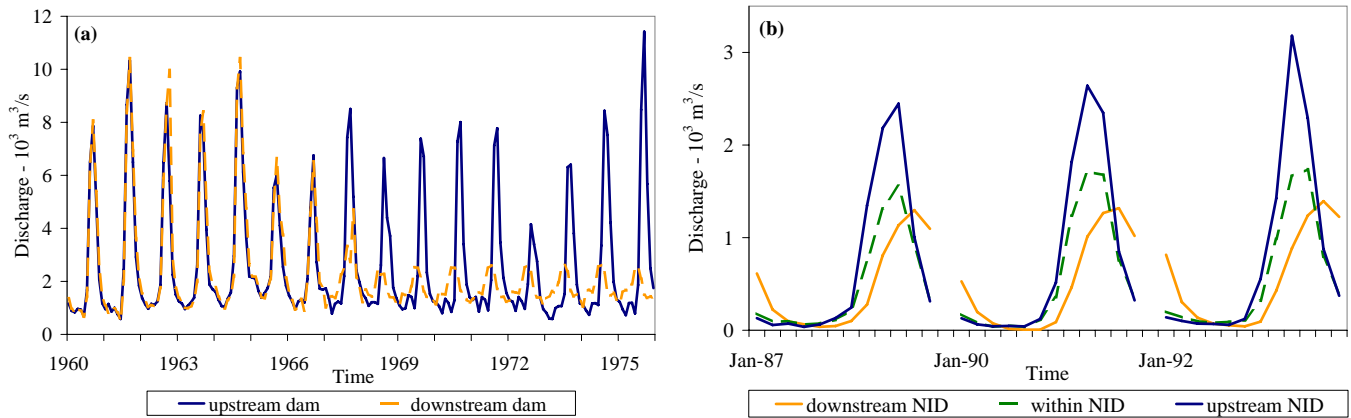use, which mostly have local importance, are currently also not accounted in the model.

Fig. 5. (a) Observed discharge up- and downstream of the Aswan dam in Egypt – before and after its closure in 1968 **(b)** Observed discharge within, up- and downstream of the Niger Inland Delta (NID).

One approach in dealing with this shortcoming would be to exclude the gauging stations affected by the reservoirs and wetlands from the calibration procedure. This would leave very few stations for calibration, and bias the water availability estimates assuming unreal watershed conditions. The other approach, which we will follow, is to include the reservoirs, wetlands and water uses as far as possible, even though detailed information on the management of the reservoirs and also on stored water in the wetland are not yet available.

5　Conclusion and outlook

Considering the example of West Africa, we showed that SWAT can be used for large-scale water quantity investigations and SUFI-2 is an efficient parameter optimization-uncertainty analysis procedure. All the uncertainties (input, parameter and model) are combined and attributed to the parameters and expressed through the 95% prediction uncertainty. While not shown in detail, data uncertainty, e.g. in the climatic input, is certainly also of great importance. For a data scarce region like West Africa previous studies showed that the discharge simulations using generated data were superior to the simulations using available measured data from local climate stations (Schuol and Abbaspour, 2006[1]).

An improved calibration, bracketing most of (>80%) the observed data will be pursuit by the addition of the largest reservoirs and the wetland (NID) in the model. In the near future we will create a model for the whole continent of Africa, making use of the experience gained in West Africa and continue to approach our global freshwater quantification goal. Based on the presented work, we will further emphasize on a quantification of the uncertainty in the freshwater availability estimates.

References

Abbaspour, K. C., Johnson, C. A., and van Genuchten, M. T.: Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure, Vadose Zone J., 3, 1340–1352, 2004.

Alcamo, J., Döll, P., Henrichs, T., Kaspar, F., Lehner, B., Rösch, T., and Siebert, S.: Development and testing of the WaterGAP 2 global model of water use and availability, Hydrol. Sci. J., 48(3), 317–337, 2003.

Arnold, J. G., Srinivasan, R., Muttiah, R. S., and Allen, P. M.: Continental scale simulation of the hydrologic balance, J. Am. Water Resour. As., 35(5), 1037–1051, 1999.

Arnold, J. G., Srinivasan, R., Muttiah, R. S., and Williams, J. R.: Large area hydrologic modeling and assessment part I: model development, J. Am. Water Resour. As., 34(1), 73–89, 1998.

Beven, K. and Binley, A.: The future of distributed models: Model calibration and uncertainty prediction, Hydrol. Processes, 6, 279–298, 1992.

Di Luzio, M., Srinivasan, R., and Arnold, J. G.: ArcView interface for SWAT2000 – User's guide, Blackland Research Center, Temple, Texas, 2001.

Döll, P., Kaspar, F., and Lehner, B.: A global hydrological model for deriving water availability indicators: model tuning and validation, J. Hydrol., 270, 105–134, 2003.

Duan, Q., Sorooshian, S., and Gupta, V.: Effective and efficient global optimization for conceptual rainfall-runoff models, Water Resour. Res., 28, 1015–1031, 1992.

Duan, Q., Sorooshian, S., Gupta, V., Rousseau, A. N., and Turcotte, R.: Calibration of watershed models, American Geophysical Union, Washington D.C., 345p., 2003.

Geng, S., Penning de Vries, F. W. T., and Supit, I.: A simple method for generating daily rainfall data, Agr. Forest Meteorol. 36, 363-376, 1986.

Mitchell, T. D. and Jones, P.: An improved method of constructing a database of monthly climate observations and associated high-resolution grids, Int. J. Climatol. 25, 693–712, 2005.

New, M., Hulme, M., and Jones P.: Representing twentieth-century space-time climate variability, Part II: Development of 1901–96 monthly grids of terrestrial surface climate, J. Clim., 13, 2217–2238, 2000.

Shiklomanov, I. A.: Appraisal and assessment of world water resources, Water Int., 25(1), 11–32, 2000.