Hydrology and Earth System Sciences

Assessing the model performance of an integrated hydrological and biogeochemical model for discharge and nitrate load predictions

T. Pohlert, L. Breuer, J. A. Huisman, and H.-G. Frede

Institute for Landscape Ecology and Resources Management, Justus-Liebig-University Gießen, Heinrich-Buff-Ring 26, 35392 Gießen, Germany

Abstract. In this study, we evaluate the performance of the SWAT-N model, a modified version of the widely used SWAT version, for discharge and nitrate predictions at the mesoscale Dill catchment (Germany) for a 5-year period. The underlying question is, whether the model efficiency is sufficient for scenario analysis of land-use changes on both water quantity and quality. The Shuffled Complex Evolution (SCE-UA) algorithm is used to calibrate the model for daily discharge at the catchments outlet. Model performance is assessed with a split-sampling as well as a proxy-basin test using recorded hydrographs of four additional gauges located within the catchment. The efficiency regarding nitrate load simulation is assessed without further calibration on a daily, log-daily, weekly, and monthly basis as compared to observations derived from an intensive sampling campaign conducted at the catchments outlet. A new approach is employed to test the spatial consistency of the model, where simulated longitudinal profiles of nitrate concentrations were compared with observed longitudinal profiles. It is concluded that the model efficiency of SWAT-N is sufficient for the assessment of scenarios for daily discharge predictions. SWAT-N can be employed without further calibration for nitrate load simulations on both a weekly and monthly basis with an acceptable degree of accuracy. However, the model efficiency for daily nitrate load is insufficient, which can be attributed to both data uncertainty (i.e. point-source effluents and actual farming practise) as well as structural errors. The simulated longitudinal profiles meet the observations reasonably well, which suggests that the model is spatially consistent.

1 Introduction

In the last decades the Dill catchment, which is located in Mid-Hesse, Germany, faces a decline in farming, which has

led to an increase of fallow land. Currently, fallow land contributes to about 9% of the area. This change in land-use has severe consequences for various landscape functions, which led to the foundation of the Collaborative Research Centre (*Sonderforschungsbereich 299*, SFB 299) at the Justus-Liebig-University Gießen. The scope of the SFB 299 is to develop and evaluate both economically and environmentally sustainable land-use options for peripheral regions, such as the Dill catchment.

Within the framework of the SFB 299, various models coming from the fields of economy, hydrology, soil science, and biology are used to assess the impact of potential land-uses on various landscape functions, such as food production, social welfare, habitat for plant and animal species, tourism and leisure, flood protection and water harvest (Frede and Bach, 1999). As members of the SFB 299, our aim is to evaluate potential land-use options with regard to water balance and alteration of river water quality on the landscape-scale, which can be achieved by using eco-hydrologic models. Due to the need for spatial transferability of the models developed within the scope the SFB299, the use of process-oriented models that require commonly available data and a minimum of calibration is required.

During the last 20 years, various nitrogen turn-over and transport models were developed for the plot and field scale (i.e. no spatial variability of site parameters) to simulate processes within agro-ecosystems. Historically, these models were developed by both crop scientists and soil scientists. Although both groups work on the same research object (i.e. agro-ecosystem), the focus (e.g. crop productivity, groundwater contamination) is different. Consequently, processes that occur within agro-ecosystems (e.g. crop-growth, soil water movement, nutrient cycle and transport) are modelled with a varying degree of detail depending on the focus.

Algorithms to simulate nitrogen cycle and transport are incorporated in crop-growth models that were developed by crop scientists, such as the WOFOST model (Supit et al., 1994), the CropSyst model (Stöckle et al., 2003), and models

Correspondence to: T. Pohlert
(thorsten.pohlert@agrar.uni-giessen.de)

of the CERES-family (Jones and Kiniry, 1986). Agro-ecosystem models developed by soil scientists focus on environmental risks due to farming, such as nitrate leaching into groundwater (e.g. HERMES, Kersebaum, 1995) and gaseous nitrous emissions into the atmosphere (Li et al., 1992). A few agro-ecosystem models were build to investigate both crop productivity and environmental pollution due to fertilisation on the plot-scale. Two examples are the EPIC model (Williams et al., 1984) and Crop-DNDC (Zhang et al., 2002), which both have incorporated concepts of the aforementioned models.

Although the modelling of water and nutrient fluxes at the plot scale is well-represented, up-scaling of the concepts underpinning these models to provide predictions at the landscape-scale has been identified as an area of needed research (Sivapalan et al., 2003). The concepts within eco-hydrologic models that simulate both water and nitrogen cycle on the catchment scale range from simple empirically based models to complex process-oriented models. Simple empirically based methods were developed for large-area applications such as source-apportionment methods (e.g. Grizetti et al., 2005), data-based nutrient retention models for various aquatic systems (EUROHARP-NUTRET, Kronvang et al., 2004), and GIS-aided static approaches, which include population equivalents for the estimation of nitrogen input in river systems from point-sources and diffuse emissions based on nitrogen surplus in soils, such as MONERIS (Behrendt et al., 2000) or the PolFlow model (de Wit, 2001). The main advantages of the aforementioned methods are their ease of use and their low computational demand. However, their use for scenario-analysis is questionable due to several incorporated simplifications and assumptions.

The process-oriented models for the application on large scale catchments have incorporated concepts of crop-growth and nutrient cycle models. Examples are the semi-distributed HBV-N model (Bergström, 1995; Arheimer and Brandt, 1998) and LASCAM model (Sivapalan et al., 2002; Viney et al., 2000), which both simulate water and nitrate dynamics on the land-phase within subbasins. A strategy to reflect the spatial variability of soil, land-use, and farming practise within subbasins are Hydrotopes or Hydrologic Response Units (HRU). Examples for models with this concept are the SWAT model (Arnold et al., 1998), and its spin-off – the SWIM model (Krysanova and Haberlandt, 2002; Hattermann et al., 2006).

Fully distributed models were developed in the hope that an improved representation of spatial variability of land-use and soil characteristics will lead to an improved model performance regarding to discharge and river pollution prediction. Examples for this type of models are the INCA model (Whitehead et al., 1998; Wade et al., 2002), which requires hydrologic effective rainfall calculated with an external water balance model as a driving input variable, the TNT model (Beaujouan et al., 2001) that is build on the TOPMODEL approach (Beven et al., 1995), and MIKE SHE (Refsgaard and Storm, 1995). Although fully distributed models are believed to be the most "realistic" models to represent the hydrologic system, inevitable constraints are their demands for both input data and computational power. Hence, these models can often only be employed for the simulation of small watersheds.

Since a process oriented model with little data demand is needed for the assessment of land-use options in meso-scale catchments on discharge and nitrate load within the research scope of the SFB299, we choose the semi-distributed SWAT-G model (Eckhardt et al., 2002), that is a modified version for mountainous catchments of the well known Soil and Water Assessment Tool (SWAT, Arnold et al., 1998). However, since the original conceptualisation of N-cycling in SWAT-G, which originates from the EPIC model (Williams et al., 1984) failed to predict N turnover accurately, we integrated algorithms of the detailed biogeochemical model Denitrification-Decompositon model (DNDC, Li et al., 1992, 2000) as well as the CropSyst model (Stöckle et al., 2003) for an improved N simulation in a previous study (Pohlert et al., 2006). The newly coupled model that is further denoted as SWAT-N is used within this study and extensively tested for applicability on the catchment scale.

We employed an extended approach for model testing. As Kuczera and Franks (2002) critised, many simulation studies compare simulated with observed discharge at the final outlet of a catchment only, and conclude that the employed model is "validated", although other catchment responses remained unconsidered in the model testing. In this study we employed conventional split-sampling tests as well as proxy-basins tests (Klemeš, 1986) to measure the accuracy and spatial consistency of predicted discharge (five gauges) as well as nitrate load (two gauges). Furthermore, we compare observed and simulated longitudinal nitrate concentration profiles (Grayson et al., 1997) taken under various flow conditions. This new approach allows the identification of the correct representation of spatially distributed sources of nitrate as well as the routing of nitrate along the main channel.

The specific research questions that we would like to address in this contribution are:

1. Is the SWAT-N model able to predict daily river discharge at multiple gauges within the Dill catchment at the same time?

2. Does the SWAT-N model accurately predict daily, weekly, and monthly nitrate loads at the catchments outlet without further calibration of parameters that govern the nitrogen cycle?

