Advances in Geosciences

Integrating heterogeneous landscape characteristics into watershed scale modelling

S. Julich[1,*], L. Breuer[1], and H.-G. Frede[1]

[1]Institute of Landscape Ecology and Resources Management, Justus-Liebig-University Gießen, Germany
[*]now at: Department for Soil Science and Site Ecology, Dresden University of Technology, Dresden, Germany

Correspondence to: S. Julich (stefan.julich@tu-dresden.de)

Abstract. Two versions of the SWAT-model with different dominating runoff generation processes have been applied. One version comprises the original available SWAT version where only the basic input data are used. In the second version SWAT has been modified, by the integration of an impermeable layer at the subbasin level, in order to better reflect the boundary between soil and bedrock that results in increased lateral flow in low mountain ranges. As well, since conventional German soil maps do not describe soil horizons beyond 2 m depth, we also added a 4 m fixed depth in the lowland areas in order to reflect the deep loess deposits in this region. The decision for the location of the impermeable and the additional loess layer was based on a GIS analysis of additional geologic information.

This study revealed that both model versions produced acceptable and comparable results regarding the evaluated goodness of fit measures. The GLUE analysis showed that the SWAT model set up with additional information about the distribution of impervious soil layers and the loess depth in the lowlands produced the highest simulation quality and the lowest uncertainty. Moreover, SWAT II version was able to better represent the spatial extend of the dominating runoff processes best. This leads to the conclusion that the SWAT II version is better suited for scenario analysis than the original model version.

al., 2005; Beven, 2001). Hence the values of such conceptual parameters need to be determined in a calibration process. A "well-calibrated" model is essential for scenario analysis. Following Gupta et al. (2005) the necessary features of such a "well-calibrated" model are the ability of the model to reproduce the measured behaviour of the study catchment with small prediction uncertainty. Another vital feature is the capability of the model to capture the dominant ecological and hydrological processes. This is a challenging task, particularly in watersheds with heterogeneous landscape structures and elements. One way to achieve this is the integration of data which can not directly used in the model (soft data) or expert knowledge into the model calibration or validation process (Seibert and McDonnell, 2002). Vaché et al. (2004, 2006) have used soft data on mean transit time to test and reject several model structures in order to find a model which represents the governing hydrological processes best. Another way to consider soft data is the use of expert knowledge, for example through the analysis of spatial data which are not imperative for the model set up, but could be used as a additional qualitative information. In the following paper we show how to use such expert knowledge, gained by the analysis of additional soil and geological data to parameterize a semi distributed hydrological model. Through an extended uncertainty analysis we show that the additional information helps to reduce the predictive uncertainty of the model.

1 Introduction

Mathematical models representing the water and nutrient balance at watershed scale are gaining importance, for example to assess measures for improving water quality in the Water Framework Directive. Models are always simplifications of natural systems which often rely on conceptual parameters which cannot be measured in the real world (Gupta et

2 Material and methods

2.1 Study area

The study was conducted in the mesoscale Wetter catchment with an area of 514 km^2, located approximately 50 km north of Frankfurt/Main. The catchment can be divided into three major landscape units which are characterised by
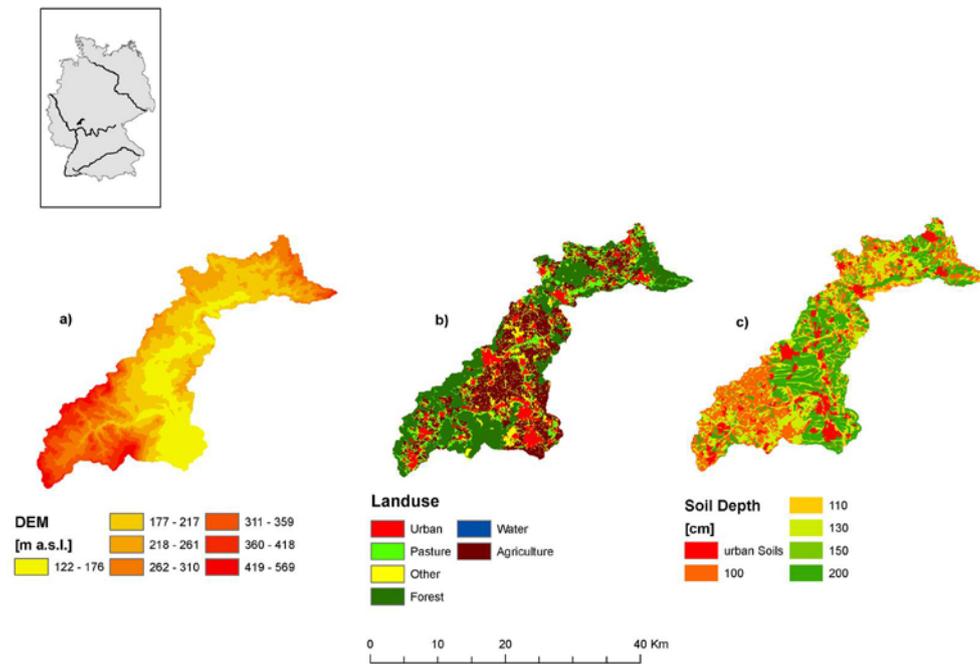
Fig. 1. Overview of the study area; with **(a)** elevation, **(b)** land use and **(c)** soil depth.

