



available at [www.sciencedirect.com](http://www.sciencedirect.com)



journal homepage: [www.elsevier.com/locate/jhydrol](http://www.elsevier.com/locate/jhydrol)



# Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT

Karim C. Abbaspour<sup>a,\*</sup>, Jing Yang<sup>a</sup>, Ivan Maximov<sup>a</sup>, Rosi Siber<sup>a</sup>,  
Konrad Bogner<sup>b</sup>, Johanna Mieleitner<sup>a</sup>, Juerg Zobrist<sup>a</sup>,  
Raghavan Srinivasan<sup>c</sup>

<sup>a</sup> Swiss Federal Institute of Aquatic Science and Technology (Eawag), Ueberlandstrasse 133,  
CH-8600 Duebendorf, Switzerland

<sup>b</sup> Bavarian Water Management Agency, Unit 16 – Catchment Hydrology, Flood Information and Warning Services,  
Lazarettstrasse 67, 80636 Munich, Germany

<sup>c</sup> Texas A&M University, Texas Agricultural Experimental Station, Spatial Science Lab, College Station, TX 77845, USA

Received 23 January 2006; received in revised form 11 September 2006; accepted 13 September 2006

## KEYWORDS

Watershed modelling;  
Water quality modelling;  
Calibration;  
Uncertainty analysis;  
SWAT;  
SUFI-2

**Summary** In a national effort, since 1972, the Swiss Government started the “National Long-term Monitoring of Swiss Rivers” (NADUF) program aimed at evaluating the chemical and physical states of major rivers leaving Swiss political boundaries. The established monitoring network of 19 sampling stations included locations on all major rivers of Switzerland. This study complements the monitoring program and aims to model one of the program’s catchments – Thur River basin (area 1700 km<sup>2</sup>), which is located in the north-east of Switzerland and is a direct tributary to the Rhine. The program SWAT (Soil and Water Assessment Tool) was used to simulate all related processes affecting water quantity, sediment, and nutrient loads in the catchment. The main objectives were to test the performance of SWAT and the feasibility of using this model as a simulator of flow and transport processes at a watershed scale. Model calibration and uncertainty analysis were performed with SUFI-2 (Sequential Uncertainty Fitting Ver. 2), which was interfaced with SWAT using the generic iSWAT program. Two measures were used to assess the goodness of calibration: (1) the percentage of data bracketed by the 95% prediction uncertainty calculated at the 2.5 and 97.5 percentiles of the cumulative distribution of the simulated variables, and (2) the *d-factor*, which is the ratio of the average distance between the above percentiles and the standard deviation of the corresponding measured variable. These statistics showed excellent results for discharge and nitrate and quite good results for sediment and total phosphorous. We concluded that: in watersheds similar to Thur – with good data quality and availability and relatively small model uncertainty – it is

\* Corresponding author. Tel.: +41 1 823 5359; fax: +41 1 823 5511.  
E-mail address: [abbaspour@eawag.ch](mailto:abbaspour@eawag.ch) (K.C. Abbaspour).

feasible to use SWAT as a flow and transport simulator. This is a precursor for watershed management studies.

© 2006 Elsevier B.V. All rights reserved.

## Introduction

During the last three decades, in Switzerland, as well as in other European countries, extensive and costly measures have been taken to reduce pollution input by point sources. The measures included installation of advanced waste water treatment and regulations for restricted use of phosphorous (phosphate ban in household detergents) and toxic substances. In 1972, the "National Long-term Monitoring of Swiss Rivers" (NADUF) program was initiated as a cooperative project between what is now the Swiss Federal Office for the Environment (FOEN), and the Swiss Federal Institute of Aquatic Science and Technology (Eawag) (Binderheim-Bankay et al., 2000) ([www.naduf.ch](http://www.naduf.ch)). The program monitored the chemical and physical states of Swiss rivers to evaluate the effectiveness of water protection measures undertaken by Swiss national environmental protection agencies. Several NADUF catchments and stations, including the one investigated in this study – the pre-alpine/alpine Thur River basin, which represents a low level of anthropogenic pollution – serve as reference stations for other international water monitoring programs, e.g., International Commission for the Protection of the Rhine (IKSR), the Global Environmental Monitoring System – United Nations Environmental Program/World Health Organization (GEMS-UNEP/WHO) (Jakob et al., 2002). After changing the water protection law in the late 1980s to early 1990s, the positive effects of these measures were reported at several monitoring stations. In the Swiss part of the Rhine watershed, the international target of 50% reduction in the total inputs into surface waters of P and N was achieved for P (reduction of 51%) but not for N (reduction of 23%) (Prasuhn and Sieber, 2005). Also, lead concentration decreased by 80–90% during the same time period. Furthermore, with the adoption of a new "ecologically oriented" agricultural management in 1993, which included animal friendly farming, balanced use of fertilizers, appropriate proportions of ecological compensation areas, suitable crop rotation, soil erosion protection, and measured use of pesticides – decreasing trends of nutrients in Swiss water bodies were reported as well (Jakob et al., 2002; SAEFL, 2002). The total phosphorous and nitrogen concentrations decreased significantly by 28% and 14%, respectively, from 1985 to 2001 (Prasuhn and Sieber, 2005). However, the problem of non-point source pollution still exists and is associated primarily with the agricultural applications of mineral (ammonium, nitrate) and organic (liquid and solid manure) fertilizers. It should be noted that the landuse change during the period of 1980–1995 has been quite insignificant in the Thur region as indicated by the first (from 1979 to 1985) and the second (from 1992 to 1995) landuse maps compiled by the Swiss Federal Statistical Office ([www.bfs.admin.ch](http://www.bfs.admin.ch)). The latest landuse map shown in Fig. 1 indicates a predominantly agricultural region.

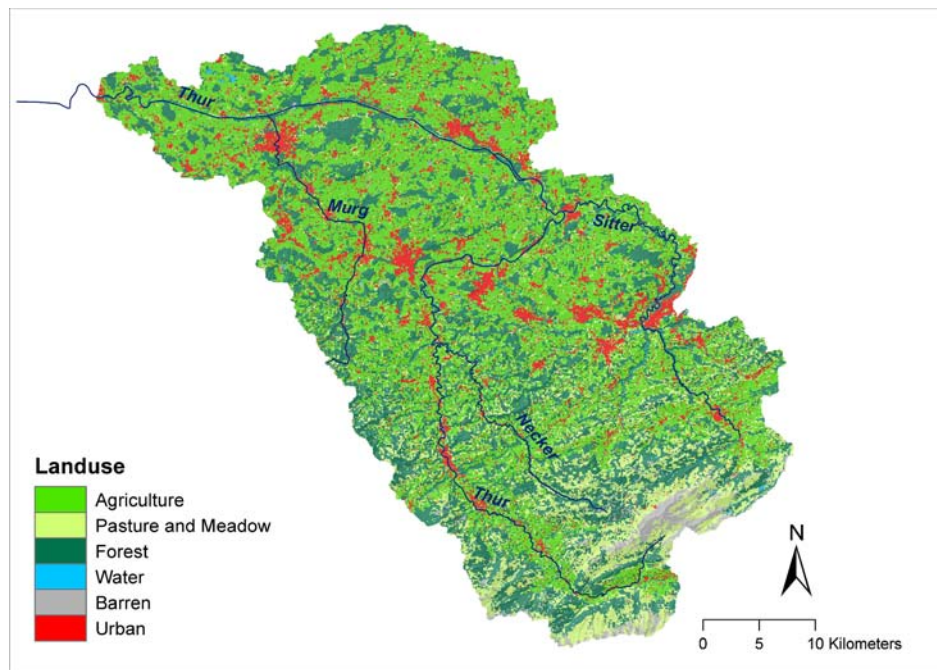
Surface runoff, especially immediately after a storm, is an important medium of transport for non-point source pol-

lution. Runoff from different landuses may be enriched with different kinds of contaminants. For example, runoff from agricultural lands is generally enriched with sediments, nutrients and pesticides, whereas runoff from actively developed urban areas contains heavy metals, hydrocarbons, chloride and other contaminants (Huber, 1993). Due to the significant reduction in the loads from point sources in the past years, the relative significance of diffuse sources of pollution in Swiss waters has increased. Presently in Switzerland, wash-out and runoff from agricultural lands contributes to a greater extent to the impairment of natural waters than it was a few decades ago (Prasuhn and Sieber, 2005).

Inverse modelling (IM) has in recent years become a very popular method for calibration (e.g., Beven and Binley, 1992; Abbaspour et al., 1997; Simunek et al., 1999; Duan et al., 2003; Gupta et al., 2003; Wang et al., 2003). IM is concerned with the problem of making inferences about physical systems from measured output variables of the model (e.g., river discharge, sediment concentration). This is attractive because direct measurement of parameters describing the physical system is time consuming, costly, tedious, and often has limited applicability. Because nearly all measurements are subject to some uncertainty, the inferences are usually statistical in nature. Furthermore, because one can only measure a limited number of (noisy) data and because physical systems are usually modelled by continuum equations, no hydrological inverse problem is really uniquely solvable. In other words, if there is a single model that fits the measurements there will be many of them. Our goal in inverse modelling is then to characterize the set of models, mainly through assigning distributions (uncertainties) to the parameters that fit the data and satisfy our presumptions as well as other prior information.

To make the parameter inferences quantitative, one must consider: (1) the error in the measured data (driving variables such as rainfall and temperature), (2) the error in the measured output variables (e.g., river discharges and sediment concentrations used for calibration), and (3) the error in the conceptual mode (inclusion of all the physics in the model that contributes significantly to the data).

The objective of this research study was to evaluate the application of a mechanistic modelling approach as a complementary technique to the monitoring program in investigating the relative impacts of different types of landuse and agricultural managements on water quality and quantity of the Thur River. A number of simulators such as SWAT (Soil Water Assessment Tool) (Arnold et al., 1998), HSPF (Hydrologic Simulation Program Fortran) (Bicknell et al., 1996), and SHETRAN (Ewen et al., 2000) could have been used in this study. Several comparisons of these models indicated similarly reasonable results in simulating discharge, phosphorous, and sediment (e.g., Singh et al., 2005; Borah and Bera, 2004). We chose SWAT because of its availability and user-friendliness in handling input data. SWAT was evaluated by performing calibration and uncertainty analysis



**Figure 1** Landuse map of the Thur watershed showing a predominantly agricultural region. Reproduced with the permission of swisstopo (BA067983).

using SUFI-2 (sequential uncertainty fitting ver. 2) algorithm (Abbaspour et al., 2004), which is a semi-automated inverse modelling procedure for a combined calibration-uncertainty analysis. The available time series data on discharge, sediment, nitrate, and total phosphorus loads at the watershed outlet as well as some constraints on sediment and nutrients from different landuses were used to perform calibration and validation studies.

## Materials and methods

### Description of the study site

The Thur watershed with an area of 1700 km<sup>2</sup> is situated in north-eastern Switzerland near the border with Germany (Fig. 2). The main river (Thur) has a total length of 127 km. The major tributaries to this river are Murg, Glatt and Sitter rivers. Mean elevation of the watershed is about 774 m above sea level and mean slope is around 7.5°. The lowest point is located at Andelfingen gauging station at 356 m above sea level and the highest point is the Saentis at 2500 m above sea level. Close to 75% of the watershed area lies below 1000 m elevation and 0.6% above 2000 m. The average daily discharge at Andelfingen is 48 m<sup>3</sup> s<sup>-1</sup> for the period of 1991–2000, with a minimum value of 3 m<sup>3</sup> s<sup>-1</sup> and a maximum value of 912 m<sup>3</sup> s<sup>-1</sup>.