3. Does SWAT-N accurately predict the longitudinal profile of nitrate concentrations along the main channel and nitrate concentrations of tributaries at stable flow conditions?
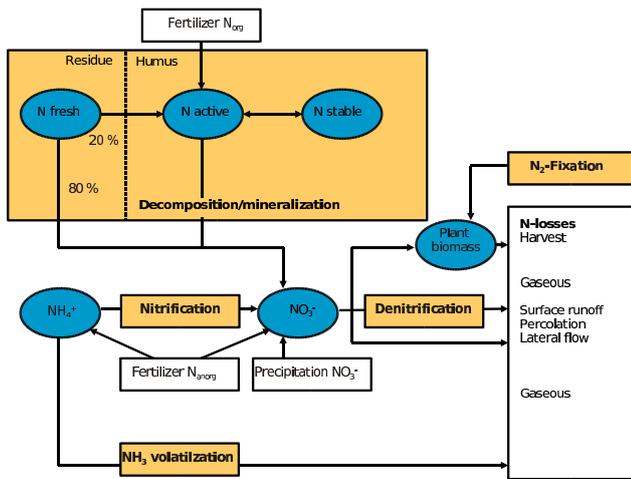
Fig. 1. Conceptualisation of the old SWAT version (modified after Pohlert et al., 2006).
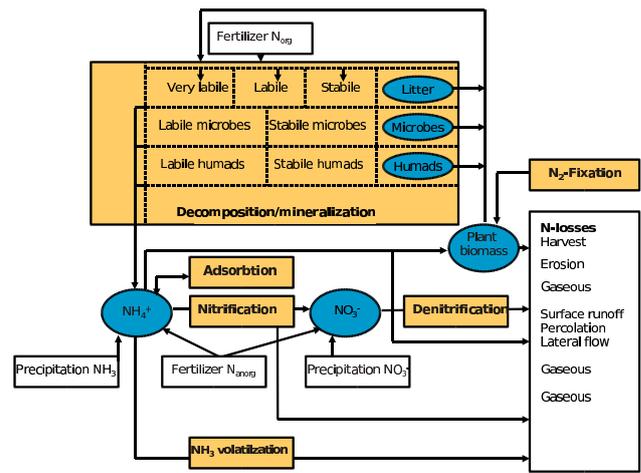

Fig. 2. Conceptualisation of the new SWAT-N version (modified after Pohlert et al., 2006).

2 Materials and methods

2.1 The SWAT-N model

The SWAT-N model (N for Nitrogen) is a modified version of the SWAT (Arnold et al., 1998) and SWAT-G model (Eckhardt et al., 2002) in which the conceptualisation of the nitrogen cycle and transport has been changed. It is a conceptually based, though process oriented semi-distributed eco-hydrologic model, which operates on a daily time step.

The smallest spatial modelling unit in SWAT-N is a hydrological response unit (HRU), which is a unique combination of land-use type and soil type derived through the overlay of a soil and a land-use map. Processes such as surface runoff estimated with the SCS curve number equation, percolation with a layered storage routing technique, lateral subsurface flow, potential evapotranspiration by the Penman-Monteith method, snow melt, crop growth, soil erosion, nitrogen and phosphorous cycle are simulated for each HRU (Arnold and Fohrer, 2005).

Based on a DEM, the catchment is delineated into several subbasins and a topological stream network is derived. The pre-processing of GIS data was done with the AVSWAT tool (Di Luzio et al., 2002). The resulting water, sediment, and nutrients fluxes are seperately summarised for each HRU within a subbasin, and allocated to the main channel of the particular subbasin. Routing of water, sediment, and solutes within the stream network is calculated with the variable storage routing method (Arnold et al., 1995). The algorithms for in-stream water quality processes such as nitrate uptake by algae are incorporated from the QUAL2E model (Brown and Barnwell, 1987).

Eckhardt et al. (2002) modified the SWAT99.2 version to yield high proportions of interflow as compared to other flow components. Interflow is the dominant flow component in mountainous regions with shallow soils, developed from periglacial layers above rocks with low permeability. The SWAT-G model (Eckhardt et al., 2002) considers the anisotropy of vertical and lateral hydraulic conductivity as well as the low permeability of the underlying bedrock, which is present in vast areas of the Dill catchment.

The original conceptualisation of the nitrogen cycle in both SWAT and SWAT-G is based on the EPIC-model (Williams et al., 1984). However, the EPIC based SWAT model failed to predict N-cycle reasonably, since high denitrification losses of up to $135 \, \mathrm{kg \, N \, ha^{-1} \, yr^{-1}}$ were simulated for single HRU's within the Dill catchment. This can be explained by the conceptualisation of denitrification in SWAT (Neitsch et al., 2002). Denitrification occurs, whenever 95% of field capacity is exceeded. Since water will only percolate in the model if soil moisture exceeds field capacity, denitrification and nitrogen leaching are two heavily competing processes in the EPIC-based SWAT versions. Under humid climatic conditions, where soils are moist for extended periods of the year, the EPIC approach leads to a rapid and complete depletion of the simulated nitrate pools in each layer due to denitrification (Pohlert et al., 2005).

To overcome the aforementioned limitations, algorithms from both the DNDC (Li et al., 1992, 2000) and the CropSyst (Stöckle et al., 2003) model were used to replace the EPIC approach for nitrogen simulation in SWAT. The new version, which includes the modifications made in SWAT-G and the algorithms for nitrogen cycling, is further denoted as SWAT-N. It simulates decomposition of organic matter from three different organic pools with first-order-kinetics, microbial immobilisation of decomposed organic nitrogen as well as mineral nitrogen, clay adsorption of ammonium, ammonium – ammonia equilibrium, ammonia volatilisation,
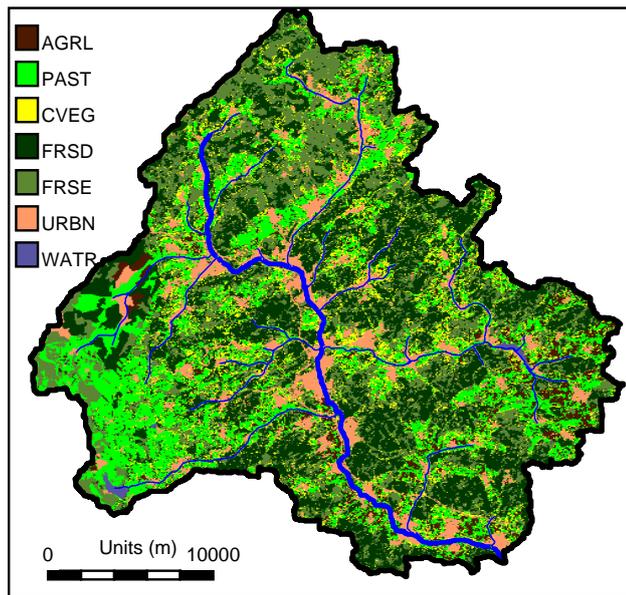
Fig. 3. Map of actual land-use in the Dill catchment based on LANDSAT TM5 data (Nöhles, 2000). AGRL means arable land, PAST is pasture, CVEG is changing vegetation, FRSD is deciduous forest, FRSE is evergreen forest, URBN means urban area, and WATR denotes surface water.

nitrification including nitrificatory nitrogen emissions to the atmosphere, and denitrification. Figures 1 and 2 illustrate the old and new conceptualisation of N-cycle within SWAT, respectively. A detailed description of the conceptualisation of the nitrogen cycle within SWAT-N as well as the sensitivity of model parameters are given by Pohlert et al. (2006).

2.2 The Dill catchment

The Dill catchment (692 km^2) is located in the mountainous area of Hesse, Germany, with elevation ranging between 155 to 675 m above sea level, and a mean slope of 8° according to the digital elevation model. Approximately 55% of the catchment is forested and both arable and pasture land contributes to about 28% of the basin based on remotely sensed data (LANDSAT TM5) taken in 1994/95 (Nöhles, 2000). The land-use distribution within the basin is shown in Fig. 3.

The geology in the south of the catchment comprises of Carboniferous clay schist and graywacke, the centre of the basins is dominated by both Devonian volcanic rocks and graywacke, and in the northern part Devonian quartzite sandstones can be found. The rocks are widely covered by periglacial deposits, which resulted from solifluction during the Pleistocene. This process lead to a considerable anisotropy of vertical and horizontal saturated hydraulic conductivity due to coarse fragments, which are aligned with their longitudinal axis in parallel to hill slopes (Eckhardt et al., 2002; Fröhlich et al., 2005). In river valleys, colluvium with larger depths developed during the Holocene. The dom-

inant soil type throughout the catchment is a shallow cambisol with underlying rocks with a low permeability (Sauer, 2002).

Annual average air temperature is 9.5°C and annual average areal precipitation is 856 mm for the period under investigation (1 July 2000 to 31 December 2004). Discharge at the river outlet Aßlar accounts for 412 mm.

2.3 Spatial input data

For this study, a 25×25 m^2 digital elevation model (DEM) and a vectorised soil type map (1:50 000) including soil-profile-data were available (BFD50, Hessisches Landesamt für Umwelt und Geologie, 2000). The soil-profile-data comprises of both textural and bulk density classes according to the German soil classification. The relevant soil physical parameters were derived from these information using the tables of AG Boden (1994). Since pH-values are not included in the BFD50 data set hitherto, they were taken from Sauer (2002) with 5.9, 4.1, 5.1, and 4.9 for arable land, forests, pasture land, and fallow land in the Dill catchment, respectively. A land-use map was compiled from multi-temporal LANDSAT TM5 images taken in 1994/95 (Nöhles, 2000). Further processing of the maps, such as the delineation of watersheds, the HRU distribution, and the interpolation of "point" rainfall at precipitation stations to each subbasion (nearest-neighbour-approach) was accomplished within the AVSWAT preprocessing-tool (Di Luzio et al., 2002).

2.4 Nitrogen input data

Data on crop rotation were taken from Lenhart (2003), where a crop rotation of winter rape – winter barley – oat was assumed to be typical for arable land in the Dill catchment. Fertilisation rates for each crop are 145 kg N ha^{-1}, 50 kg N ha^{-1}, and 50 kg N ha^{-1}, respectively. A fertilisation rate of 55 kg N ha^{-1} was specified for pasture land.