heterogeneous characteristics. The SW part is formed by the low mountain range of the Taunus. The central part of the watershed belongs to the lowland of the Wetterau. The NW region is formed by the foothills of the Vogelsberg, a volcanic low mountain range. Elevation ranges from 122 m to 568 m a.s.l. The mean annual temperature is around 8–10 °C and the mean annual precipitation is around 500 mm in the central part and up to 1000 mm at higher elevations (Fig. 1).

The low mountain ranges are characterized by steep slopes and shallow soils with an average soil depth of 1 m. In these parts the soils are mostly underlain by bedrock at a general depth of 1 m and the lowest soil horizons exhibit rock contents of more than 70 %. Here, lateral subsurface stormflow seems to be the dominating hydrological process. Soils in the undulating lowland part of the watershed are deep with an average soil depth of around 2 m. These soils were formed mostly from large Pleistocene loess deposits and can exhibit a thickness of up to 10 m in extremes. Vertical water movement in these soils is expected to be dominant.

The land use in the area under study reflects the heterogeneous soil characteristics. The central part with the deep loess soils is dominated by intensive agriculture whereas in the low mountain ranges with the shallow soils mainly forests or pastures can be found (Fig. 1).

2.2 Model

In the study we applied the widely known Soil and Water Assessment Tool (SWAT, Arnold et al., 1998) is a semi-distributed mathematical model to represent the water and

nutrient cycle at the watershed scale. It relies on empirical as well as physical based process descriptions. The watershed is divided into subcatchments based on topography and the river network. These sub catchments are further splitted into HRU (hydrological response units) consisting of similar slope, land use and soils. The model consists of several submodules which describe processes regarding soil hydrology, plant growth, nutrient cycling (nitrogen and phosphorus), land management (agricultural practices), pollutant fate (mainly pesticides), in-stream water quality and water routing. All processes considered in SWAT are calculated at a daily time step. Model outputs of SWAT are water discharge, nutrient and pollutant loads at the subbasins scale as well as parameters of the water and nutrient balance at the HRU scale (Neitsch et al., 2005).

2.3 Input data and model set up

Figure 2a shows the location of the climate and gauging stations used for model input and validation of model simulations. Climate data from two temperature stations and seven precipitation stations were used for the period 2005 to 2008. Discharge measurements at daily time steps were available for the same period for the gauging stations at the outlet of the watershed as well as for four additional stations distributed within the watershed. During the period 2006 to 2008 snapshot measurements of nitrate concentrations were taken at 4 locations in the catchment on an almost biweekly basis. Additionally loads are calculated from the measured discharge and nitrogen concentration at 4 sites.
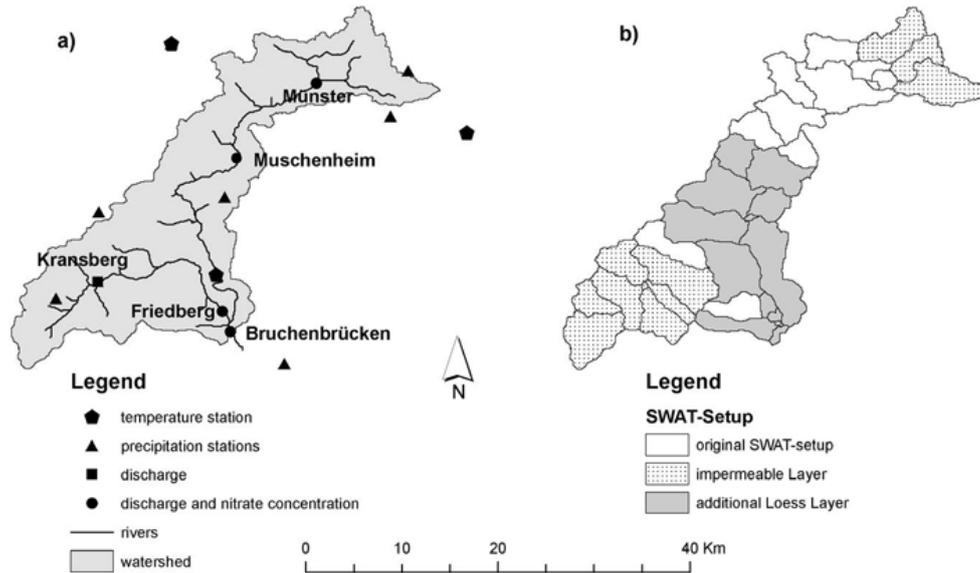
Fig. 2. Location of gauging stations **(a)** and model setup for model structure SWAT II **(b)**. In addition to the general model set up of SWAT I, impermeable soil layers in the low mountain region as well as additional deep soil layers in the centre of the watershed have been added in SWAT II.