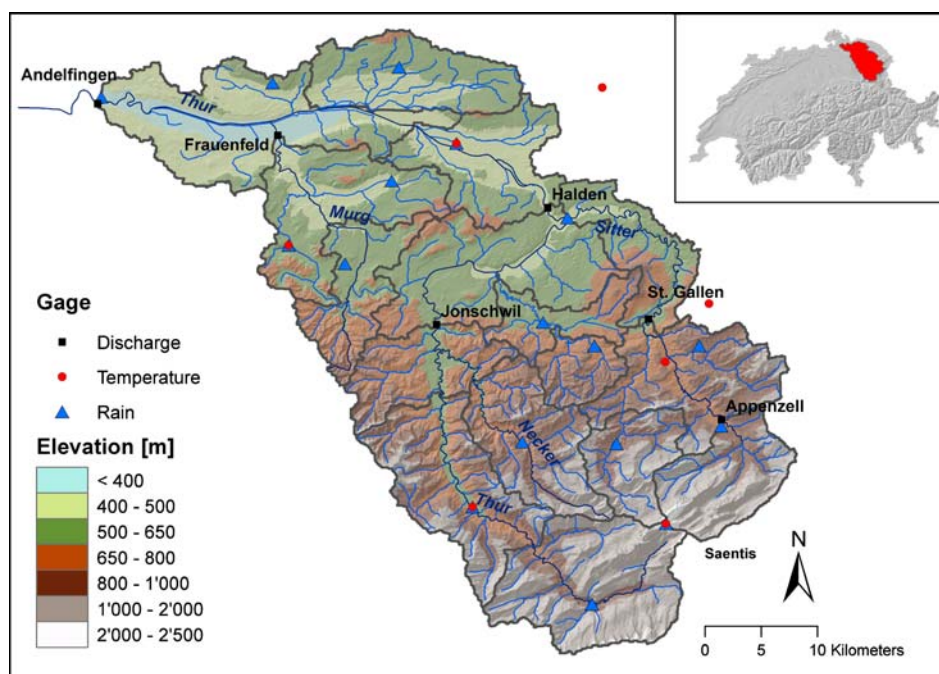
The study area has a pre-alpine/alpine climate, which is characterized by moderate winters in hilly dissected terrain area, cold winters in mountainous areas and summer seasons with relatively large annual temperature variations. Topographic effect of the terrain plays a significant role in moisture regime dynamics in the basin. The mountain climate is fairly cool and characterized by high precipitation (about 2200–2500 mm year<sup>-1</sup>), most of which falls during

the summer months. The lower (sub-mountain) portion of the watershed receives about 1000 mm year<sup>-1</sup>, and also, mostly during summer months. The mean annual precipitation for the watershed is 1460 mm year<sup>-1</sup> and the mean potential evapotranspiration estimated by Thornthwaite (1948) method is 667 mm year<sup>-1</sup>. Mean actual evapotranspiration is about 565 mm year<sup>-1</sup>, runoff 895 mm year<sup>-1</sup>. The runoff coefficient is relatively high, 0.61, and index of dryness (Budyko, 1974), i.e. the ratio of potential evapotranspiration to precipitation is relatively low, 0.46. Mean monthly temperature ranges from about 10 °C to 25 °C in the summer and from –15 °C to 7 °C during the winter. Mean annual temperature ranges from 0.02 °C at Saentis to 15.1 °C at Taenikon with an average of 7.5 °C for the catchment.

Agriculture is the dominant landuse within the area of study. Approximately 60% of the land within the basin is used for agricultural activities; these are mostly meadows for feeding cows, alpine pastures, and arable lands. Close to 30% of the total area is covered by forests, about 3% under orchards. The rest of the area is occupied by barren land, surface waters, and urban areas. Hogs and cattle are the main livestock raised in the study area.

Most of the Thur basin is underlain by conglomerates, marl incrustations and sandstone with medium to low storage capacity and rather high permeability. Groundwater is mainly found in areas with fluvio-glacial deposits of gravel and sands (Gurtz et al., 1999).

The upper (mountainous) part of the Thur River watershed is fairly uniform in terms of soil cover, i.e. covered by shallow mountain soils (about 10 cm of rooting depth in soil profile), whereas middle and lower part of the basin is more diverse and covered by more developed soils (more than 3 horizons with the rooting depth in the range of 90–140 cm).



**Figure 2** The Thur river basin with SWAT-delineated subbasins, digital elevation model, river network, and meteorological stations. Reproduced with the permission of swisstopo (BA067983).

### SWAT development and interface

SWAT (Arnold et al., 1998) is a semi-distributed, time continuous watershed simulator operating on a daily time step. It is developed for assessing the impact of management and climate on water supplies, sediment, and agricultural chemical yields in watersheds and larger river basins. The model is semi-physically based, and allows simulation of a high level of spatial detail by dividing the watershed into a large number of sub-watersheds. The major components of SWAT include hydrology, weather, erosion, plant growth, nutrients, pesticides, land management, and stream routing.

The program is provided with an interface in ArcView GIS (AVSWAT2000, Di Luzio et al., 2002) for the definition of watershed hydrologic features and storage, as well as the organization and manipulation of the related spatial and tabular data.

### Theoretical description of SWAT

The large scale spatial heterogeneity of the study area is represented by dividing the watershed into subbasins. Each subbasin is further discretised into a series of hydrologic response units (HRUs), which are unique soil-landuse combinations. Soil water content, surface runoff, nutrient cycles, sediment yield, crop growth and management practices are simulated for each HRU and then aggregated for the subbasin by a weighted average. Physical characteristics, such as slope, reach dimensions, and climatic data are considered for each subbasin. For climate, SWAT uses the data from the station nearest to the centroid of each subbasin. Calculated flow, sediment yield, and nutrient loading obtained for each subbasin are then routed through the river system. Channel routing is simulated using the variable storage or Muskingum method.

The water in each HRU in SWAT is stored in four storage volumes: snow, soil profile (0–2 m), shallow aquifer (typically 2–20 m), and deep aquifer. Surface runoff from daily rainfall is estimated using a modified SCS curve number method, which estimates the amount of runoff based on local landuse, soil type, and antecedent moisture condition. Peak runoff predictions are based on a modification of the Rational Formula (Chow et al., 1988). The watershed concentration time is estimated using Manning's formula, considering both overland and channel flow.

The soil profile is subdivided into multiple layers that support soil water processes including infiltration, evaporation, plant uptake, lateral flow, and percolation to lower layers. The soil percolation component of SWAT uses a water storage capacity technique to predict flow through each soil layer in the root zone. Downward flow occurs when field capacity of a soil layer is exceeded and the layer below is not saturated. Percolation from the bottom of the soil profile recharges the shallow aquifer. Daily average soil temperature is simulated as a function of the maximum and minimum air temperature. If the temperature in a particular layer reaches less than or equal 0 °C, no percolation is allowed from that layer. Lateral sub-surface flow in the soil profile is calculated simultaneously with percolation. Groundwater flow contribution to total stream flow is simulated by routing a shallow aquifer storage component to the stream (Arnold and Allen, 1996). A provision for estimating runoff from frozen soil is also included. Snow melts on days when the maximum temperature exceeds a prescribed value. Melted snow is treated the same as rainfall for estimating runoff and percolation.

The model computes evaporation from soils and plants separately. Potential evapotranspiration can be modelled with the Penman–Monteith (Monteith, 1965), Priestley–Taylor (Priestley and Taylor, 1972), or Hargreaves methods



(Hargreaves and Samani, 1985), depending on data availability. Potential soil water evaporation is estimated as a function of potential ET and leaf area index (area of plant leaves relative to the soil surface area). Actual soil evaporation is estimated by using exponential functions of soil depth and water content. Plant water evaporation is simulated as a linear function of potential ET, leaf area index, and root depth, and can be limited by soil water content. More detailed descriptions of the model can be found in Arnold et al. (1998).

Sediment yield in SWAT is estimated with the modified soil loss equation (MUSLE) developed by Williams and Berndt (1977). The sediment routing model consists of two components operating simultaneously: deposition and degradation. The deposition in the channel and floodplain from the sub-watershed to the watershed outlet is based on the sediment particle settling velocity. The settling velocity is determined using Stoke's law (Chow et al., 1988) and is calculated as a function of particle diameter squared. The depth of fall through a reach is the product of settling velocity and the reach travel time. The delivery ratio is estimated for each particle size as a linear function of fall velocity, travel time, and flow depth. Degradation in the channel is based on Bagnold's stream power concept (Bagnold, 1977; Williams, 1980).

Nutrient cycles are similar to those of the EPIC model (Williams et al., 1984). SWAT allows crop rotations and management practice combinations. As nutrient inputs, the model takes into consideration natural sources such as organic matter mineralization, N-fixation, wet deposition of nitrate, and anthropogenic contributions such as fertilizer applications (diffuse sources) and waste water from treatment plants (point sources). The biochemical transformations of nitrogen and phosphorus simulated by SWAT are shown in Fig. 3. The in-stream water quality modelling is based on QUAL2E (Brown and Barnwell, 1987). The QUAL2E model includes the major interactions of the nutrient cycles, algae production, and benthic oxygen demand.

## The interface program iSWAT

Discretization of the watershed into HRUs results in the generation of numerous SWAT input files in ASCII format. Each subbasin requires four files to specify the subbasin, weather, water use, and water quality parameters, while each HRU requires six files to store information on chemistry, groundwater, topography, management, routing, and soil properties. For the current project, the discretised watershed properties were stored in 809 files. Because any calibration procedure requires repeated changing of the parameter values followed by model run, the use of an interface to automate this procedure is essential. For this purpose a generic interface (iSWAT) program was developed. This program allows parameter aggregation on the basis of hydrologic group, soil, landuse, and subbasin specifications formulated as:

$$x_{-}(\text{parname})_{-}(\text{ext})_{-}(\text{hydrogrp})_{-}(\text{soltext})_{-}(\text{landuse})_{-}(\text{subbsn})$$

where  $x_{-}$  is a code to indicate the type of change to be applied to the parameter. If replaced by  $v_{-}$  it would mean the existing parameter value is to be replaced by a given value, while  $a_{-}$  would mean a given quantity should be added to the existing parameter value, and  $r_{-}$  would mean the existing parameter value should be multiplied by  $(1 + a_{-}$  given value);  $\langle \text{parname} \rangle$  is the SWAT parameter name;  $\langle \text{ext} \rangle$  is the SWAT file extension code for the file containing the parameter;  $\langle \text{hydrogrp} \rangle$  is the soil hydrological group ('A', 'B', 'C' or 'D');  $\langle \text{soltext} \rangle$  is the soil texture;  $\langle \text{landuse} \rangle$  is the landuse category; and  $\langle \text{subbsn} \rangle$  is the subbasin number, crop index, or fertilizer index. Any combination of the above factors can be used to describe a parameter identifier; hence, providing the opportunity for a detailed parameterization of the system. Omitting the identifiers  $\langle \text{hydrogrp} \rangle$ ,  $\langle \text{soltext} \rangle$ ,  $\langle \text{landuse} \rangle$ , and  $\langle \text{subbsn} \rangle$  would allow global assignment of parameters. The iSWAT interface has a generic format that allows for easy coupling with any optimization program.