Data of average monthly point source effluents were available from three municipal sewage treatment plants as well as from one steel mill, which uses nitric acid to harden the steel. The nitrate rich wastewater of the steel mill is released into the Dietzhölze (see Fig. 4). The effluents of these four point sources comprise about 90% of total nitrate released by point sources in the entire catchment (Lenhart, 2003). The average monthly data of point source effluents were assumed to be representative for the entire simulation period.

A constant nitrate concentration in precipitation of 1.5 mg N l^{-1} was specified for the model, which was calculated from annual average precipitation and reported atmospheric nitrogen deposition (Gauger et al., 2001).

2.5 Hydrologic and meteorological data

The Dill catchment is monitored by five river gauges, which are operated by the *Hessisches Landesamt für Umwelt und Geologie* (HLUG). The location of the river gauges as well

Hydrol. Earth Syst. Sci., 11, 997–1011, 2007

www.hydrol-earth-syst-sci.net/11/997/2007/

as their drained subbasins are given by Fig. 4. For this study, daily discharge data for the period 1 July 2000–31 December 2004 were used for the assessment of the model.

Furthermore, records of daily precipitation of 12 stations from the precipitation network of the German Weather Service were used. The meteorological data set was completed with records of daily maximum and minimum temperature, wind speed, air humidity, and sun shine duration of two climatic stations.

Three of the 12 precipitation stations were abandoned during the research period. The series of these precipitation stations were extended with estimated daily precipitation based on multiple linear regression models using the records of the surrounding stations as predictor variables. Cross-validation provided root-mean-squared-errors (RMSE) between 1.63 to 2.20 mm.

The SWAT-N model will require daily global radiation records if the Penman-Monteith method is used for the calculation of evapotranspiration. Therefore sun shine duration was converted to daily global radiation. Hereby, the relation as proposed by Angstrøm (1924) was used, where the ratio of global radiation and extra-terrestrial radiation equals the proportion of hours of bright sunshine for a given location and day of the year.

2.6 Hydrochemical data

From April 2000 until December 2005, an automated sampling station (ISCO 3700, ISCO, Lincoln, NE, USA) at the gauge Aßlar (see Fig. 4) took hourly time-proportional mixed samples. These samples were collected twice a week, bulked to yield a composite daily sample, filtered through 0.45 μm polypropylene-membrane filters (Whatman purdisc, Whatman Inc., Clifton, NJ, USA), and stored at −20°C until the water samples were chemically analysed. For the period of April 2000 to April 2002, an automated photometric method (Technicon Autoanalyzer N, Technicon Industrial Systems, Tarrytown, NY, USA) was used to measure nitrate. Nitrate is hereby reduced to nitrite by Cu-II-sulphate, colourised to yield a diazo dye, and photomotrically measured at 520 nm. From May 2002 onwards, the samples were analysed with an ion chromatograph (Dionex DX-120, Dionex Corp., Sunnyvale, CA, USA) in accordance to DIN EN ISO 10304-1 (1995). To account for the systematic difference between the two chemical analysis methods, a simple linear regression model was employed to homogenise the measured nitrate concentration series. The homogenised nitrate series refer to the photometric method ($y=0.91+0.83x$, $R^2=0.72$).

Highest nitrate concentrations at Aßlar are measured in summer due to a low capacity for the dilution of point source effluents. The 10% percentile, median and 90% percentile for nitrate concentrations at gauge Aßlar are $Q_{10}=1.8$ mg N l^{-1}, $Q_{50}=3.2$ mg N l^{-1}, and $Q_{90}=5.9$ mg N l^{-1}, respectively.
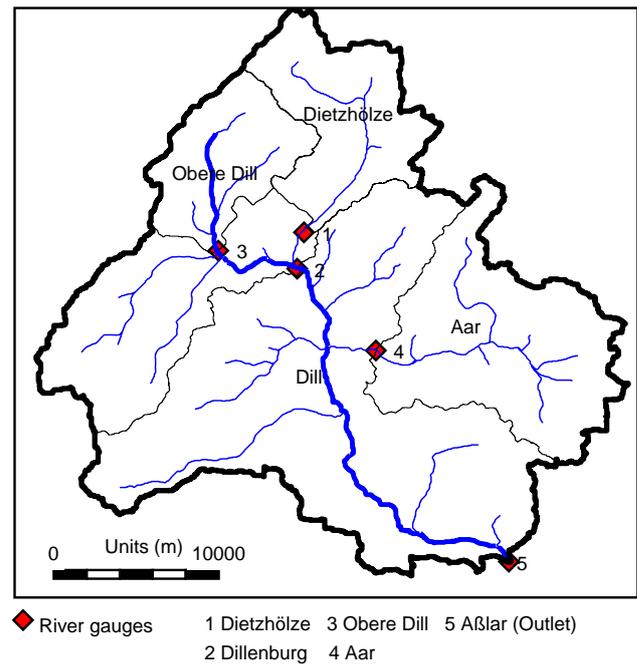

Fig. 4. Map of the Dill catchment and its gauged subbasins.

Furthermore, values of instantaneous nitrate samples, which were usually collected on a monthly basis at gauge Obere Dill (2000–2003, $n=37$), were taken from the online publication of the *Hessische Gewässergütebericht* (http://www.hlug.de). In contrast to gauge Aßlar, the highest concentrations at this up-stream gauge are present during wintertime, which indicates the absence of major point sources. The nitrate concentrations are lower as compared to Aßlar with $Q_{10}=0.4$ mg N l^{-1}, $Q_{50}=0.6$ mg N l^{-1}, and $Q_{90}=0.9$ mg N l^{-1}.

Additionally, four longitudinal profiles were sampled during two low flow conditions, one high flow and one intermediate flow condition to check the internal consistency of the model with regard to the geographical sources of the nitrate emissions. The samples were also analysed for nitrate with the ion chromatograph. The sampling sites at the main channel and at the tributaries are depicted in Fig. 5.

2.7 Calibration

With the term "model calibration" we mean the adjustment of model parameters within physically reasonable spans to yield a better match between an observed and simulated variable for a specific time period. Only the most sensitive parameters with regard to the simulated hydrologic cycle were calibrated in this study. The parameters, which control N-cycle, were manually set after an extensive sensitivity analysis (Pohlert et al., 2006).

Table 1. Calibrated parameters for the Dill catchment (1 July 2000–31 January 2003). The soil-unit 2458 refers to a shallow soil that is present on both upper slopes and the top of the hills. Soil-unit 202 is typical for lower slope positions. Spatially distributed parameters were calibrated as proposed by Eckhardt and Arnold (2001).

Parameter	Lower limit	Upper limit	Calibrated value
Baseflow recession constant (d^{-1})	0.03	0.06	0.034
Groundwater delay time for aquifer reacharge (d)	1	20	19.8
Bulk density moist, soil-unit 2458, layer 4 $(g\,cm^{-3})$	2.51	2.64	2.64
Saturated hydraulic conductivity, soil-unit 2458, layer 3 $(mm\,h^{-1})$	10	85	83.3
Saturated hydraulic conductivity, soil-unit 202, layer 3 $(mm\,h^{-1})$	1	45	44.9
Anisotropy, soil-unit 2458, layer 3	2	8	4.65
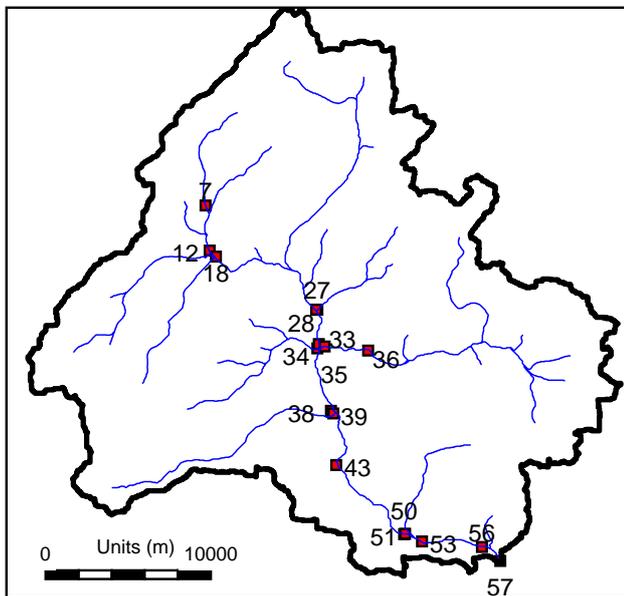
Fig. 5. Locations for snapshot-sampling along the river Dill and its tributaries.

In this study, a split-sampling-test (Klemeš, 1986) was conducted using the recorded hydrograph at gauge Aßlar for the period from 1 July 2000–31 January 2003 and the period 1 February 2003–31 December 2004 for calibration and model assessment, respectively. Additionally, a proxy-basin test (Klemeš, 1986) was employed, where simulated hydrographs were compared with observations of four interior gauges (Fig. 5) using the same parametrisation as for the gauge Aßlar. A two years warm-up period was used to initialise the pools of the model.