Two different model versions of SWAT were applied in this study. The first version (further called model SWAT I) is based on the established SWAT model set up with basic spatial input data needed on topography, soils and land use (Fig. 1). In the second model, SWAT II (Fig. 2b), we included additional information on the geologic characteristics of the watershed with the goal to better reflect the landscape characteristics and the dominating hydrologic processes. In low mountainous areas we added an impermeable layer at a depth of 1.2 m at the subbasin scale in order to better reflect the boundary between soil and bedrock. This stresses the dominant hydrologic process of subsurface flow in SWAT II. The decision whether to add the impermeable layer or not is made based on GIS-analyses of slope, soil depth and rock content. Only those subbasins received an impermeable layer where steep slopes (slopes larger than 20 %), soil depths of at most 1 m and rock contents of more than 75 % in the lowest horizon are dominant. In the low land we added an additional soil layer of 4 m depth with the properties of the typical loess soils for a better representation of the deep pleistocenic deposits in the catchment. The layers were only added to subbasins and those soils which have a reported soil depth of 2 m and where the depth of the loess deposits are larger than 4 m. The subbasins where neither of the aforementioned landscape features are edominant (no signature in Fg. 2b) are set up with the basic input data used also in SWAT I. Information for the SWAT II set up was gained from the analysis of available soil data and as well as from the explanation reports from bore hole drillings and geological maps (Kegel, 1979; Kümmerle, 1976, 1981).

2.4 GLUE approach

To compare both model versions and to identify the model structure with the better we carried out an uncertainty analysis, by applying the GLUE approach to both models (Beven and Binley, 1992; Beven and Freer, 2001). The GLUE method is based on Monte-Carlo simulations where an arbitrary number of parameters are varied within user-defined ranges. For each combination of parameters a model realization is obtained. The simulated and measured outputs for a variable of interest are compared for one or more goodness-of-fit measures (e.g. Legates and McCabe, 1999). By means of a user-defined threshold for the evaluated goodness-of-fit (GOF) measure all realizations are discriminated into behavioural and non-behavioural model runs. All non-behavioural model runs are discarded. The goodness-of-fit values for the behavioural runs are weighted and rescaled in a way that their sum equals one. With the rescaled goodness-of-fit measures a cumulative distribution is calculated from which user-defined uncertainty bounds can be calculated.

For the GLUE analysis we generated 5000 parameter sets with a Latin hypercube sampling scheme by varying model parameters concerning channel routing, groundwater description, soil physical properties and nitrogen balance. The varied parameters and their ranges can be obtained from Table 1. For each parameter set one model run has been realized and a set of goodness-of-fit measures has been calculated for the time period 2007–2008. The GOF are the model efficiency Nash Sutcliff Efficiency (NSE) and logarithmized model efficiency NSE-log following Nash and

Table 1. SWAT parameter ranges used for the Monte-Carlo simulation.