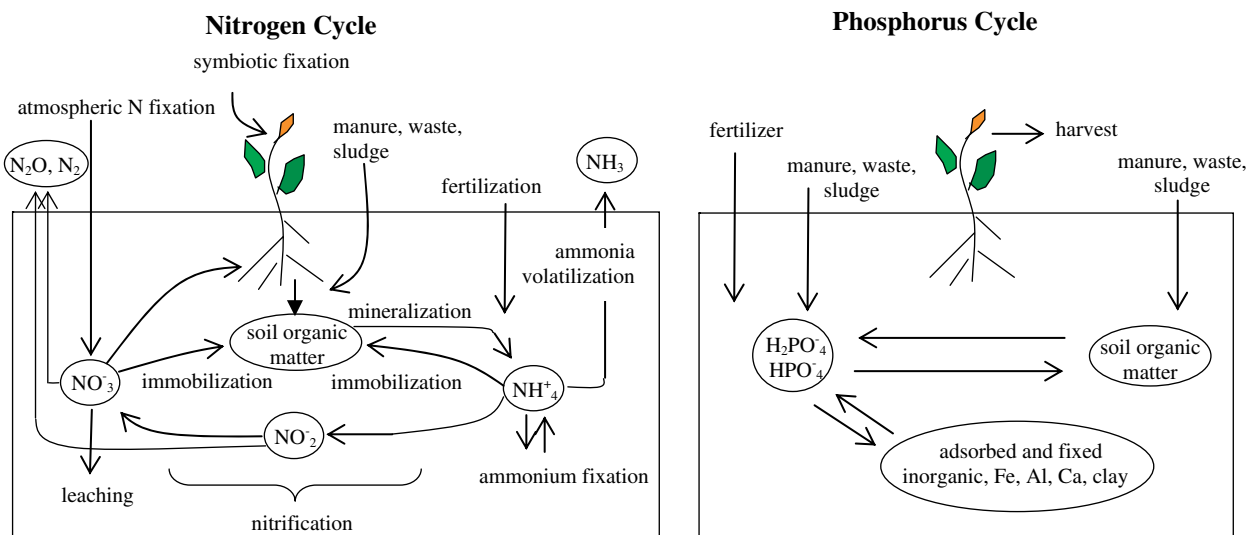


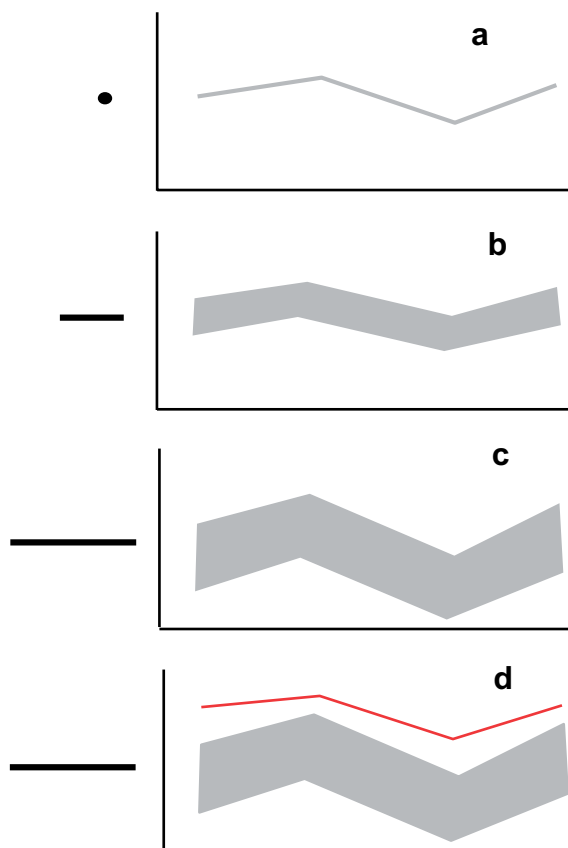
Figure 3 Schematic representations of nitrogen and phosphorus cycles (after Arnold et al., 1998).

In the current project iSWAT was coupled with SUFI-2 optimization program (Abbaspour et al., 2004). A brief description of SUFI-2 is presented below.

#### Conceptual basis of the SUFI-2 uncertainty analysis routine

In SUFI-2, uncertainty of input parameters are depicted as uniform distributions, while model output uncertainty is quantified by the 95% prediction uncertainty (95PPU) calculated at the 2.5% and 97.5% levels of the cumulative distribution of output variables obtained through Latin hypercube sampling. The concept behind the uncertainty analysis of the SUFI-2 algorithm is depicted graphically in Fig. 4. This figure illustrates that a single parameter value (shown by a point) leads to a single model response (Fig. 4a), while propagation of the uncertainty in a parameter (shown by a line) leads to the 95PPU illustrated by the shaded region in Fig. 4b. As parameter uncertainty increases (Fig. 4c), the output uncertainty also increases.

Plotting the measured data alongside the 95PPU can be quite revealing with respect to the choice of the parameter uncertainty ranges. For example, if the situation in Fig. 4d



**Figure 4** Illustration of the relationship between parameter uncertainty and prediction uncertainty. A single-valued parameter results in a single response (a), whereas an uncertain parameter results in uncertainty in prediction depicted by the 95%PPU (b and c). The larger the parameter uncertainty, the larger will be the 95PPU (c). If parameters are at their maximum physical limits and the 95PPU does not bracket the measured response, then model must be re-evaluated (d).

occurs, then parameter range must be shifted in an appropriate direction, and if the range of the parameter uncertainty already corresponds to the limits of physically meaningful values, then the problem is not one of parameter calibration and the conceptual model must be re-examined.

SUFI-2 starts by assuming a large parameter uncertainty, so that the measured data initially falls within the 95PPU, then decrease this uncertainty in steps until two rules are satisfied: (1) the 95PPU band brackets “most of the observations” and (2) the average distance between the upper (at 97.5% level) and the lower (at 2.5% level) parts of the 95PPU is “small”. Quantification of the two rules is somewhat problem dependent. If measurements are of high quality, then 80–100% of the measured data should be bracketed by the 95PPU, while a low quality data may contain many outliers and it may be sufficient to account only for 50% of the data in the 95PPU. For the second rule we require that the average distance between the upper and the lower 95PPU be smaller than the standard deviation of the measured data. This is a practical measure based on our experience. A balance between the two rules ensures bracketing most of the data within the 95PPU, while seeking the smallest possible uncertainty band. We use the above two measures to quantify the strength of calibration and accounting of the combined parameter, model, and input uncertainties.

#### SUFI-2 as an optimization algorithm

A short step-by-step description of SUFI-2 algorithm is as follows:

**Step 1.** In the first step an objective function is defined. The literature shows many different ways of formulating an objective function (e.g., Legates and McCabe, 1999; Gupta et al., 1998). Each formulation may lead to a different result; hence, the final parameter ranges are always conditioned on the form of the objective function. To overcome this problem, some studies (e.g., Yapo et al., 1998) combine different types of functions (e.g., based on root mean square error, absolute difference, logarithm of differences,  $R^2$ ,  $\chi^2$ , etc.) to yield a “multi-criteria” formulation. The use of a “multi-objective” formulation (Duan et al., 2003; Gupta et al., 1998) where different variables are included in the objective function is also important to reduce the non-uniqueness problem. The objective function used in this project is described later in model application.

**Step 2.** The second step establishes physically meaningful absolute minimum and maximum ranges for the parameters being optimized. There is no theoretical basis for excluding any one particular distribution. However, because of the lack of information, we assume that all parameters are uniformly distributed within a region bounded by minimum and maximum values. Because the absolute parameter ranges play a constraining role, they should be as large as possible, yet physically meaningful:

$$b_j : b_{j,\text{abs\_min}} \leq b_j \leq b_{j,\text{abs\_max}}, \quad j = 1, \dots, m, \quad (1)$$

where  $b_j$  is the  $j$ -th parameter and  $m$  is the number of parameters to be estimated.

**Step 3.** This step involves an optional, yet highly recommended “absolute sensitivity analysis” for all parameters in the early stages of calibration. We maintain that no automated optimization routine can replace the insight from

physical understanding and knowledge of the effects of parameters on the system response. The sensitivity analysis is carried out by keeping all parameters constant to realistic values, while varying each parameter within the range assigned in step one. For each parameter about five simulations are performed by simply dividing the absolute ranges in equal intervals and allowing the midpoint of each interval to represent that interval. Plotting results of these simulations along with the observations on the same graph gives insight into the effects of the parameters on observed signals.

**Step 4.** Initial uncertainty ranges are next assigned to parameters for the first round of Latin Hypercube sampling, i.e.

$$b_j : [b_{j,\min} \leq b_j \leq b_{j,\max}], \quad j = 1, m. \quad (2)$$

In general, the above ranges are smaller than the absolute ranges, are subjective, and are dependent upon experience. The sensitivity analysis in step 3 can provide a valuable guide for selecting appropriate ranges. Although important, these initial estimates are not crucial as they are updated and allowed to change within the absolute ranges.

**Step 5.** A Latin Hypercube (McKay et al., 1979) sampling is carried out next; leading to  $n$  parameter combinations, where  $n$  is the number of desired simulations. This number should be relatively large (approximately 500–1000). The simulation program is then run  $n$  times and the simulated output variable(s) of interest, corresponding to the measurements, are saved.

**Step 6.** As a first step in assessing the simulations, the objective function,  $g$ , is calculated.

**Step 7:** In this step a series of measures is calculated to evaluate each sampling round. First, the sensitivity matrix,  $J$ , of  $g(\mathbf{b})$  is computed using:

$$J_{ij} = \frac{\Delta g_i}{\Delta b_j}, \quad i = 1, \dots, C_2^n, \quad j = 1, \dots, m, \quad (3)$$

where  $C_2^n$  is the number of rows in the sensitivity matrix (equal to all possible combinations of two simulations), and  $j$  is the number of columns (number of parameters). Next, equivalent of a Hessian matrix,  $H$ , is calculated by following the Gauss–Newton method and neglecting the higher-order derivatives as:

$$H = J^T J. \quad (4)$$

Based on the Cramer–Rao theorem (Press et al., 1992) an estimate of the lower bound of the parameter covariance matrix,  $C$ , is calculated from:

$$C = s_g^2 (J^T J)^{-1}, \quad (5)$$

where  $s_g^2$  is the variance of the objective function values resulting from the  $n$  runs. The estimated standard deviation and 95% confidence interval of a parameter  $b_j$  is calculated from the diagonal elements of  $C$  (Press et al., 1992) from:

$$s_j = \sqrt{C_{jj}} \quad (6)$$

$$b_{j,\text{lower}} = b_j^* - t_{v,0.025} s_j, \quad (7)$$

$$b_{j,\text{upper}} = b_j^* + t_{v,0.025} s_j, \quad (8)$$

where  $b_j^*$  is the parameter  $b$  for one of the best solutions (i.e. parameters which produce the smallest value of the objective function), and  $v$  is the degrees of freedom

$(n - m)$ . Parameter correlations can then be assessed using the diagonal and off-diagonal terms of the covariance matrix as follows:

$$r_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}} \sqrt{C_{jj}}}. \quad (9)$$

It is important to note that the correlation matrix  $r$  quantifies the change in the objective function as a result of a change in parameter  $i$ , relative to changes in the other parameters  $j$ . As all parameters are allowed to change, the correlation between any two parameters is quite small.