A single criteria calibration was conducted with the Shuffled Complex Evolution Algorithm developed at the University of Arizona (SCE-UA, Duan et al., 1992) for minimising the sum-of-squared-residuals between the simulated hydro-

graph and observations. The SCE-UA method is an optimisation algorithm, which searches for the "global optimum" within a given parameter space based on a random search strategy combined with a downhill simplex method. SCE-UA has been successfully applied for hydrologic calibration (e.g. Hogue et al., 2000; Duan et al., 1992; Eckhardt and Arnold, 2001). Table 1 provides the parameters and their upper and lower bounds, which were selected for calibration. We choose the upper and lower limits for "moist bulk density" and "saturated hydraulic conductivity" as given by AG Boden (1994). The spans for the "anisotropy factor" that denotes the ratio between lateral and vertical saturated hydraulic conductivity, the "baseflow recession constant" as well as the "groundwater delay time" were similarly set as given by Eckhardt and Arnold (2001) and Huisman et al. (2004).

Only six parameters were selected for calibration in this study, whereby four of these parameters vary in space (i.e. refer to HRU's, see also Table 1). Since the number of HRU's in the Dill model is very high (795 HRU's), an independent calibration of these parameters is neither meaningful nor feasible. Therefore, the calibration methodology as proposed by Grayson and Blöschl (2000) and Eckhardt and Arnold (2001) was applied. In this methodology a pattern of a spatially distributed parameter is calibrated by varying a single parameter that sets the values for the whole pattern according to predefined ratios. This is similar to a percent-change-calibration, where a single distributed parameter is varied for each run by multiplying each value of that parameter with one factor. Consequently, the factor only needs to be calibrated.

2.8 Assessment of model performance

Various goodness-of-fit measures were calculated, and both graphs and thematic maps were visually analysed for the assessment of model efficiency to predict both discharge and nitrate load as recommended by Legates and McCabe (1999).

The model efficiency (E) after Nash and Sutcliffe (1970) was calculated for daily observed and simulated discharge, where $E = 1 - \sum_{i=1}^{n} (O_i - P_i)^2 / \sum_{i=1}^{n} (O_i - \bar{O})^2$. O_i and P_i

Table 2. Goodness-of-fit measures for simulated and observed discharge at various gauges for the calibration period (1 July 2000–31 January 2003) and assessment period (1 February 2003–31 December 2004). E denotes model efficiency after Nash and Sutcliffe (1970), the indices d, log, w, and m denote daily data, log-transformed data, weekly averages, and monthly averages, respectively.

		Aßlar	Dillenburg	Dietzhölze	Aar	Obere Dill
Calibration	E_d	0.72	0.76	0.66	0.66	0.51
	E_{log}	0.85	0.84	0.71	0.14	0.14
	E_w	0.93	0.92	0.89	0.79	0.67
	E_m	0.96	0.97	0.91	0.80	0.74
Assessment	E_d	0.63	0.74	0.56	0.68	0.66
	E_{log}	0.80	0.74	0.60	−0.43	0.36
	E_w	0.86	0.89	0.81	0.88	0.87
	E_m	0.87	0.92	0.84	0.90	0.94

are observed and predicted values, respectively, and $\bar{O}$ is the observed average. E ranges between 1 and $-\infty$, where $E=1$ denotes a "perfect" model fit and $E=0$ means that the value for average observed discharge is as good a predictor as the hydrologic model.

A disadvantage of E is the fact that due to the squaring of residuals, E is highly sensitive to mis-predicted flood peaks (Legates and McCabe, 1999), and is almost insensitive to mis-matches during low flows. Since low flows are of great importance for the capacity of the river to dilute point source effluents, where relatively small residuals of predicted flows can result in large discrepancies between simulated and observed nitrate concentrations, a log-transformation of both observations and simulations was applied. The model efficiency of the log-transformed data (E_{log}) was calculated using the above given equation. The log-transformation emphasises the differences between observed and simulated low flows as well as recession curves, whereas the impact of flood peaks is less emphasised since they are flattened. Consequently, E_{log} is an aid for the interpretation, whether possible discrepancies between simulated and observed nitrate concentrations are due to erroneous water flow or N turnover predictions. E was additionally calculated for both weekly and monthly averaged discharge.

To assess the model performance with regard to nitrate predictions at the gauge Aßlar and Obere Dill, the index of agreement (D) according to Wilmott et al. (1985) was additionally calculated, with

$$D=1-\sum_{i=1}^{n} \mid P_i - O_i \mid^2 / \sum_{i=1}^{n} \left(\mid P_i - \bar{O} \mid + \mid O_i - \bar{O} \mid \right)^2 \quad (1)$$

The index of agreement ranges between 0 and 1. It should be noted that the values of the measures E and D are not directly comparable, though both measures are correlated due to the "squared residuals" term, which is apparent in both numerators of the equations.

For visual inspection, comparative plots of measured and simulated nitrate concentrations for 11 sites along the main

channel and six tributaries sampled at four different dates were drawn. The post-processing tool hru2map (Pohlert, 2005) was used to re-assemble model output for HRU's to a grid map, by overlaying the subbasin map, land-use map as well as the soil map as proposed by Haverkamp et al. (2005). The site-specific visualisation of hydrologic variables aids for error identification in the model (Haverkamp et al., 2005).

3 Results and discussions

3.1 Discharge

Table 1 gives the upper and lower bounds as well as the final calibrated values for the parameters considered for the hydrologic calibration. The calibrated values are similar and consistent as compared to those reported by previous studies conducted in the Dill catchment (e.g. Huisman et al., 2004; Eckhardt et al., 2005), although different calibration periods were used in these studies. This confirms both the robustness of the SWAT code as well as the SCE-UA calibration algorithm.

The calibrated values of the parameters "bulk density moist" as well as "saturated hydraulic conductivity" for both selected soil units and layers are close to the upper limit of the given parameter span (Table 1). The phenomena that several "optimised" parameters are close to one bound of the selected parameter span using automated single or multi-objective calibration routines has been frequently reported in literature (e.g. van Griensven et al., 2002; Vrugt et al., 2003; Yapo et al., 1998; Eckhardt et al., 2005). Possible reasons for that may be: (i) an insensitivity of the specific parameter (i.e. values of the parameter within a broad range leads to similar results, though the calibration algorithm further searches and stops when the termination criteria is reached), (ii) the global optimum of the parameterisation is beyond the predefined parameter span, or (iii) an effect of model and data error compensation.
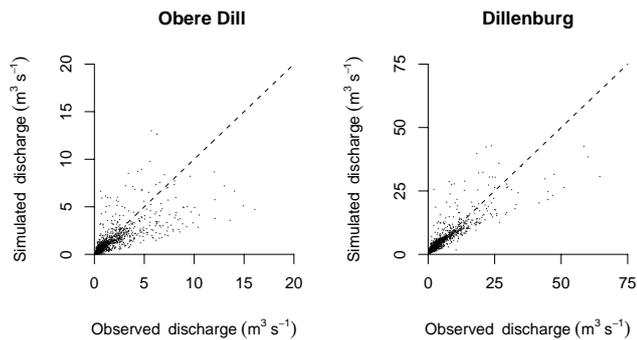
Fig. 6. Scatterplots of simulated and observed discharge at gauges Obere Dill (left) and Dillenburg (right) for the period 1 July 2000–31 December 2004. The dashed line indicates the one-to-one line.

Table 2 provides the goodness-of-fit measures for simulated and observed discharge for each gauge within the Dill catchment. As explained in Sect. 2.7, only the hydrograph at gauge Aßlar was considered for the single-criteria calibration. For the calibration period the efficiency measure E for daily data is lower at the gauges of the head catchments (Diethölze, Obere Dill, Aar) as for those, which drain larger proportions of the catchment (Aßlar, Dillenburg). The low flows at Aßlar and Dillenburg are well simulated, which is indicated by $E=0.85$ and $E=0.84$ for log-transformed values, respectively. In contrast, the predictions for low flows at the gauges Aar and Obere Dill are less accurate, as indicated by $E=0.14$ for both gauges.

The poor model performance at the head-catchments for low flows can be explained by the fact that the geology in these areas (Devonian quartzite sandstones) differ as compared to the rocks of the remaining catchment areas. A single-criteria calibration focusing on discharge at the catchments outlet, as it has been done in this study, leads to an "optimised" integral parameterisation of groundwater related parameters (i.e. "baseflow recession constant", and "groundwater delay time for aquifer recharge") for the entire catchment, which are not representative for the head-catchments.

As expected, the goodness-of-fit increases for both weekly and monthly averaged discharge data, which can be explained by the effect of smoothing. The model efficiencies slightly decline for the assessment period for all gauges except for the Obere Dill gauge.

It is noteworthy that Huisman et al. (2004) found model efficiencies of $E=0.85$ and $E=0.8$ for the calibration (1986–1988) and validation period (1989–1991) for the gauge Aßlar, respectively, which suggests a "better" model performance as compared to this study. As mentioned previously, the "optimal" parameterisation in their work was very similar to the parameterisation found in the current study. This illustrates the importance of the selected simulation period with regard to model performance, since weather patterns as well as the quality of weather data (three rain gauges went
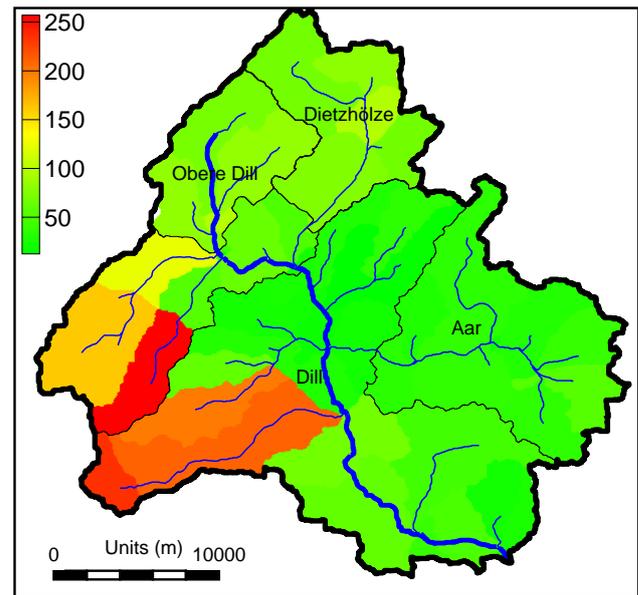

Fig. 7. Map of simulated annual average groundwater contribution to the reach of each subbasin (2000–2004).

out of service in our study period), have a strong impact on the accuracy of simulated hydrologic processes.