SWAT-model parameter [units]	Lower bound	Upper bound
Surface run off lag time [d]	1	10
Manning's roughness coefficient main channel [–]	0.01	0.3
Hydraulic conductivity channel bottom [mm h^{-1}]	0	25
Baseflow alpha factor [d^{-1}]	0.01	1
Groundwater delay time [d]	1	25
Groundwater revap coefficient [–]	0.02	0.2
Threshold for baseflow [mm]	0	100
Available water capacity [mm mm^{-1}]	−15*	15*
soil hydraulic conductivity I [mm h^{-1}] (<75% rock)	−15*	15*
soil hydraulic conductivity II [mm h^{-1}] (>75% rock)	10	150
Rate factor humus mineralization [–]	0.0001	0.001
Nitrogen uptake distribution parameter [–]	10	30
Nitrate percolation coefficient [–]	0.01	1
Residue decomposition coefficient [–]	0.01	0.1
Denitrification rate coefficient [–]	0	3.0
Threshold for denitrification [–]	0.5	1.10

* value altered ±15 % from input

Sutcliffe (1970), the Index of Agreement (d) following Willmott (1981) as well as the absolute model bias (PBIAS). Equations (1)–(4) describe the calculated GOF with S_i being the simulated value (discharge, nitrate load) at time step i and O_i representing the observed value at time step i. $\bar{O}$ is the mean of all measured values.

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{N}(S_i - O_i)^2}{\sum_{i=1}^{N}(O_i - \bar{O})^2} \qquad (1)$$

$$\mathrm{NSE-log} = 1 - \frac{\sum_{i=1}^{N}(\log S_i - \log O_i)^2}{\sum_{i=1}^{N}(\log O_i - \log \bar{O})^2} \qquad (2)$$

$$d = 1 - \frac{\sum_{i=1}^{N}(O_i - S_i)^2}{\sum_{i=1}^{N}(|S_i - \bar{O}| + |O_i - \bar{O}|)^2} \qquad (3)$$

$$\mathrm{PBIAS} = \frac{|\sum_{i=1}^{N}(S_i - O_i)|}{\sum_{i=1}^{N} O_i} \cdot 100 \qquad (4)$$

We calculated NSE, NSE-log, d and PBIAS to compare simulated and observed discharge for the five discharge gauging stations. For the evaluation of simulated nitrate loads we calculated d and PBIAS for the four sites where nitrate concentrations and calculated loads were available. To achieve a good simulation quality for the whole watershed we calculated the mean value for each GOF over all measurement sites, resulting in a total of six GOF values. These mean values were then taken as objective function for the GLUE analysis. We performed one GLUE analysis for each GOF. To make the GLUE analysis comparable to the applied model

versions we decided to set the threshold for identification of acceptable and non-acceptable runs to 70 % of the maximum value achieved for each GOF measure. For the further uncertainty analysis we only selected those parameter sets which were identified as behavioural for all six GOF.

3 Results and discussion

The GLUE analysis revealed that for SWAT I no parameter set could be identified acceptable for all six objective functions. After omitting NSE-log as objective function for the analysis 325 parameter sets were identified as acceptable for all remaining GOF measures. We used the remaining set of model runs in the further analyses of the SWAT I model structure, keeping in mind the reduced requirement we attributed to this ensemble. For SWAT II in contrast 19 parameter sets were recognized to be behavioural for all six GOFs.