Parameter sensitivities were calculated by calculating the following multiple regression system, which regresses the Latin hypercube generated parameters against the objective function values:

$$g = \alpha + \sum_{i=1}^m \beta_i b_i. \quad (10)$$

A  $t$ -test is then used to identify the relative significance of each parameter  $b_j$ . We emphasize that the measures of sensitivity given by [10] are different from the sensitivities calculated in step 3. The sensitivities given by [10] are estimates of the average changes in the objective function resulting from changes in each parameter, while all other parameters are changing. Therefore, [10] gives relative sensitivities based on linear approximations and, hence, only provides partial information about the sensitivity of the objective function to model parameters. Furthermore, the relative sensitivities of different parameters, as indicated by the  $t$ -test, depend on the ranges of the parameters. Therefore, the ranking of sensitive parameters may change in every iteration.

**Step 8.** In this step measures assessing the uncertainties are calculated. Because SUFI-2 is a stochastic procedure, statistics such as percent error,  $R^2$ , and Nash–Sutcliffe, which compare two signals, are not applicable. Instead, we calculate the 95% prediction uncertainties (95PPU) for all the variable(s) in the objective function. As previously mentioned, this is calculated by the 2.5th ( $X_L$ ) and 97.5th ( $X_U$ ) percentiles of the cumulative distribution of every simulated point. The goodness of fit is, therefore, assessed by the uncertainty measures calculated from the percentage of measured data bracketed by the 95PPU band, and the average distance  $\bar{d}$  between the upper and the lower 95PPU (or the degree of uncertainty) determined from:

$$\bar{d}_x = \frac{1}{k} \sum_{l=1}^k (X_U - X_L)_l, \quad (11)$$

where  $k$  is the number of observed data points. The best outcome is that 100% of the measurements are bracketed by the 95PPU, and  $\bar{d}$  is close to zero. However, because of measurement errors and model uncertainties, the ideal values will generally not be achieved. A reasonable measure for  $\bar{d}$ , based on our experience, is calculated by the  $d$ -factor expressed as:

$$d\text{-factor} = \frac{\bar{d}_x}{\sigma_x}, \quad (12)$$

where  $\sigma_X$  is the standard deviation of the measured variable  $X$ . A value of less than 1 is a desirable measure for the  $d$ -factor.

**Step 9.** Because parameter uncertainties are initially large, the value of  $\bar{d}$  tends to be quite large during the first sampling round. Hence, further sampling rounds are needed with updated parameter ranges calculated from:

$$\begin{aligned} b'_{j,\min} &= b_{j,\text{lower}} - \max\left(\frac{(b_{j,\text{lower}} - b_{j,\min})}{2}, \frac{(b_{j,\max} - b_{j,\text{upper}})}{2}\right), \\ b'_{j,\max} &= b_{j,\text{upper}} + \max\left(\frac{(b_{j,\text{lower}} - b_{j,\min})}{2}, \frac{(b_{j,\max} - b_{j,\text{upper}})}{2}\right), \end{aligned} \quad (13)$$

where  $b'$  indicate updated values. The top  $p$  solutions are used to calculate  $b_{j,\text{lower}}$  and  $b_{j,\text{upper}}$ , and the largest  $(b'_{j,\max} - b'_{j,\min})$  is used for the updated parameter range. The above criteria, while producing narrower parameter ranges for each subsequent iteration, ensure that the updated parameter ranges are always centered on the top  $p$  current best estimates, where  $p$  is a user defined value. In the final step, parameters are ranked according to their sensitivities, and highly correlated parameters are also identified. Of the highly correlated parameters, those with the smaller sensitivities should be fixed to their best estimates and removed from additional sampling rounds.

### Model parameterisation

In this study, the Thur watershed was subdivided into 16 subbasins and 149 HRUs. The watershed parameterisation and the model input were derived using the SWAT ArcView Interface (Di Luzio et al., 2002), which provides a graphical support to the disaggregation scheme and allows the construction of the model input from digital maps. The basic data sets required to develop the model input are: topography, soil, landuse and climatic data. The data used in modelling are as follows:

- (i) *Digital elevation model (DEM)*, produced by the swisstopo (grid cell: 25 m × 25 m) (DHM25@2004swisstopo) (<http://www.swisstopo.ch/en/products/digital/height/dhm25>).
- (ii) *Digital stream network*, produced by the swisstopo at a scale of 1:25,000 (Vector25@2004 swisstopo) (<http://www.swisstopo.ch/en/products/digital/landscape/vec25/vec25gwn>).
- (iii) *Soil map*, produced by the Swiss Federal Statistical Office at a scale of 1:200,000 (BEK200, BFS Geostat, CH) (<http://www.bfs.admin.ch/bfs/portal/en/index.html>), and soil data from the Kanton Zürich Office of Planning and Measurement (Bodenkarte 1:5000, ARV Kanton Zurich) (<http://www.arv.zh.ch/>).
- (iv) *Landuse map*, produced by the Swiss Federal Statistical Office (grid cell: 100 m × 100 m) (Arealstatistik 1992/1997 BFS Geostat) ([www.bfs.admin.ch](http://www.bfs.admin.ch)).
- (v) *Agricultural census data*, produced by Swiss Federal Statistical Office at a municipality level ([www.bfs.admin.ch](http://www.bfs.admin.ch)).
- (vi) *Climate data*, records from 17 precipitation, eight air temperature, five solar radiation, five relative humidity, and five wind speed gages over a period of 20

years (1980–2000) were used in the model, data were obtained from the Swiss Federal Office of Meteorology and Climatology (<http://www.meteoschweiz.ch/web/en/weather.html>).

- (vii) *Point source emissions* include monthly organic nitrogen, nitrates, nitrites and nitrogen ammonia discharges obtained from available records at several cantonal stations for a period of 10 years (1991–2000) (Kanton Zurich, St. Gallen, Thurgau). All above websites are active and have been last accessed on June 2006.

The soil map includes 17 types of soils. Soil texture, available water content, hydraulic conductivity, bulk density, and organic carbon content information were available for different layers (between two and five layers) for each soil type. A generalization of land management was thus established, considering five main classes: agriculture, range, forest-deciduous, forest-evergreen, urban, and water. In the Thur watershed wheat occupied about 40% of crop areas, corn 20–25%, sugar beat 10–13% and potato about 5%. Representative crops for each subbasin/HRU were selected according to the available cantonal agricultural management data. Wheat was chosen as a representative crop in the lower portion of the Thur watershed, whereas the middle and small fragments of the upper parts were delineated as meadows.

In our simulation, the following management scheme was adopted. Winter wheat was planted in mid-November, after a tillage operation, followed by a fertilizer application of 150 kg N ha<sup>-1</sup> and 90 kg P ha<sup>-1</sup>. Winter wheat was harvested in late August to early September and the soil was tilled again, to incorporate plant residues. An average value of 9 kg N ha<sup>-1</sup> and 1 kg P ha<sup>-1</sup> were applied on range grasses during spring.

Available point source discharge records consisted of monthly values for the period of 1991–2000 on the subbasin level. Calculated average annual loads from available reported point sources were: 210 t year<sup>-1</sup> for nitrates, 360 t year<sup>-1</sup> for total nitrogen and about 20 t year<sup>-1</sup> for total phosphorus.

An initial concentration of 1 ppm nitrogen was assumed in precipitation, but later this was calibrated to 1.3 ppm. Considering that the mean annual precipitation in the Thur basin is 1460 mm, this corresponds to an average input of about 19 kg N ha<sup>-1</sup> per year. This value is in agreement with data given in the literature. The typical range of the total nitrogen wet deposition for Switzerland is between 10 and 20 kg N ha<sup>-1</sup> per year (EAWAG News Information Bulletin, 2000).

### Model application

SWAT was calibrated based on the biweekly measured discharge, sediment, nitrate, and total phosphorous loads at the watershed outlet at Andelfingen station (Fig. 2). Water discharge was measured continuously. Concentrations of sediments (suspended solids), nitrate, and total phosphorous in the river water were determined in biweekly composite flow proportional samples. Corresponding biweekly loads were calculated as the product of biweekly average water discharge times concentration.



A constrained objective function was used to ensure correct loads were being simulated for different landuses. The objective function,  $g$ , and the constraints were formulated as follows:

Minimize :

$$g = \frac{1}{\sigma_{Q_m}^2} \sum_{i=1}^{130} (Q_m - Q_s)_i^2 + \frac{1}{\sigma_{S_m}^2} \sum_{i=1}^{130} (S_m - S_s)_i^2 + \frac{1}{\sigma_{N_m}^2} \sum_{i=1}^{130} (N_m - N_s)_i^2 + \frac{1}{\sigma_{P_m}^2} \sum_{i=1}^{130} (P_m - P_s)_i^2$$

$$\begin{aligned} \text{Subject to : } & 0.1 \leq S_{\text{Forest}} \leq 0.3 \quad 1.5 \leq S_{\text{Agricultural}} \\ & \leq 6 \quad (\text{t ha}^{-1}) \quad 2.2 \leq N_{\text{Forest}} \leq 16 \quad 19 \\ & \leq N_{\text{Agricultural}} \leq 47 \quad 15 \leq N_{\text{Pasture}} \\ & \leq 25 \quad (\text{kg N ha}^{-1}) \quad 0.02 \leq P_{\text{Forest}} \\ & \leq 0.1 \quad 0.5 \leq P_{\text{Agricultural}} \leq 2.4 \quad 0.3 \\ & \leq P_{\text{Pasture}} \leq 1.2 \quad (\text{kg P ha}^{-1}) \end{aligned} \quad (14)$$

where  $Q$  is the average biweekly discharge ( $\text{m}^3 \text{s}^{-1}$ ),  $S$  is the total biweekly sediment load in the river (t),  $N$  is the total biweekly nitrate ( $\text{NO}_3\text{-N}$ ) load in the river (kg),  $P$  is the total

**Table 1** List of SWAT's parameters that were fitted and their final calibrated values

Variable	Sensitive parameters	Final parameter value
Parameters sensitive to all four variables	– snowfall temperature, SFTMP.bsn <sup>a</sup>	–1.1 <sup>b</sup>
	– Melt factor for snow on December 21, SMFMN.bsn	0.36
	– Melt factor for snow on June 21, SMFMX.bsn	2.84
	– Snowmelt base temperature, SMTMP.bsn	2.8
	– Snowmelt temperature lag factor, TIMP.bsn	0.29
	– Baseflow alpha factor, v_ALPHA_BF.gw <sup>c</sup>	[0.17, 0.34]
	– Groundwater delay time, v_GW_DELAY.gw	0.74
	– Curve number, r_CN2.mgt	[0.085, 0.045]
	– Manning's n value for the main channel, v_CH_N2.rte	[0.0, 0.3]
	– Effective hyd. cond. in the main channel, v_CH_K2.rte	[4, 14]
	– Soil available water storage capacity, r_SOL_AWC.sol	[–0.17, 0.3]
	– Soil hydraulic conductivity, r_SOL_K.sol	[–0.19, 0.5]
	– Soil bulk density, r_SOL_BD.sol	[–0.02.7, 0.3]
	– Maximum canopy storage, v_CANMX.hru__AGRR <sup>d</sup>	2.8
– Maximum canopy storage, v_CANMX.hru__FRST	4.8	
– Maximum canopy storage, v_CANMX.hru__PAST	4.1	
Parameters sensitive to sediment only	– Sediment routing factor in main channels, v__PRF.bsn	[0.2, 0.25]
	– Channel re-entrained exponent parameter v__SPEXP.bsn	[1.35, 1.47]
	– Channel re-entrained linear parameter v__SPCON.bsn	[0.001, 0.002]
	– Channel erodability factor, v__CH_EROD.rte	[0.12, 0.14]
	– Channel cover factor, v__CH_COV.rte	[0.2, 0.25]
Parameters sensitive to total phosphorus only	– Phosphorus availability index, v__PSP.bsn	[0.5, 0.7]
	– P enrichment ratio with sediment loading, ERORGP.hru	[2.0, 4.0]
	– Rate constant for mineralization of organic P, BC4.swq	[0.3, 0.5]
	– Organic P settling rate, RS5.swq	[0.08, 0.1]
Parameters sensitive to nitrate only	– Nitrogen in rain, RCN.bsn	1.3
	– Nitrogen uptake distribution parameter, UBN.bsn	9.4
	– Concentration of $\text{NO}_3$ in groundwater, r__GWNO3.gw	[–0.3, 0.5]
	– Organic N enrichment for sediment, ERORGN.hru	2.75
	– Nitrate percolation coefficient, NPERCO.bsn	0.223
Parameters sensitive to sediment and total phosphorus	– support practice factor r__USLE_P.mgt	[–0.6, –0.1]
	– water erosion factor v__USLE_C.crp__AGRR	[0.03, 0.3]
	– water erosion factor v__USLE_C.crp__PAST,ORCD	[0.07, 0.2]
	– water erosion factor v__USLE_C.crp__FRST	[0.0, 0.1]
	– soil erodability factor, r__USLE_K.sol	[–0.19, 0.5]

<sup>a</sup> The extension (.bsn) refers to the SWAT file type where the parameter occurs.