Scatterplots of simulated and observed discharge for the entire simulation period at the gauges Obere Dill and Dillenburg are depicted in Fig. 6. The larger scatter for daily data in the plot for gauge Obere Dill as compared to the gauge Dillenburg agrees with the model efficiencies for daily data presented in Table 2. It is obvious that SWAT-N underestimates low flows at the gauge Obere Dill, whereas low flows at the gauge Dillenburg are more accurately predicted.

The average groundwater contribution (2000–2004) to the reach for each delineated subbasin is given in Fig. 7. The simulated average annual groundwater contribution to the reach for the catchments Dietzhölze (73 mm), Aar (33 mm), Dillenburg (88 mm), Haiger (76 mm) and Aßlar (76 mm) fall within the range of reported annual groundwater contributions based on various estimation methods after Kaviany (1978). The high groundwater contribution in the south-west of the Dill catchment can be explained by the fact that this area receives more precipitation (station Driedorf 1230 mm) as compared to the aerial average precipitation of the catchment with 856 mm, and the hillslope in the south-east is lesser than throughout the catchment.

3.2 Nitrate loads

The performance of SWAT-N with regard to nitrate load predictions (1 July 2000–31 December 2004) at the gauges Aßlar and Obere Dill is summarised in Table 3. It it worth noting that the parameters, which control N-cycling within
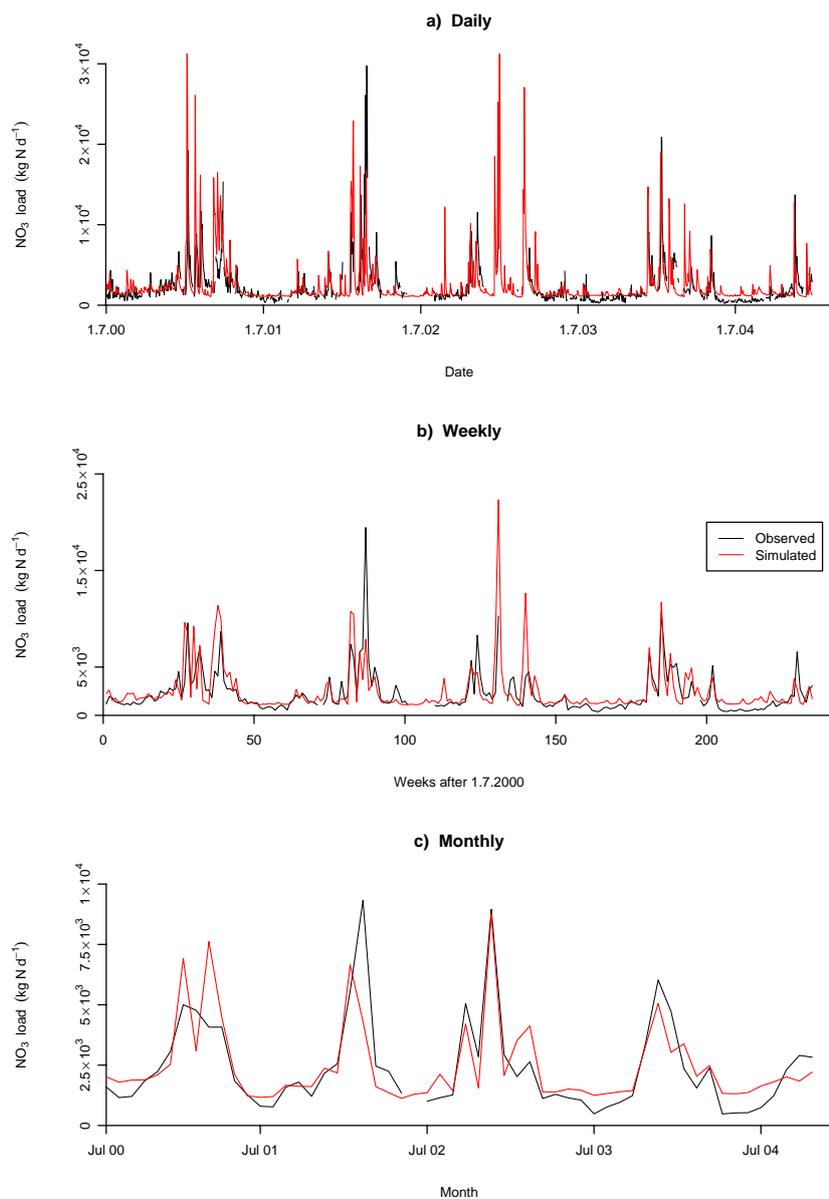
Hydrol. Earth Syst. Sci., 11, 997–1011, 2007

www.hydrol-earth-syst-sci.net/11/997/2007/

Fig. 8. Simulated and observed nitrate load at gauge Aßlar (1 July 2000–31 December 2004).

Table 3. Goodness-of-fit measures for nitrate load predictions (1 July 2000–31 December 2004) at gauges Aßlar and Obere Dill. E means model efficiency (Nash and Sutcliffe, 1970), D is the index of agreement (Wilmott et al., 1985), $\bar{N}$ is average nitrate load (kg N d^{-1}), sd is standard deviation (kg N d^{-1}), n is the number of data pairs, and the indices o and p mean observed and simulated, respectively. Only monthly instantaneous samples were available for the Obere Dill.

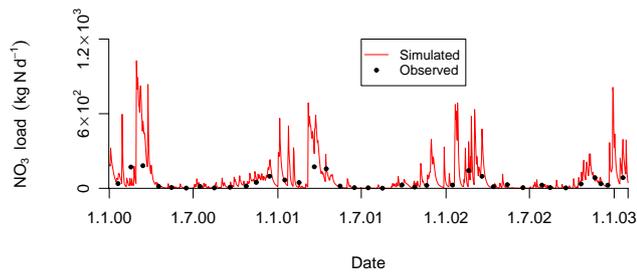| Gauge	Values	E	D	$\bar{N}_{\mathrm{o}}$	$\bar{N}_{\mathrm{p}}$	sd_{o}	sd_{p}	n
Aßlar	daily	−0.29	0.62	2247	2527	2649	3003	1249
	log daily	0.25	0.73	7.4	7.6	0.8	0.6	1249
	weekly	0.22	0.81	2231	2523	2171	2509	223
	monthly	0.65	0.89	2397	2502	1896	1690	55
Obere Dill	daily	−3.44	0.64	48	85	54	141	37
	log daily	0.07	0.87	1.4	1.2	0.6	1.0	37

Fig. 9. Simulated and observed daily nitrate load at gauge Obere Dill (1 January 2000–28 February 2003).

SWAT-N, were not calibrated. Hence, the presented results can be directly interpreted for model assessment.

The SWAT-N model overestimates average daily nitrate load ($\bar{N}_O$) at gauge Aßlar and Obere Dill by ca. $+13\%$ and $+77\%$, respectively. However, it should be noted that only 37 instantaneous nitrate samples were available for the gauge Obere Dill to calculate average daily nitrate load.

The simulated and observed nitrate load for gauge Aßlar is shown in Fig. 8. SWAT-N simulates the seasonal cycle of nitrate load accurately, which can be taken from the graphs of monthly (Fig. 8c) and weekly (Fig. 8b) loads, respectively. At week 90 and 140 SWAT-N under- and overestimates the average weekly nitrate loads, respectively (Fig. 8b). The daily variability of nitrate loads are accurately simulated (Fig. 8a, and Table 3), though the timing as well as the absolute amount of nitrate load peaks are mis-predicted.

A lower limit of simulated nitrate loads is visible in each of the graphs. This limit can be explained by the fact that only monthly average data on point source effluents were available for model input, which leads to an overestimation of nitrate loads during low flows in the model. As illustrated in Fig. 9, SWAT-N accurately predicts low and average nitrate loads for the gauge Obere Dill. However, the model overpredicts high nitrate loads on a daily basis.

The unsatisfactory model efficiency for the prediction of daily data can be partly explained by input data uncertainty. Only long-term averaged input data of point-source effluents were available for this study, but we attempted to predict nitrate loads on a daily basis. Furthermore, assumptions regarding crop rotations as well as timing and amount of fertilisation have to be made (Lenhart, 2003).

Apart from data uncertainty, the model performance in springtime points to a structural problem as well. The SWAT-N model hypothesises that frost kills the entire microbes population, which immobilises nitrogen as well as carbon during the microbial growth. As soon as temperatures rise above $0°C$, the organic nitrogen in the form of dead microbes is rapidly mineralised in the model. The impact of freeze-thaw events on rapid decomposition is known in soil science, though there is a large uncertainty in the amount of nitrifica-

tory as well as denitrificatory nitrogen losses, and the contribution to mineral nitrogen compounds (Ludwig et al., 2004). The large mineral nitrogen supply in the model due to freeze-thaw events during spring leads to high amounts of leachable nitrate. The simulated wash-out of nitrate in spring is visible in both graphs (Figs. 8a and 9), though the observed spring peak is less.