Figure 3 presents the values of the two SWAT ensembles for the GOF concerning the simulated discharge at all five gauging stations for the simulation period 2007–2008. Both model versions yield the best values for all objective functions at the main outlet. Considerable differences between both models can be seen for PBIAS and NSE-log where SWAT II has a better performance predominantly in the three headwater catchments. These headwater catchments are mainly dominated by the low mountain ranges where SWAT I predicts groundwater flow and SWAT II lateral flow as the dominant runoff process. As the NSE-log criterion especially considers periods with low discharge (Bekele and Nicklow, 2007) we conclude that SWAT II is more capable to reproduce low flow dynamics as compared to SWAT I. Slightly less good results of discharge simulations for NSE
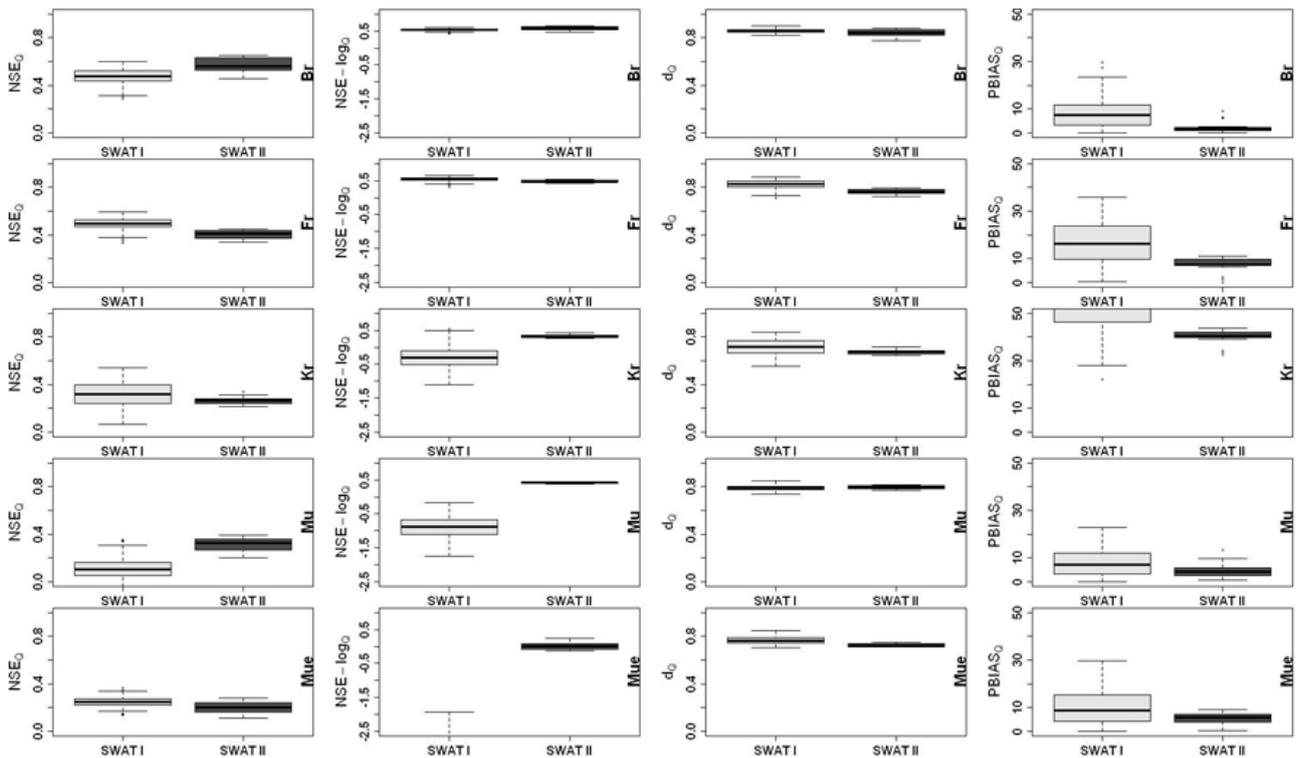
Fig. 3. Performance of the parameter ensembles of SWAT I and SWAT II for the goodness-of-fit measures concerning discharge; the respective gauging stations are indicated with Br = Bruchenbrücken, Fr = Friedberg, Kr = Kransberg, Mu = Muschenheim, Mue = Münster.

for SWAT II can be explained by a offset of simulated and measured discharge peak, indicating that the absolute amount predicted discharge is in agreement but that wrong timing is responsible for the reduced model performance.

Figure 4a–d presents the average annual groundwater flow and the average annual lateral flow at the subbasin scale for the two SWAT structures derived from the GLUE analysis. SWAT I (Fig. 4a and c) simulates high groundwater flow and low lateral flow in the low mountain ranges. SWAT II (Fig. 4b and d) shows the opposite behaviour in predicting mainly lateral subsurface flow in the hilly regions. These results are in agreement with other studies which also predicted lateral flow as major runoff process in low mountain ranges (Hergesell, 2003; Eckhardt et al., 2002). For the central part of the watershed SWAT II predicts low groundwater flow as well as low lateral flow. This could be expected from the landscape features of this area with the large loess deposits. These deposits act as soil storages which only slowly releases water. These findings are supported by observations of Kümmerle (1976, 1981) who stated that the amount of discharge of rivers flowing into this region is not increasing and that the river are losing water via infiltration through the river bed, indicating that barely no baseflow contribution takes place in that part of the watershed.