<sup>b</sup> The fixed values indicate that a parameter was fitted and then fixed.

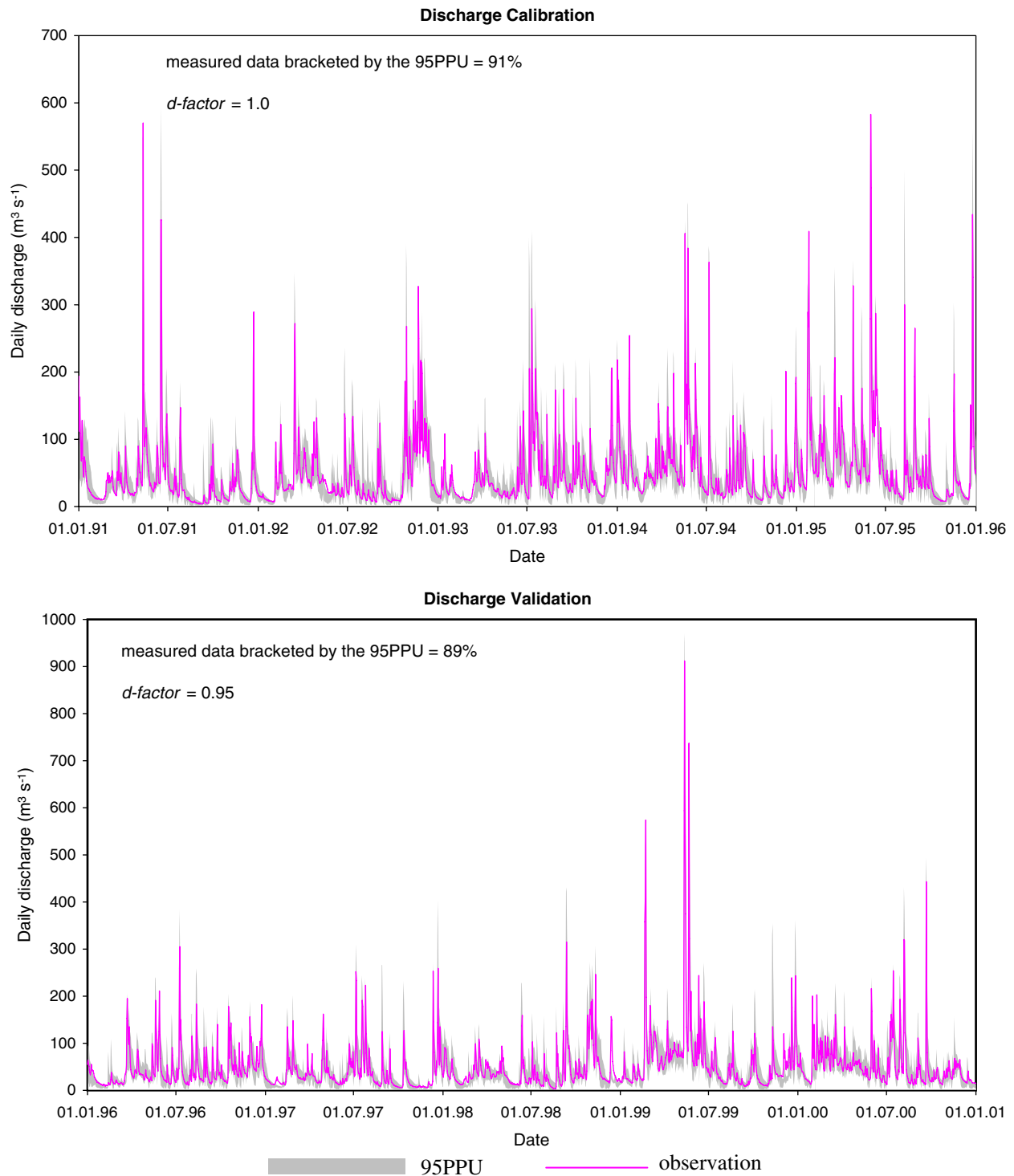
<sup>c</sup> The qualifier (v\_\_) refers to the substitution of a parameter by a value from the given range, while (r\_\_) refers to a relative change in the parameter where the current values is multiplied by 1 plus a factor in the given range.

<sup>d</sup> AGRR = agricultural, PAST = pasture, ORCD = orchard, FRST = forest.

biweekly total phosphorus load (kg),  $\sigma^2$  is the variance, and 'm' and 's' subscripts stand for measured and simulated, respectively. In the constraints,  $S_{\text{Landuse}}$  is the average annual sediment load of the landuse in the watershed ( $\text{t ha}^{-1}$ ),  $N_{\text{Landuse}}$  is the average annual nitrate load of the landuse ( $\text{kg N ha}^{-1}$ ), and  $P_{\text{Landuse}}$  is the average annual total phosphorus load of the landuse ( $\text{kg P ha}^{-1}$ ) all in the period of

1991–1995. Values of the constraints were obtained from Prasuhn et al. (1996), Prasuhn (1999), and Zobrist and Reichert (2006).

A split sample procedure was used for calibration and validation. Data from the period of 1991–1995 were used for calibration, and data from 1996–2000 were used to validate the model. It should be noted that we use the term



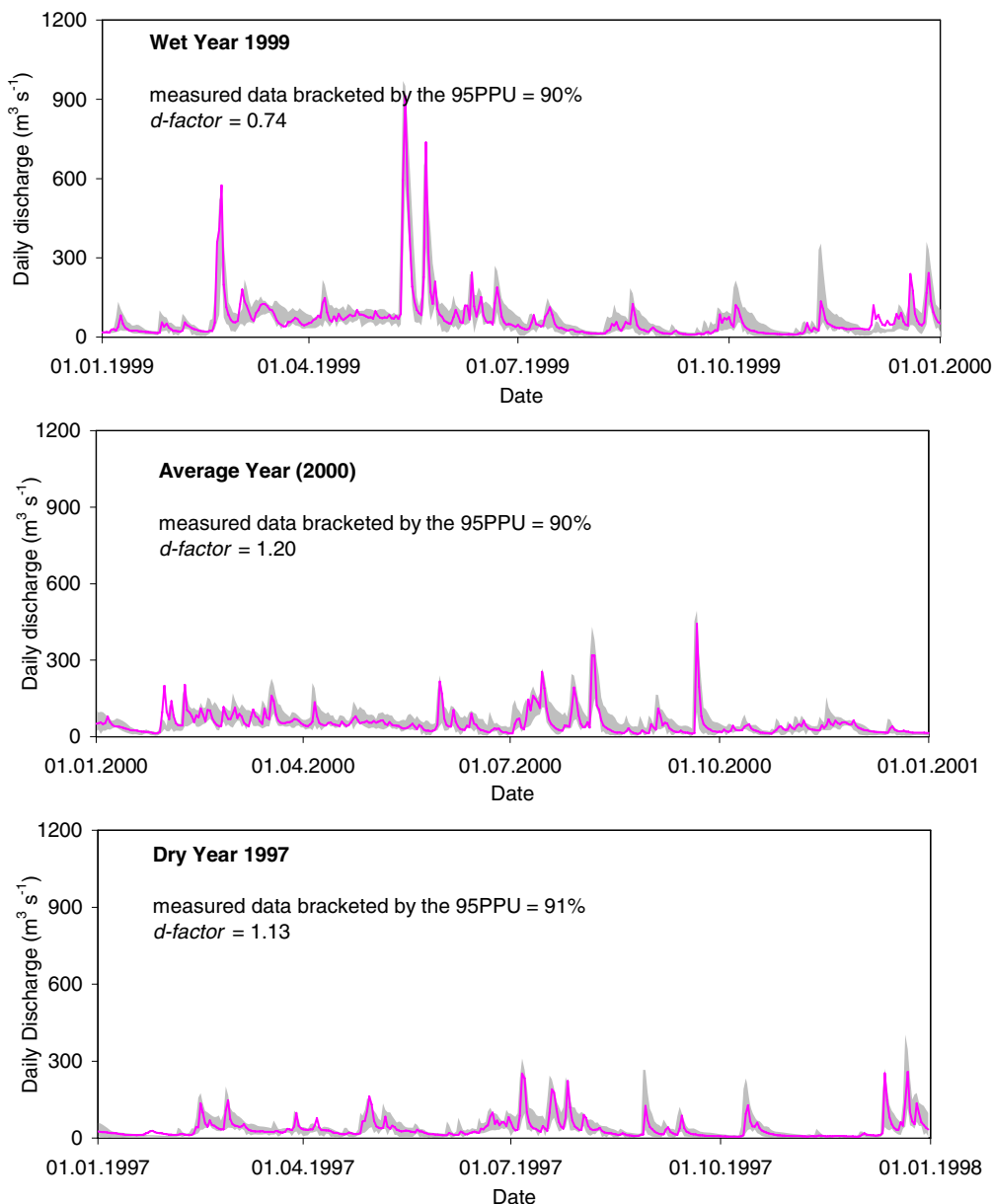
**Figure 5** Simulated and observed daily discharges at the watershed outlet. Top, discharge calibration (1991–1995); and bottom, discharge validation (1996–2000).

‘validation’ to comply with the traditional literature syntax. We are well aware that a watershed model can never be fully validated.

### Results and discussion

Calibration of models at a watershed scale is a challenging task because of the possible uncertainties that may exist in the form of process simplification, processes not accounted for by the model, and processes in the watershed that are unknown to the modeller. Some examples of the above mentioned model uncertainties are: effects of wetlands and reservoirs on hydrology and chemical transport; interaction between surface and groundwater; occurrences of landslides, and large constructions (e.g., roads, dams, tun-

nels, bridges) that could produce large amounts of sediments for a number of years affecting water quantity and quality; unknown wastewater discharges into water streams from factories and water treatment plants; and unaccounted for fertilization, irrigation and water diversions, and other activities in the river flood planes such as agricultural activities and dumping of construction materials. In a separate project in the Chaohe Basin in North China, we experienced insurmountable difficulties with simulation of sediment load in the river because of activities such as construction, material dumping, etc. The Thur watershed, however, during the period of study was relatively free of such activities; hence, model uncertainties were limited to the errors in the process simplifications alone, e.g., the simplification in the universal soil loss equation used in SWAT.