The efficiency of SWAT-N predictions found in this study is similar as reported in other eco-hydrolocical simulation experiments in mesoscale catchments. Wade et al. (2002) used the INCA model to predict weekly nitrate concentrations at the river Kennet ($1200\,km^2$) in the UK for various sampling sites along the main channel, and found Nash and Sutcliffe efficiencies between <0 to 0.8. Jarvie et al. (2002) found E's between <0 to 0.5 when INCA was used for the simulation of nitrate concentrations in the river Tweed ($4400\,km^2$), UK. The model LASCAM predicted monthly loads for the Avon River ($119\,000\,km^2$) and Ellen Brook ($700\,km^2$), Australia, with efficiencies of $E=0.61$ and 0.89, respectively (Viney et al., 2000).

Other authors working with the previous SWAT-EPIC version, partly found lower agreements between simulated and observed nitrate loads as compared to this study. Grizzetti et al. (2003) found an $E=0.30$ when they used SWAT at the Vataanjoki watershed ($1680\,km^2$), Finland, to model diffuse emissions and retentions of nutrients on a monthly basis. Chaplot et al. (2004) found $R^2=0.73$ for the prediction of monthly nitrate loads in the Walnut Creek watershed ($51\,km^2$), Iowa, and Santhi et al. (2001) reported a model efficiency ($E=0.64$) for monthly nitrate loads at the Bosque River watershed ($4277\,km^2$).

In comparison to a previously conducted simulation experiment with the SWAT-EPIC model in the Dill catchment by Lenhart et al. (2003), who achieved a model efficiency for monthly nitrate load of $E=0.31$, the modifications implemented in SWAT-N improved the model efficiency. Lenhart et al. (2003) used annual estimated point-source effluents derived from population equivalents to calculate a point-source background concentration of nitrate by division with simulated discharge. This background concentration was then subtracted from observed nitrate concentrations at gauge Aßlar, and multiplied with simulated discharge to yield nitrate loads for reference. Hence, Lenhart et al. (2003) did not account for any simulations error regarding discharge predictions, whereas both discharge prediction as well as nitrate load prediction reflecting monthly point-source inputs were evaluated in this study.

3.3 Longitudinal profile of nitrate concentrations

Figure 10 illustrates observed and simulated nitrate concentrations along the main channel Dill as well as its tributaries at four different dates. The sampling sites are given in Fig. 5. Two sampling campaigns were conducted during low flow phases on 10 September 2003 and 15 September 2004. The
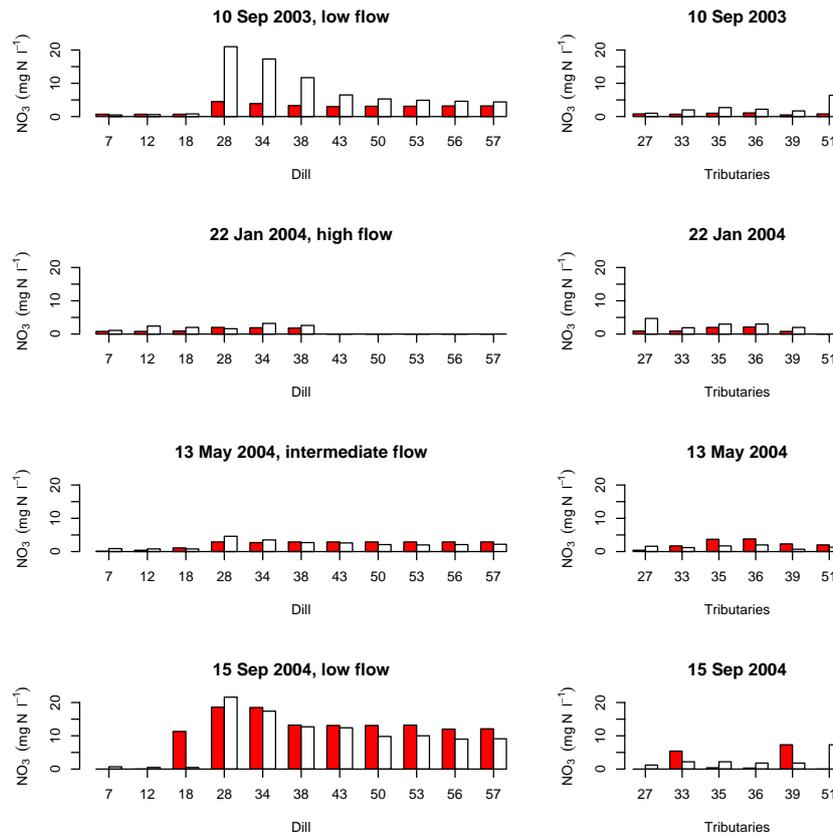
Fig. 10. Comparison of observed (white) and simulated (red) nitrate concentrations at various sampling locations along the river channel Dill (left) and at various tributaries (right) on four different dates. The sampling sites are given in Fig. 5.

white columns of the top graph indicate that observed nitrate concentrations are lowest at the Obere Dill (sites 7, 12, 18), strongly increase at site 28, which is located below the Dietzhölze tributary, and decrease along the main channel until the gauge Aßlar (site 57). The high nitrate concentrations at site 28 can be explained by the nitrate rich wastewater released by the steel mill, which is located at the mouth of the Dietzhölze.

The sampling campaign on 22 January 2004 and the campaign on 13 May 2004 are referred to high flow and intermediate flow conditions, respectively. Although the nitrate profile peaks at site 28 in May 2004 during intermediate flow conditions, the clear impact of point source effluents on nitrate concentrations along the longitudinal profile disappears due to the higher dilution capacity of the stream during intermediate and high flow periods.

SWAT-N simulates the longitudinal profiles for the dates 22 January 2004 (D=0.74, R^2=0.57), 13 May 2004 (D=0.70, R^2=0.30), and 15 September 2004 (D=0.90, R^2=0.66) accurately. However, SWAT-N over-predicts nitrate concentration for site 18 on 15 September 2004, which can be attributed to incorrect input data regarding point-source effluents, because only long-term average

monthly data were available for the municipal sewage treatment plant located at this site.

Although SWAT-N performs well at the main channel Dill, the predictions at the tributaries are less accurate (e.g. 15 September 2004). Furthermore, SWAT-N fails to predict the longitudinal profile on 10 September 2003 (D=0.52, R^2=0.64). This is because SWAT-N does not accurately simulate both discharge and nitrate emissions for the days when the longitudinal profile was sampled, although the overall model efficiency with regard to time series predictions is acceptable. It should be noted that the sampling campaign on 10 September 2003 was conducted after an exceptionally long dry period. Higher discharge predictions as compared to observed discharge under low flow conditions led to a higher dilution capacity in the model for point source effluents and, hence, lower nitrate concentrations than observed. The opposite is true for lower flow predictions as compared to observations. It is noteworthy that SWAT-N simulates nitrate concentrations in the same range as observed on 10 September 2003 with a time lag of about 10 days (data not presented).
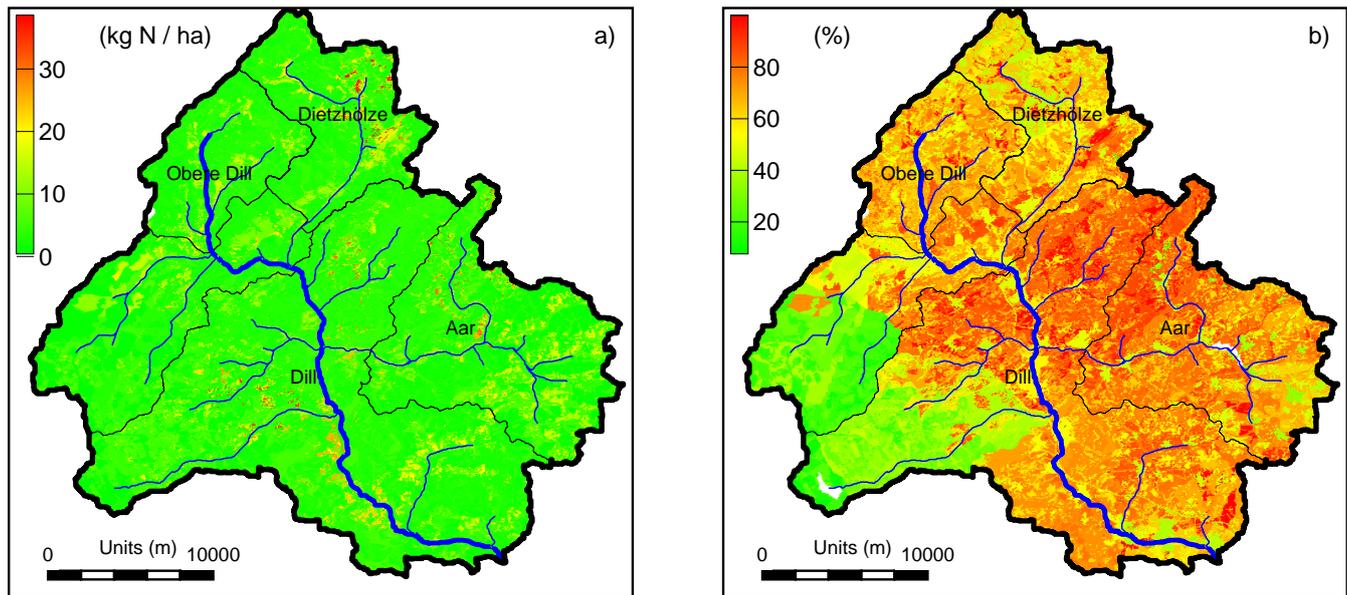
Fig. 11. Simulated average annual nitrate leaching through lateral flow (kg N ha^{-1}) for 2000 to 2004 (left). Ratio between nitrate leaching through lateral flow and total nitrate leaching (%) for the same period (right).