Figure 5 presents the simulated and observed hydrographs for the gauging stations Bruchenbrücken (main outlet) and
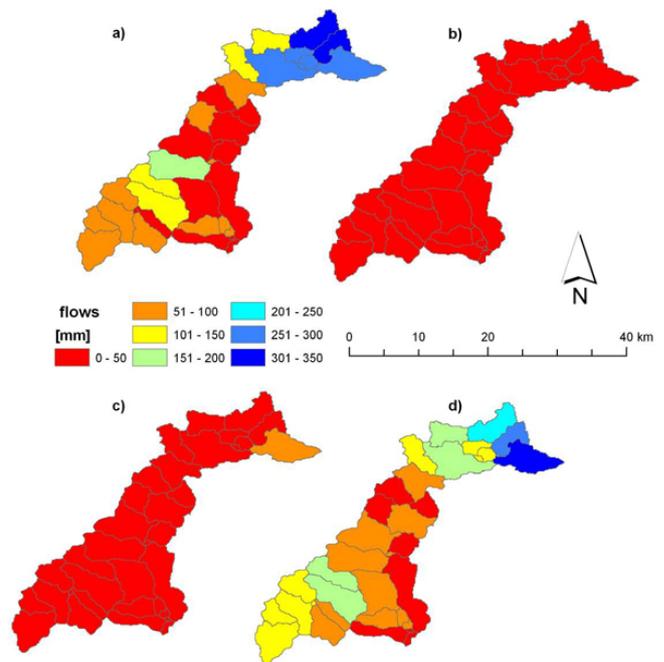
Fig. 4. Mean annual predicted groundwater flow for SWAT I **(a)** and SWAT II **(c)** and mean annual predicted lateral subsurface flow for SWAT I **(b)** and SWAT II **(d)**.
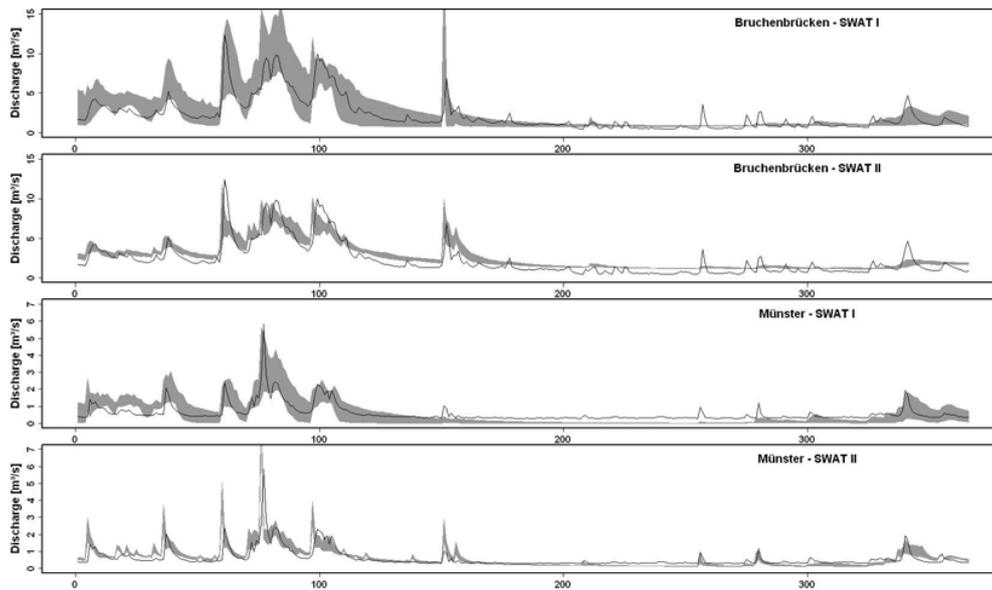
Fig. 5. Span of simulated (shaded grey) and observed (black line) discharge for the gauging stations Bruchenbrücken and Münster for the year 2008 and the parameter ensembles of SWAT I and SWAT II.

Münster (headwater subbasin) for the two SWAT versions for the last year (2008) of the simulation period. The simulated hydrographs are depicted as 95 % uncertainty bands in order to express the predictive uncertainty of the considered model versions. In general both models show a sufficient agreement with the seasonal variations and dynamics of the observed hydrographs. This applies also for the gauging stations and periods not presented here. The major difference between the individual SWAT versions is the width of the calculated uncertainty bands shown in Fig. 5. The large discrepancy is partly explained by the additional GOF of the SWAT II model that further constraints behavioural parameter sets (325 models set ups for SWAT I as compared to 19 models set ups for SWAT II). The parameter ensemble of SWAT I yields a maximum value for NSE of 0.59 whereas the ensemble for SWAT II reaches slightly higher values of NSE with 0.65. Results of both SWAT predictions concerning NSE are in the range of SWAT applications published elsewhere (Rouhani et al., 2007; Moriasi et al., 2007; Fohrer et al., 2002) and can be considered as acceptable in this regard. However, the two SWAT models structures fail to reproduce single storm events during the summer time, most likely due to the fact that no precipitation is recoded around these dates by any of the precipitation stations. Over estimations of low flow periods especially at the main outlet is attributed to the over prediction of effluent discharges from waste water treatment plants (WWTP) by the model. WWTP located in the catchment have a total size of 250 251 population equivalents. To account for the WWTP input we used a constant daily discharge according to their size as no further detailed data are available.