**Figure 6** Breakdown of simulated daily discharge into wet, average, and dry years.

Further sources of uncertainties in distributed models are due to inputs such as rainfall and temperature. Rainfall and temperature data are measured at local stations and regionalization of these data may introduce large errors, especially in mountainous regions. If an anomalous site is used, then runoff results may be skewed high or low. In SWAT, climate data for every subbasin is furnished by the station nearest to the centroid of the subbasin. Direct accounting of rainfall or temperature distribution error is quite difficult as information from many stations would be required. But the "elevation band" option in SWAT could to some extent alleviate this error by adjusting the temperature and rainfall to account for orographic effects of a subbasin.

Given the above possible errors, calibration and validation results of the Thur watershed could be qualified as "excellent" in this study. This indicates a good quality of the input data as well as small conceptual model errors in the dominant processes in the watershed.

We began the calibration process by initially including some 50 parameters in the SUFI-2 algorithm, but in the fifth and last iteration only 30 were found to be sensitive to discharge, sediment, nitrate, and total phosphorus. In each iteration, 1000 model calls were performed, for a total of 5000 simulations, attesting the efficiency of SUFI-2. An "absolute sensitivity analysis" (changing the parameters one at a time while keeping other parameters constant) was performed for all 50 parameters after the second iteration. The response of all four variables to changes in each parameter was plotted. This helped to identify the insensitive parameters (causing no or very small changes to variables), the sensitive parameters to all four variables, and the parameters that were sensitive to sediment only, total phosphorus only, and nitrate only (Table 1). This information proved to be quite useful in operating SUFI-2, which is a partially automated procedure requiring the analyst's attention in parameter updating at the end of each iteration.

It is worth mentioning that the results of the absolute and relative sensitivity analysis conducted for the Thur wa-

tershed may not be directly applicable to other sites. In the first procedure all parameters except one are kept constant; hence, the sensitivity of the varying parameter is conditional upon the values of all others. While in the later case, the sensitivity of parameters depends on the ranges that are assigned to the parameters. As the values of the fixed parameters or the ranges change, the sensitivity of the parameters may also change. Hence, such analysis must be performed for each site locally.

Other important considerations in the calibration of the pre-alpine/alpine Thur watershed were the corrections applied through the use of "elevation band" option in SWAT. We assigned four elevation bands with centers at 612, 115, 1691, and 2230 m for the subbasins 2, 5, 8, 9, 11, and 15 located at higher elevations (Fig. 2). The lapse rates of  $1 \text{ mm km}^{-1}$  and  $-6 \text{ }^\circ\text{C km}^{-1}$  were applied to rainfall and temperature, respectively. The use of elevation band was necessary to correct a shift in the discharge data and the overall dynamics of flow. The calibrated snow-related parameters are given in Table 1.

The results of the daily discharge simulation are shown in Fig. 5. These simulations are based on a calibration that used biweekly discharge, sediment, nitrate, and total phosphorus in the objective function. The calibration and validation statistics are also given in the figures for ease of referencing. The shaded region (95PPU), which is the simulation result, quantifies all uncertainties because it brackets a large amount of the measured data, which contains all uncertainties. The parameter ranges leading to the 95PPU are also given in Table 1. In SUFI-2, parameter uncertainty accounts for all sources of uncertainty, e.g., input uncertainty, conceptual model uncertainty, and parameter uncertainty, because disaggregation of the error into its source components is difficult, particularly in cases common to hydrology where the model is nonlinear and different sources of error may interact to produce the measured deviation (Gupta et al., 2005).

In discharge calibration, 91% of the measured data were bracketed by the 95PPU while the *d-factor* had a desired va-

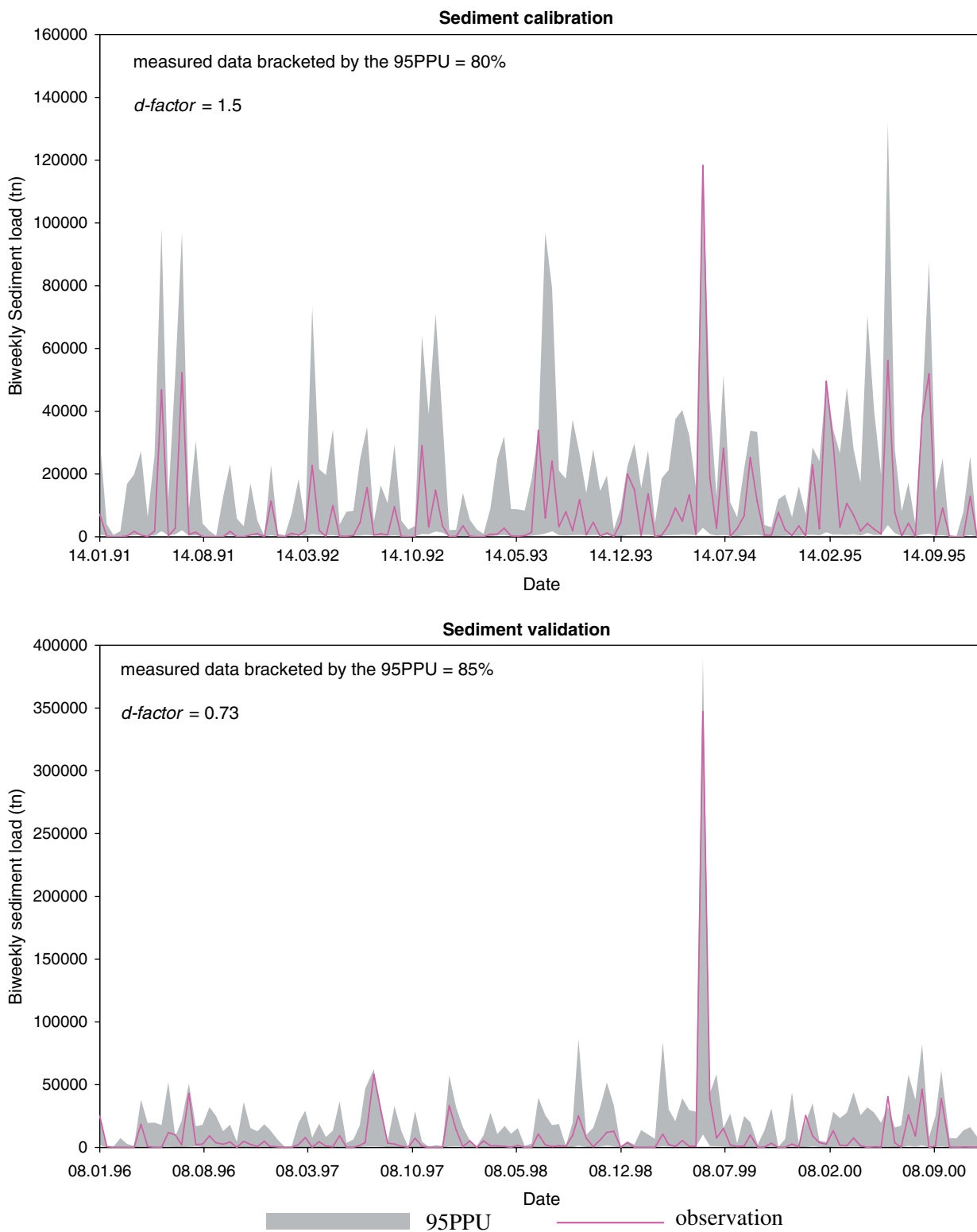
**Table 2** Break down of water fluxes and sediment and nutrient loads into different components for major landuses for a dry, a wet, and an average year

Landuse	Year	Rain <sup>a</sup> (mm)	ET (mm)	SURQ (mm)	LATQ (mm)	GWQ (mm)	WYLD (mm)	SW (mm)	PERC (mm)	SYLD (t ha <sup>-1</sup> )	TN (kg ha <sup>-1</sup> )	TP (kg ha <sup>-1</sup> )
FRST	Dry	1225	627	257	331	98	678	100	105	0.1	9	0.08
	Wet	1844	600	693	351	209	1238	118	220	0.3	21	0.15
	Average	1521	678	436	394	176	994	116	189	0.2	13	0.1
AGRR	Dry	1077	648	181	211	86	471	112	93	1.9	16	0.9
	Wet	1656	636	520	247	226	980	132	241	4.6	42	2.7
	Average	1355	716	287	264	166	708	127	177	3.7	24	1.3
PAST	Dry	1457	580	472	452	79	990	69	84	1.5	19	0.8
	Wet	2174	552	1095	449	117	1640	82	121	2.5	27	2.5
	Average	1806	621	785	513	113	1392	81	120	1.6	19	1

FRST = forest, AGRR = agricultural, PAST = summer pasture, ET = evapotranspiration, SURQ = surface runoff, LATQ = lateral flow into stream, GWQ = groundwater contribution to stream flow, WYLD = SURQ + LATQ + GWQ-LOSSES, SW = soil water, PERC = percolation below root zone (groundwater recharge), SYLD = sediment yield, TN = total nitrogen, TP = total phosphorus.

<sup>a</sup> All entries are averages for landuses occurring in different subbasins.





**Figure 7** Simulated and observed biweekly sediment loads carried by the river at the watershed outlet. Top, sediment calibration (1991–1995); and bottom, sediment validation (1996–2000).

lue of 1. The validation results were also quite excellent with 89% of the data bracketed with a  $d$ -factor equal to 0.95. To understand the flow processes during different seasons, we plotted the discharges for a wet (1999), a dry

(1997), and an average (2000) year for the validation period (Fig. 6). For these years, various water flux components are also given in Table 2. In all three years, over 90% of the observations were bracketed by the 95PPU, which is an

excellent statistic. The dry and average years, however, have slightly larger prediction uncertainties associated with them. It should be noticed that the use of the term “dry” is relative as the rainfall is still greater than 1077 mm. The fluxes in Table 2 reveal that in an average or a wet year surface runoff dominates water yield. In a dry year, however, lateral flow contribution makes up a larger part of the water yield from the landuses. Perhaps one important reason for the favourable discharge simulation of the Thur watershed is the fact that most of the rainfall in the region translates into runoff and lateral flow, and parameters dealing with the lesser understood processes such as groundwater recharge, and groundwater–river interaction were not as important as they would have been if surface runoff and lateral flow did not dominate the flow processes.