3.4 Geographical sources of nitrate leaching

Figure 11a shows that simulated average annual nitrate leaching through lateral flow that enters the stream is low for large areas in the Dill catchment (0 to 10 kg N ha^{-1}) for the period 2000 to 2004. At river valleys, in the north of the Dietzhölze catchment, and in the east of the Aar catchment patches with nitrate leaching between 15 to 40 kg N ha^{-1} are present according to the model. These patches correspond to pasture land as well as arable land, where fertilisation takes place.

Figure 11b illustrates the proportion of predicted lateral nitrate transport to predicted total nitrate transport to the stream. A zone with proportions of up to 90% is present in the centre of the catchment. In the north, the proportion accounts for up to 50%, and in the west of the catchments proportions of below 50% are simulated. As previously shown in Fig. 7, SWAT-N simulates highest groundwater contributions to the streams in the west of catchment, and, thus, lowest lateral flow in this area. Consequently, lateral nitrate transport is less in the west. The high proportion of laterally transported nitrate predicted by the model is consistent with the high lateral flow, which is the dominant run-off component in mountainous areas.

4 Conclusions

Performance regarding to daily discharge predictions at various sites. The model efficiency for daily discharge predictions at the catchments outlet is on an acceptable level. The selected time-period and, hence, the selected weather period for calibration has a significant impact on the model

performance but it does not significantly influence the "optimal" parametrisation found by SCE-UA.

It can be concluded from the proxy-basin test that an "optimised" parameter set for the entire catchment will lead to accurate predictions of daily discharge at up-stream gauges, if the geology and, hence, hydrologically effective parameters of the specific subbasin do not significantly differ from the entire catchment, as is the case for the Obere Dill.

Performance regarding daily, weekly, and monthly nitrate load predictions and geographical sources of nitrate emissions.

The modifications incorporated in SWAT-N improved the simulation of the nitrogen cycle as demonstrated by Pohlert et al. (2005) and as compared to a previous study with the SWAT-EPIC version by Lenhart et al. (2003).

A visual inspection of simulated nitrate load and observations at gauge Aßlar suggests a reasonable performance of the model for daily nitrate load predictions. However, the model efficiency ($E=-0.29$) for daily data indicates that the average of observations is a "better" predictor than the SWAT-N model!

The values of goodness-of-fit measures will increase to an acceptable degree, if the nitrate load predictions are averaged to a weekly or a monthly basis for the gauge Aßlar. Consequently, we conclude that SWAT-N can be used for monthly, as well as weekly predictions of nitrate load, but should be avoided for daily predictions.

The comparison of sampled longitudinal profiles of nitrate with model predictions suggests that simulated nitrate emissions from both non-point and point sources are spatially consistent. This new approach to test the internal

Hydrol. Earth Syst. Sci., 11, 997–1011, 2007

www.hydrol-earth-syst-sci.net/11/997/2007/

model performance with one-point-in-time many-points-in-space was a useful addition to the "classical" split sample techniques, where one-point-in-space and many-points-in-time are considered.

Can SWAT-N be used for scenario assessment? Based on an evaluation of the performance of any eco-hydrologic model, the modeller has to decide whether to reject the model or to apply the model for a specific purpose.

We conclude that SWAT-N can be used to predict changes on discharge caused by land-use changes on a daily basis. The accuracy of nitrate load predictions is acceptable when SWAT-N is applied on a monthly or weekly basis. The model can be used for scenario assessment, because it uses an enhanced process-description and it is a robust approach that needs no further calibration. It provides a promising tool for the joint research within the SFB 299 as well as the assessment of action plans for the implementation of the EU-Water Framework Directive (2000/60/EG, 2000). However, the use of the model should be avoided for assessing land-use impacts on nitrate loads on a daily basis.

Edited by: A. Ghadouani

References

2000/60/EG: Richtlinie 2000/60/EG des europäischen Parlaments und des Rates vom 23. Oktober 2000 zur Schaffung eines Ordnungsrahmens für Maßnahmen der Gemeinschaft im Bereich der Wasserpolitik, ABl. EG, pp. L 327/1–L 327/27, 2000.

AG Boden: Bodenkundliche Kartieranleitung, Schweizbart'sche Verlagsbuchhandlung, Hannover, 4 edn., 1994.

Angstrøm, A.: Solar and terrestrial radiation, Q. J. R. Meteorol. Soc., 50, 121–125, 1924.

Arheimer, B. and Brandt, M.: Modelling nitrogen transport and retention in the catchments of southern Sweden, Ambio, 27, 471–480, 1998.

Arnold, J., Srinivasan, R., Muttiah, R., and Williams, J.: Large area hydrologic modeling and assessment – Part 1: Model development, J. Am. Water Resour. Assoc., 34, 73–89, 1998.

Arnold, J. G. and Fohrer, N.: SWAT2000: current capabilities and research opportunities in applied watershed modelling, Hydrol. Process., 19, 563–572, doi:10.1002/hyp.5611, 2005.

Arnold, J. G., Williams, J. R., and Maidment, D.: Continuous-time water and sediment routing model for large basins, J. Hydrol. Eng., 121, 171–183, 1995.

Beaujouan, V., Durand, P., and Ruiz, L.: Modelling the effect of the spatial distribution of agricultural practices on nitrogen fluxes in rural catchments, Ecol. Model., 137, 93–105, 2001.

Behrendt, H., Huber, P., Kornmilch, M., Opitz, D., Schmoll, O., Scholz, G., and Uebe, R.: Nutrient emissions into river basins of Germany, no. 23/00 in UBA-Texte, Berlin, http://www.umweltbundesamt.de/search-public.php, 2000.

Bergström, S.: The HBV model, in: Singh (1995), pp. 443–476, 1995.

Beven, K. J., Lamb, R., Quinn, P., Romanowicz, R., and Freer, J.: TOPMODEL, in: Singh (1995), pp. 627–668, 1995.

Brown, L. C. and Barnwell, T. O.: The enhanced water quality models QUAL2E and QUAL2E-UNCAS documentation and user manual, USEPA, Athens, GA, epa document epa/600/3-87/007 edn., 1987.

Chaplot, V., Saleh, A., Jaynes, D., and Arnold, J.: Predicting water, sediment and NO3-N loads under scenarios of land-use and management practices in a flat watershed, Water Air Soil Pollut., 154, 271–293, 2004.

de Wit, M. J. M.: Nutrient fluxes at the river basin scale. I: the PolFlow model, Hydrol. Process., 15, 743–759, 2001.

Di Luzio, M., Srinivasan, R., and Arnold, J.: Integration of watershed tools and SWAT model into BASINS, J. Am. Water Resour. Assoc., 38, 1127–1141, 2002.

DIN EN ISO 10304-1: Bestimmung der gelösten Anionen Fluorid, Chlorid, Nitrit, Orthophosphat, Bromid, Nitrat und Sulfat mittels Ionenchromatographie. Teil 1: Verfahren für gering belastete Wässer, 1995.

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

Eckhardt, K. and Arnold, J.: Automatic calibration of a distributed catchment model, J. Hydrol., 251, 103–109, 2001.

Eckhardt, K., Haverkamp, S., 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.

Eckhardt, K., Fohrer, N., and Frede, H.-G.: Automatic model calibration, Hydrol. Process., 19, 651–658, doi:10.1002/hyp.5613, 2005.

Frede, H.-G. and Bach, M.: Perspectives for peripheral regions, Z. Kulturtech. Landentw., 40, 193–196, in German, 1999.

Fröhlich, H., Opp, C., and Scholten, T.: Periglacial layer or anthropogenic layer? – A small scale analysis of their spatial distribution under forest, Z. Geomorph. N. F., 139, 203–219, 2005.

Gauger, T., Köble, R., Spranger, T., Bleeker, A., and Draaijers, G.: Deposition of sulphur and nitrogen in Germany, Water Air Soil Pollut., 1, 353–373, 2001.

Grayson, R. and Blöschl, G.: Spatial modelling of catchment dynamics, in: Spatial patterns in catchment hydrology, edited by: Grayson, R. and Blöschl, G., pp. 51–81, Cambridge University press, Cambridge, 2000.

Grayson, R. B., Gippel, C. J., Finlayson, B. J., and Hart, B. T.: Catchment-wide impacts on water quality: the use of 'snapshot' sampling during stable flow, J. Hydrol., 199, 121–134, 1997.

Grizetti, B., Bouraoui, F., de Marsily, G., and Bidoglio, G.: A statistical method for source apportionment of riverine nitrogen loads, J. Hydrol., 304, 302–315, doi:10.1016/j.jhydrol.2004.07.036, 2005.