Figure 6 presents the predicted nitrate loads for the gauging stations Bruchenbrücken (main outlet) and Münster (nested headwater subbasin) for the two SWAT versions for the last year of the simulation period. Comparable to the predicted flows SWAT I exhibit a wider uncertainty band compared to SWAT II. Nitrate is highly soluble and leaves the soils via leaching. Therefore the nitrate export in a watershed is closely linked to the dominant hydrological runoff processes. As demonstrated above the SWAT I was not able to predict the expected hydrological processes in the watershed which propagates also in a flawed depiction of nitrate transport pathways. SWAT II was found to be better representing the runoff generation processes in the catchment which speaks also for a better representation of the nitrate export in the model, reflected in the smaller uncertainty bands of SWAT II in Fig. 6. The credibility of both model structures with respect to the simulation of higher temporal dynamics or seasonal variations cannot be evaluated due to missing hydro-chemical data. In response, we did not evaluate NSE or NSE-log for nitrate and suggest to improve monitoring schemes of N solutes in the catchment. For example, the sampling of single rain storm events as proposed by Seibert and Beven (2009) for discharge and nitrate could be helpful to examine the behaviour of the predicted loads in short temporal resolution and under various flow conditions.

4   Conclusions

In this study we compare the performance of two SWAT models regarding their accuracy, prediction uncertainty and ability to reflect the governing runoff generation processes in
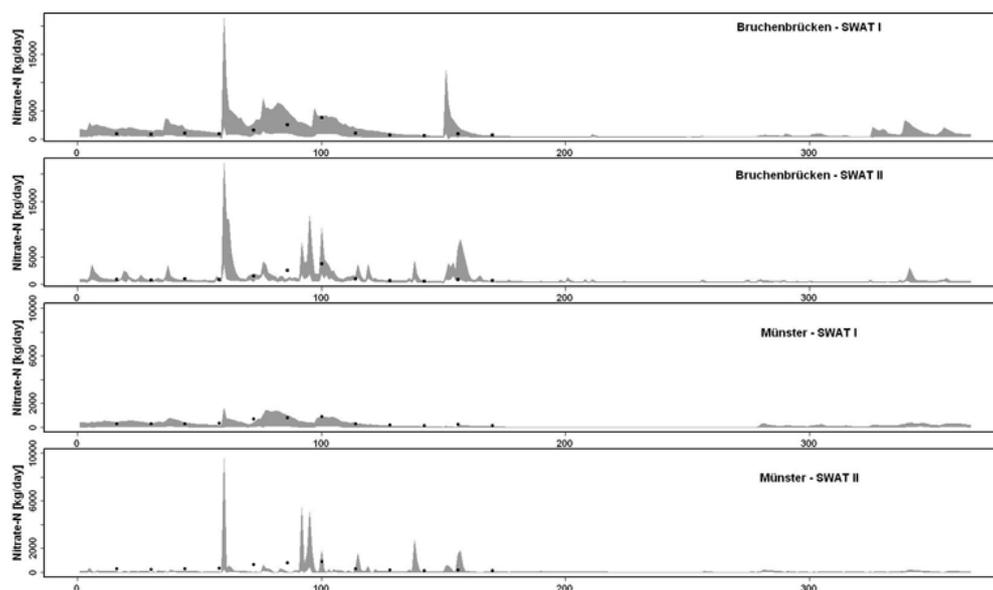
Fig. 6. Span of simulated (shaded grey) and observed (black dots) nitrate loads for the gauging stations Bruchenbrücken and Münster for the year 2008 and the parameter ensembles of SWAT I and SWAT II.