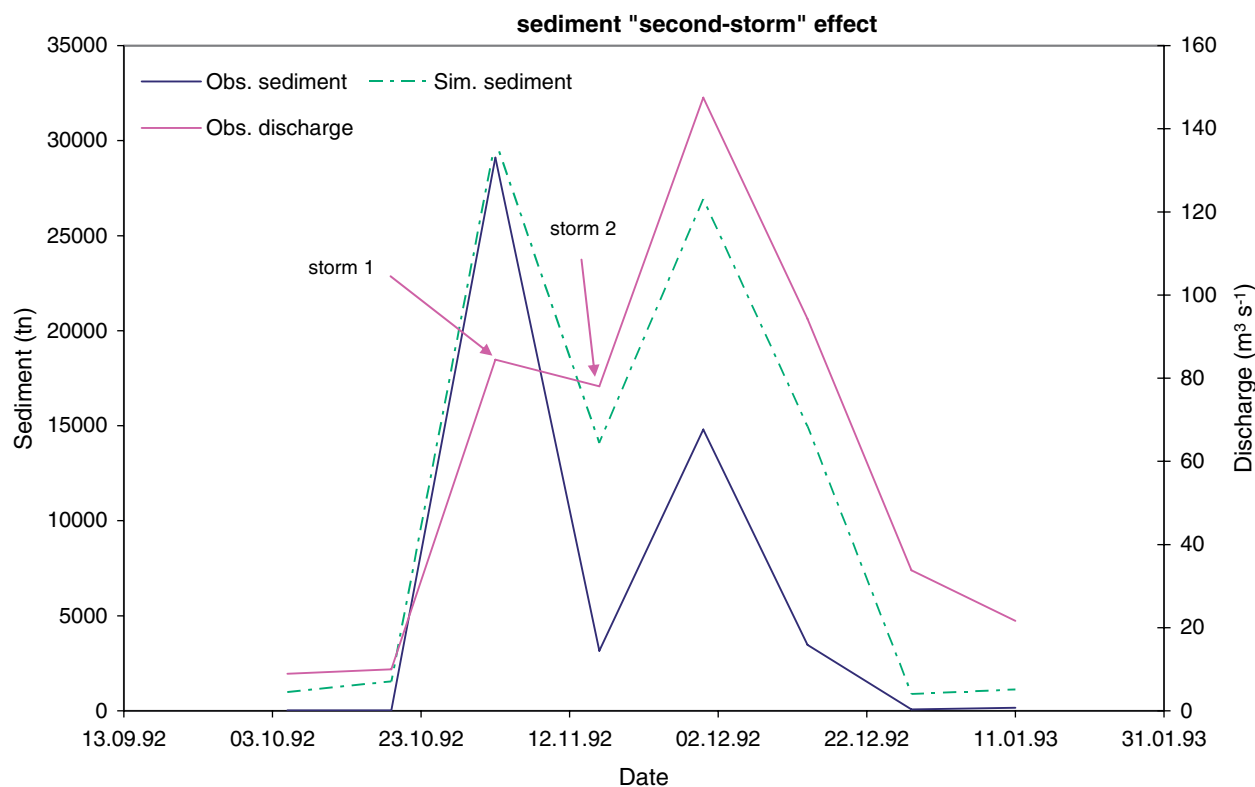
The results for biweekly sediment are shown in Fig. 7. About 80% of the data were bracketed by the 95PPU and the *d*-factor had a value of 1.5. Most of the data missing the 95PPU band were from the very small sediment loads, while all of the peaks were accounted for. The calibration and validation statistics show larger uncertainties than discharge. In the validation, 85% of the data were bracketed by the 95PPU, although the *d*-factor is small – primarily due to one large observation resulting in a large standard deviation. Removing this observation gives a larger value of 1.4 for the *d*-factor. A common problem in the prediction of particulates such as sediment and organic phosphorus is that of the “second-storm” effect. After a storm, there is less sediment to be moved, and the remaining surface layer is much more difficult to mobilize. Hence, a similar size storm, or even a bigger size second or third storm could

actually result in smaller sediment loads. The model, however, does not account for this effect as illustrated in Fig. 8. The model produces a good simulation of sediment load for the first storm, while in the second and the third storms it overestimated the load.

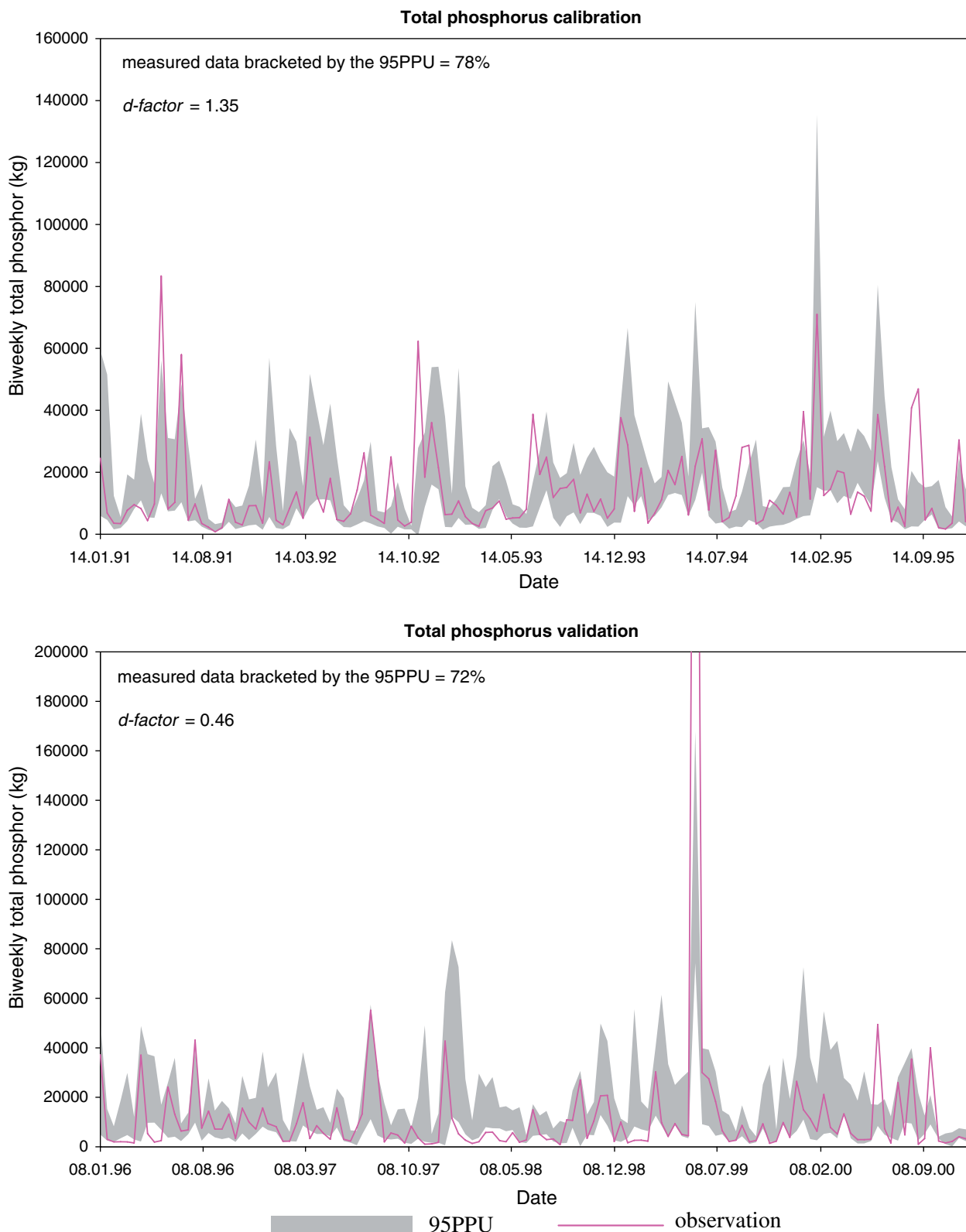
Five parameters were found to be sensitive to sediment only. These included sediment routing factor in the main channels (PRF.bsn), channel re-entrained exponent parameter (SPEXP.bsn), channel re-entrained linear parameter (SPCON.bsn), channel erodability factor (CH\_EROD.rte), and channel cover factor (CH\_COV.rte). Three other parameters were found to be sensitive to both sediment and total phosphorus. These included support practice factor (USLE\_P.mgt), water erosion factor (USLE\_C.crp), and soil erodability factor (USLE\_K.sol). Fourteen other discharge related parameters listed in Table 1 were also sensitive to sediment.

Results of the total phosphorus (TP) simulation in the river discharge are shown in Fig. 9. As a large part of TP is the organic component transported by sediment, the “second-storm” effect as described for the sediment also applies to TP. For this reason the uncertainty in TP is also large as indicated by a *d*-factor of 1.35 for calibration while bracketing only 78% of the data. As in the sediment, the validation *d*-factor would also have been quite large without the large observation occurring in the June of 1999 where a large TP load of 478,054 kg was reported.

Four parameters were found to be sensitive to TP only (Table 1). These included phosphorus availability index (PSP.bsn), P enrichment ratio with sediment loading (ERORGP.hru), rate constant for mineralization of organic P



**Figure 8** Illustration of the “second-storm” effect on sediment. After a storm there is less sediment to be carried. The model does not account for this phenomenon, hence, overestimating the observation.



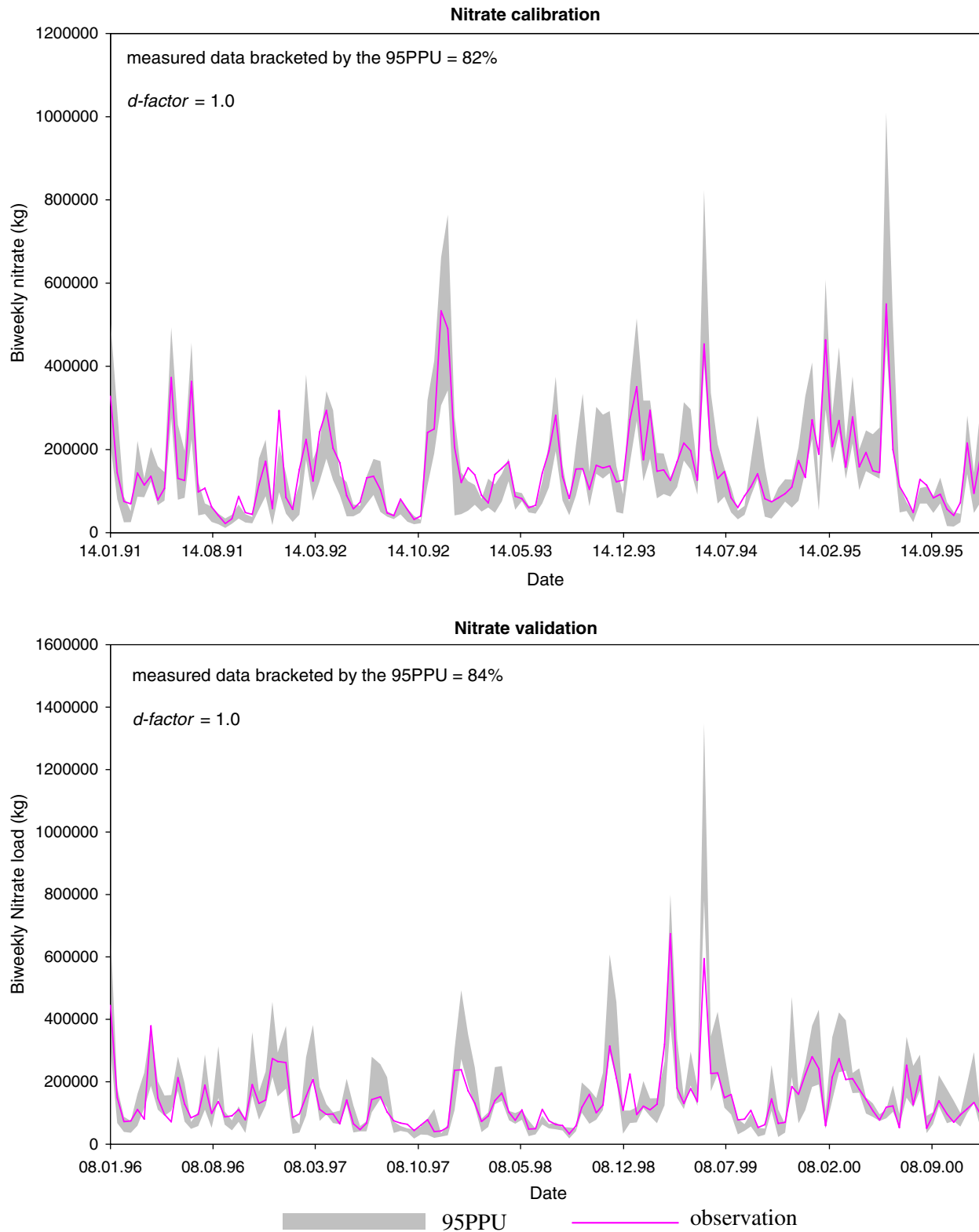
**Figure 9** Simulated and observed biweekly total phosphorus loads carried by the river at the watershed outlet. Top, total phosphorus calibration (1991–1995); and bottom, total phosphorus validation (1996–2000).