Grizzetti, B., Bouraoui, F., Granlund, K., Rekolainen, S., and Bidoglio, G.: Modelling diffuse emission and retention of nutrients in the Vantaanjoki watershed (Finland) using the SWAT model, Ecol. Model., 169, 25–38, 2003.

Hattermann, F. F., Krysanova, V., Habeck, A., and Bronstert, A.: Integrating wetlands and riparian zones in river basin modelling,

Ecol. Model., 199, 379–392, doi:10.1016/j.ecolmodel.2005.06.012, 2006.

Haverkamp, S., Fohrer, N., and Frede, H.-G.: Assessment of the effect of land use patterns on hydrologic landscape functions: a comprehensive GIS-based tool to minimize model uncertainty resulting from spatial aggregation, Hydrol. Process., 19, 715–727, doi:10.1002/hyp.5626, 2005.

Hessisches Landesamt für Umwelt und Geologie, ed.: Digitale Bodenflächendaten von Hessen 1 : 50 000 (BFD50), Version: 1, 2000.

Hogue, T. S., Sooroshian, S., Gupta, H., Holz, A., and Braatz, D.: A multistep automatic calibration scheme for river forecasting models, J. Hydromet., 1, 524–542, 2000.

Huisman, J. A., Breuer, L., Eckhardt, K., and Frede, H.-G.: Spatial consistency of automatically calibrated SWAT simulations in the Dill catchment and three of its sub-catchments, in: 2003 International SWAT Conference, edited by: Srinivasan, R., Jacobs, J. H., and Jensen, R., vol. 266 of *TWRI Technical Report*, pp. 168–173, Bari, 2004.

Jarvie, H. P., Wade, A. J., Butterfield, D., Whitehead, P. G., Tindall, C. I., Virtue, W. A., Dryburgh, W., and McGraw, A.: Modelling nitrogen dynamics and distributions in the River Tweed, Scottland: and application of the INCA model, Hydrol. Earth. Syst. Sci., 6, 433–453, 2002.

Jones, C. A. and Kiniry, J. R.: CERES-Maize: A simulation model of maize growth and development, Tech. rep., Texas A&M University Press, College Station, TX, 1986.

Kaviany, E.: Zur Hydrogeologie im Niederschlagsgebiet der Dill (Hessen), Ph.D. Thesis, Univ. of Gießen, 1978.

Kersebaum, K. C.: Application of a simple management model to simulate water and nitrogen dynamics, Ecol. Model., 81, 145–156, 1995.

Klemeš, V.: Operational testing of hydrological simulation models, Hydrol. Sci. J., 31, 13–24, 1986.

Kronvang, B., Hezlar, J., Boers, P., Jansen, J. P., Behrendt, H., Anderson, T., Arheimer, B., Venohr, M., and Hoffmann, C. C.: Nutrient retention handbook. Software manual for EUROHARP-NUTRET and scientific review on nutrient retention, no. 9-2004 in EUROHARP report, Oslo, http://www.euroharp.org, 2004.

Krysanova, V. and Haberlandt, U.: Assessment of nitrogen leaching from arable land in large river basins. Part I. Simulation experiment using a process-based model, Ecol. Model., 150, 255–275, 2002.

Kuczera, G. and Franks, S. W.: Testing hydrologic models: fortification or falsification, in: Singh and Frevert (2002), pp. 141–186, 2002.

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

Lenhart, T.: Modellierung des Stoffhaushaltes in mesoskaligen Einzugsgebieten, Ph.D. Thesis, Justus-Liebig-University Gießen, Gießen, 2003.

Lenhart, T., Fohrer, N., and Frede, H.-G.: Effects of land use changes on the nutrient balance in mesoscale catchments, Phys. Chem. Earth, 28, 1301–1309, doi:10.1016/j.pce.2003.09.006, 2003.

Li, C., Frolking, S., and Frolking, T. A.: A model of nitrous oxide evolution from soil driven by rainfall events: 1. Model structure and sensitivity, J. Geophys. Res., 97, 9759–9776, 1992.

Li, C., Aber, J., Stange, F., Butterbach-Bahl, K., and Papen, H.: A process-oriented model of N_2O and NO emissions from forest soils: 1. Model development, J. Geophys. Res., 105, 4369–4384, 2000.

Ludwig, B., Wolf, I., and Teepe, R.: Contribution of nitrification and denitrification to the emission of N_2O in a freeze-thaw event in an agricultural soil, J. Plant Nutr. Soil Sci., 167, 678–684, doi:10.1002/jpln.200421462, 2004.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models. Part 1. A discussion of principles, J. Hydrol., 10, 282–290, 1970.

Neitsch, S. L., Arnold, J. G., Kinry, J. R., and Williams, J. R.: Soil and water assessment tool. Version 2000. Theoretical documentation, USDA-ARS, Temple, TX, 2002.

Nöhles, I.: Landnutzungsklassifikation mit multitemporalen Landsat TM-Szenen in einer kleinstrukturierten Agrarregion, Ph.D. Thesis, Justus-Liebig-University Gießen, Gießen, 2000.

Pohlert, T.: Manual for hru2map, Institute for Landscape Ecology and Resources Management, Gießen, http://www.uni-giessen.de/ilr/frede/swat/swat.html, 2005.

Pohlert, T., Huisman, J. A., Breuer, L., and Frede, H.-G.: Evaluation of the soil nitrogen balance model in SWAT with lysimeter data, in: 3rd International SWAT Conference, 2005, edited by: Srinivasan, R., Jacobs, J., Day, D., and Abbaspour, K., pp. 498–508, Zurich, http://www.brc.tamus.edu/swat/index.html, visited on 4 May 2006, 2005.

Pohlert, T., Breuer, L., Huisman, J. A., and Frede, H.-G.: Integration of a detailed biogeochemical model into SWAT for improved nitrogen predictions – model development, sensitivity and uncertainty analysis, Ecol. Model., doi:10.1016/j.ecolmodel.2006.11.019, in press, 2006.

Refsgaard, J. C. and Storm, B.: MIKE SHE, in: Singh (1995), pp. 809–846, 1995.

Santhi, C., Arnold, J., Williams, J., Dugas, W., Srinivasan, R., and Hauck, L.: Validation of the SWAT model on a large river basin with point and nonpoint sources, J. Am. Water Resour. Assoc., 37, 1169–1188, 2001.

Sauer, D.: Genese, Verbreitung und Eigenschaften periglaziärer Lagen im Rheinischen Schiefergebirge – anhand von Beispielen aus Westerwald, Hunsrück und Eifel , Ph.D. Thesis, Justus-Liebig-University Gießen, Gießen, 2002.

Singh, V. P. (Ed.): Computer models of watershed hydrology, Water Resources Publications, Highlands Ranch, CO, 1995.

Singh, V. P. and Frevert, D. K. (Eds.): Mathematical models of large watershed hydrology, Water Resources Publication, Highlands Ranch, CO, 2002.

Sivapalan, M., Viney, N. R., and Zammit, C.: LASCAM: Large scale catchment model, in: Singh and Frevert (2002), chap. 14, pp. 579–648, 2002.

Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J. J., Mendiondo, E. M., O'Connell, P. E., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S., and Zehe, E.: IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences, Hydrol. Sci. J., 48, 857–880, 2003.

Stöckle, C. O., Donatelli, M., and Nelson, R.: CropSyst, a cropping systems simulation model, Europ. J. Agronomy, 18, 289–307, 2003.

Supit, I., Hooijer, A. A., and van Diepen, C. A.: System description of the WOFOST 6.0 crop simulation model implemented in CGMS, Tech. rep., Joint Research Centre EC, 1994.

van Griensven, A., Francos, A., and Bauwens, W.: Sensitivity analysis and auto-calibration of an integral dynamic model for river water quality, Water Sci. Technol., 45, 325–332, 2002.

Viney, N. R., Sivapalan, M., and Deeley, D.: A conceptual model of nutrient mobilisation and transport applicable at large catchment scales, J. Hydrol., 240, 23–44, 2000.

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

Wade, A., Durand, P., Beaujouan, V., Wessel, W., Raat, K., Whitehead, P., Butterfield, D., Rankinen, K., and Lepisto, A.: A nitrogen model for European catchments: INCA, new model structure and equations, Hydrol. Earth Syst. Sci., 6, 559–582, 2002, http://www.hydrol-earth-syst-sci.net/6/559/2002/.

Whitehead, P. G., Wilson, E. J., and Butterfield, D.: A semi-distributed Integrated Nitrogen model for multiple source assessment in Catchments (INCA): Part 1. – model structure and process equations, Sci. Total Environ., 210/211, 547–558, 1998.

Williams, J. R., Jones, C. A., and Dyke, P. T.: A modeling approach to determining the relationship between erosion and soil productivity, Trans. ASAE, 27, 129–144, 1984.

Wilmott, C. J., Ackleson, S. G., Davis, R. E., Feddema, J. J., Klink, K. M., Legates, D. R., O'Donnel, J., and Rowe, C. M.: Statistics for the evaluation and comparison of models, J. Geophys. Res., 90, 8995–9005, 1985.

Yapo, P. O., Gupta, H. V., and Sorooshian, S.: Multi-objective global optimization for hydrologic models, J. Hydrol., 204, 83–97, 1998.

Zhang, Y., Li, C., Zhou, X., and Moore, B.: A simulation model linking crop growth and soil biogeochemistry for sustainable agriculture, Ecol. Model., 151, 75–108, 2002.