the watershed under study. These features are essential for a model to be considered well calibrated and therefore applicable for scenario analysis. The basic SWAT version (SWAT I) fails to reproduce the expected major hydrological processes in the catchment mainly in the low mountain ranges, leading to high prediction uncertainties for simulated discharge and nitrate loads. Moreover this model version is associated with larger uncertainties, reflected by the high number of acceptable parameter sets and the resulting wide variation of predicted discharge and nitrate loads. The second model version (SWAT II) was extended with expert knowledge regarding the landscape characteristics. This knowledge was gained by an intensive spatial analysis of additional soil and geological data that are not required for a SWAT model setup. The inclusion of additional information leads to a satisfactorily reproduction of measured discharge and nitrogen loads with a relatively low degree of uncertainty. The model is also able to better resemble the major landscape characteristics and the associated hydrological processes. The constraint number of acceptable parameter sets additionally expresses a low predictive uncertainty of the model. According to the above mentioned criteria SWAT II can be considered as a "well-calibrated" model and is therefore more suitable for scenario analysis than SWAT I. With the study we could demonstrate that expert knowledge about the governing hydrologic processes helped to refine model structures and hence reduced predictive uncertainty of model outputs.

References

Arnold, J. G., Srinivasan, R., Muttiah, R. S., and Williams, J. R.: Large area hydrologic modeling an assessment, Part I: model development, J. Am. Water Resour. Assoc., 34, 73–89, 1998.

Bekele, E. G. and Nicklow, J. W.: Multi-objective automatic calibration of SWAT using NSGA-II, J. Hydrol., 341, 165–176, 2007.

Beven, K. J.: Rainfall-runoff modelling – The Primer, New York, 2001.

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

Beven, K. and Freer, J.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modeling of complex environmental systems using the GLUE methodology, J. Hydrol., 249, 11–29, 2001.

Eckhardt, K., Haverkamp, V, Fohrer, N., and Frede, H.-G.: SWAT-G, a version of SWAT99.2 modified for application to low mountain range catchments, Phys. Chem. Earth, 27, 641–644, 2002.

Fohrer, N., Möller, D., and Steiner, N.: An interdisciplinary model approach to evaluate the effects of land use change, Phys. Chem. Earth, 27, 655–662, 2002.

Gupta, V. H., Beven, K. J. and Wagener, T.: Model Calibration and Uncertainty Estimation, in: Encyclopedia of Hydrological Sciences, edited by: Anderson, M. G., Chichester, doi:10.1002/0470848944.hsa138, 2005.

Hergesell, M.: GIS-based modelling of regional groundwater recharge in Hesse, Germany – Hydrologie in Hessen, 1, Wiesbaden, 2003.

Kegel, W.: Erläuterungen zur Geologischen Karte von Hessen 1:25 000 – Blatt Nr. 5517 Cleeberg, Wiesbaden, 1979.

Kümmerle, E.: Erläuterungen zur Geologischen Karte von Hessen 1:25 000 – Blatt Nr. 5618 Friedberg, Wiesbaden, 1976.

Kümmerle, E.: Erläuterungen zur Geologischen Karte von Hessen 1:25 000 – Blatt Nr. 5518 Butzbach, Wiesbaden, 1981.

Legates, D. R. and McCabe Jr., G. J.: Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation, Water Resour. Res., 35(1), 233–241, 1999.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., and Veith, T. L.: Model evaluation guidelines for systematic quantification of accuracy in watershed simulations, Trans. ASABE, 50, 885–900, 2007.

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. User's manual. Grassland, soil and water research laboratory, Agricultural Research Service, Temple, Texas, 2005.

Rouhani, H., Willems, P., Wyseure, G., and Feyen, J.: Parameter estimation in semi-distributed hydrological catchment modeling using a multi-criteria objective function, Hydrol. Process., 21, 2998–3008, 2007.

Seibert, J. and Beven, K. J.: Gauging the ungauged basin: how many discharge measurements are needed?, Hydrol. Earth Syst. Sci., 13, 883–892, doi:10.5194/hess-13-883-2009, 2009.

Seibert, J. and McDonnell, J. J.: On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration, Water Resour. Res., 38, 1241, doi:10.1029/2001WR000978, 2002.

Vaché, K. B. and McDonnell, J. J.: A process-based rejectionist framework for evaluating catchment runoff model structure, Water Resour. Res., 42, W02409, doi:10.1029/2005WR004247, 2006.

Vaché, K. B., McDonnell, J. J., and Bolte, J.: On the use of multiple criteria for a posteriori model rejection: soft data to characterize model performance, Geophys. Res. Lett., 31, L21504, doi:10.1029/2004GL021577, 2004.

Willmott, C. J.: On the validation of models, Phys. Geogr., 2, 184–194, 1981.