(BC4.swq), and organic P settling rate (RS5.swq). The last two of these parameters are related to in-stream processes.

Results of the nitrate simulation are given in [Fig. 10](#). Similar to the discharge, the nitrate simulation is also very good with small uncertainties, *d-factor* = 1, while bracketing 82% of the

data for calibration and 84% for validation. Five parameters were found to be sensitive to nitrate only. These included nitrogen in rain (RCN.bsn), nitrogen uptake distribution

parameter (UBN.bsn), concentration of nitrate in groundwater (GWNO3.gw), organic N enrichment for sediment (ERO-RGN.hru), and nitrate percolation coefficient (NPERCO.bsn).



**Figure 10** Simulated and observed biweekly nitrate loads carried by the river at the watershed outlet. Top, nitrate calibration (1991–1995); and bottom, nitrate validation (1996–2000).



In Table 2, sediment and nutrient loads released from the major landuses into the river is also reported for a wet, a dry, and an average year. As expected, the largest loads are produced in the wet year from the agricultural landuse followed by the summer pasture.

Two points are worth mentioning here. First, the objective function contained four variables, calibrating the model for any one variable would produce much better results for that variable but would not give as good a simulation result for other variables (the conditionality problem, Abbaspour et al., 1999). Second, ignoring the constraints would also produce better calibration and validation results at the watershed outlet, but the simulated loads from various landuses would not comply with our previous knowledge. Both these points indicate the importance of adding more variables in and constraining the objective function. This produces parameters reflecting more of the local processes, hence, providing more reasonable simulations. The down side is that more data are required for a reliable model calibration at the watershed scale. This also raises the important questions: when is a watershed model calibrated? And for what purpose can it be used for?

## Conclusions

Given the complexities of a watershed and the large number of interactive processes taking place simultaneously and consecutively at different times and places within a watershed, it is quite remarkable that the simulated results comply with the measurements to the degree that they do. Based on the results obtained in this study, SWAT is assessed to be a reasonable model to use for water quality and water quantity studies in the Thur watershed. On that positive note, however, a careful calibration and uncertainty analysis and proper application of modelling results should be exercised. The following conclusions could be drawn from the present study:

1. A watershed model calibrated based on measured data at the outlet of the watershed may produce erroneous results for various landuses and subbasins within the watershed, unless the objective function was constrained to produce correct results. This means that a large amount of measured data are necessary for a proper model calibration.
2. Simulation of particulates such as sediment and phosphorus are subject to large model uncertainties because of the "second-storm" effect, among others.
3. Large-scale watershed models could be effective for simulating watershed processes and therefore watershed management studies. The simulation of hydrology, sediment, and nutrient loads were of reasonable accuracy, allowing such integrated models to be used in scenario analysis.

## Acknowledgments

This study is a part of the NADUF program and supported by the Swiss Federal Office for Environment (FOEN), the Swiss

Federal Institute of Aquatic Science and Technology (Eawag), and MeteoSwiss. Special thanks to the graduate students at Eawag for their valuable contribution in collecting and processing environmental data.

## References

- Abbaspour, K.C., van Genuchten, M. Th., Schulin, R., Schläppi, E., 1997. A sequential uncertainty domain inverse procedure for estimating subsurface flow and transport parameters. *Water Resour. Res.* 33, 1879–1892.
- Abbaspour, K.C., Sonnleitner, M., Schulin, R., 1999. Uncertainty in estimation of soil hydraulic parameters by inverse modeling: example Lysimeter experiments. *Soil Sci. Soc. Am. J.* 63, 501–509.
- Abbaspour, K.C., Johnson, A., van Genuchten, M.Th., 2004. Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. *Vadose Zone J.* 3, 1340–1352.
- Arnold, J.G., Allen, P.M., 1996. Estimating hydrologic budgets for three Illinois watersheds. *J. Hydrol.* 176, 57–77.
- Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large area hydrologic modeling and assessment. Part I: model development. *J. Am. Water Resour. Assoc.* 34 (1), 73–89.
- Bagnold, R.A., 1977. Bedload transport in natural rivers. *Water Resour. Res.* 13 (2), 303–312.
- Beven, K., Binley, A., 1992. The future of distributed models: model calibration and uncertainty prediction. *Hydrol. Process.* 6, 279–298.
- Bicknell, B.R., Imhoff, J.C., Kittle, Donigian, A.S., Johanson, R.C., 1996. Hydrologic simulation program-FORTRAN user's manual, v.11, Athens, GA, USEPA.
- Binderheim-Bankay, E., Jakob, E., Liechti, P., 2000. NADUF – Messresultate 1977–1998 – Schriftenreihe Umwelt, Nr. 319, BUWAL, Bern (report in German).
- Borah, D.k., Bera, M., 2004. Watershed-scale hydrologic and nonpoint-source pollution models: review of applications. *Trans. ASAE* 47 (3), 789–803.
- Brown, L.C., Barnwell, T.O., 1987. The Enhanced Stream Water Quality Models QUAL2E and QUAL2E-UNCAS: Documentation and User Manual, Report EPA/600/3/87/007, US Environmental Protection Agency, Athens, GA.
- Budyko, M., 1974. *Climate and Life*. Academic Press, San Diego, CA, 508p.
- Chow, V.T., Maidment, D.R., Mays, L.W., 1988. *Applied Hydrology*. McGraw-Hill, New York.
- Di Luzio, M., Srinivasan, R., Arnold, J.G., Neitsch, S.L., 2002. ArcView Interface for SWAT2000, Blackland Research & Extension Center, Texas Agricultural Experiment Station and Grassland, Soil and Water Research Laboratory, USDA Agricultural Research Service.
- Duan, Q., Sorooshian, S., Gupta, H.V., Rousseau, A.N., Turcotte, R., 2003. Calibration of Watershed Models. American Geophysical Union, Washington, DC.
- EAWAG News Information Bulletin, 2000. Ground Water Research in Practice, No. 49, EAWAG, Duebendorf, Switzerland.
- Ewen, J., Parkin, G., O'Connell, E., 2000. SHETRAN: distributed basin flow and transport modelling system. *J. Hydrol. Eng.* 5 (3), 250–258.
- Gupta, H.V., Sorooshian, S., Yapo, P.O., 1998. Toward improved calibration of hydrologic models: multiple and noncommensurable measures of information. *Water Resour. Res.* 34, 751–763.
- Gupta, H.V., Sorooshian, S., Hogue, T.S., Boyle, D.P., 2003. Advances in automated calibration of watershed models. In: Duan, Q., Gupta, H.V., Sorooshian, S., Rousseau, A.N., Turcotte, R. (Eds.), *Calibration of Watershed Models*. American Geophysical Union, Washington, DC.

- Gupta, H.V., Beven, K.J., Wagener, T., 2005. Model calibration and uncertainty estimation. In: Anderson, M.G. (Ed.), *Encyclopedia of Hydrological Sciences*. Wiley, New York, pp. 2015–2031.
- Gurtz, J., Baltensweiler, A., Lang, H., 1999. Spatially distributed hydrotope-based modelling of evapotranspiration and runoff in mountainous basins. *Hydrol. Process.* 13, 2751–2768.
- Hargreaves, G., Samani, Z.A., 1985. Reference crop evapotranspiration from temperature. *Appl. Eng. Agric.* 1, 96–99.
- Huber, W.C., 1993. Contaminant transport in surface water. In: Maidment, D.R. (Ed.), *Handbook of Hydrology*. McGraw-Hill, New York, pp. 14.1–14.50.
- Jakob, A., Binderheim-Bankay, E., Davis, J.S., 2002. National long-term surveillance of Swiss Rivers. *Verh. Internat. Verein. Limnol.* 28, 1101–1106.
- Legates, D.R., McCabe, G.J., 1999. Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* 35, 233–241.
- McKay, M.D., Beckman, R.J., Conover, W.J., 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21, 239–245.
- Monteith, J.L., 1965. Evaporation and environment. In: Fogg, G.F. (Ed.), *The State and Movement of Water in Living Organisms*. Cambridge University Press, Cambridge, pp. 205–234.
- Prasuhn, V., 1999. Phosphor und Stickstoff aus diffusen Quellen im Einzugsgebiet des Bodensees, Ber. Int. Gewässerschutzkomm. Bodensee: V.51, Eidg. Forschungsanstalt für Agrarökologie und Landbau (FAL), Institut für Umweltschutz und Landwirtschaft (IUL) (report in German).
- Prasuhn, V., Sieber, U., 2005. Changes in diffuse phosphorus and nitrogen inputs into surface waters in the Rhine watershed in Switzerland. *Aquat. Sci.* 67 (3), 363–371.
- Prasuhn, V., Spiess, E., Braun, M., 1996. Methoden zur Abschätzung der Phosphor- und Stickstoffeinträge aus diffusen Quellen in den Bodensee, Ber. Int. Gewässerschutzkomm. Bodensee: V.45, Institut für Umweltschutz und Landwirtschaft (IUL) (report in German).
- Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T., 1992. *Numerical Recipe: The Art of Scientific Computation*, second ed. Cambridge University Press, Cambridge.
- Priestley, C.H.B., Taylor, R.J., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Rev.* 100, 81–92.
- SAEFL (Swiss Agency for the Environment, Forests and Landscapes) (SAEFL), 2002. *Environment – Switzerland 2002 report*, vol. 1.
- Simunek, J., Wendroth, O., van Genuchten, M.Th., 1999. Estimating unsaturated soil hydraulic properties from laboratory tension disc infiltrometer experiments. *Water Resour. Res.* 35, 2965–2979.
- Singh, J., Knapp, H.V., Arnold, J.G., Demissie, M., 2005. Hydrological modeling of the Iroquois river watershed using HSPF and SWAT. *J. Am. Water Resour. Assoc.* 41 (2), 343–360.
- Thorntwaite, C.W., 1948. An approach toward a rational classification of climate. *Geogr. Rev.* 38, 55–94.
- Wang, W., Neuman, S.P., Yao, T., Wierenga, P.J., 2003. Simulation of large-scale field infiltration experiments using a hierarchy of models based on public, generic, and site data. *Vadose Zone J.* 2, 297–312.
- Williams, J.R., 1980. SPNM, a model for predicting sediment, phosphorous, and nitrogen from agricultural basins. *Water Resour. Bull.* 16 (5), 843–848.
- Williams, J.R., Berndt, H.D., 1977. Sediment yield prediction based on watershed hydrology. *Trans. ASAE* 20 (6), 1100–1104.
- Williams, J.R., Jones, C.A., Dyke, P.T., 1984. A modeling approach to determining the relationship between erosion and soil productivity. *Trans. ASAE* 27, 129–144.
- Yapo, P.O., Gupta, H.V., Sorooshian, S., 1998. Multi-objective global optimization for hydrologic models. *J. Hydrol.* 204, 83–97.
- Zobrist, J., Reichert, P., 2006. Bayesian estimation of export coefficients from Diffuse and Point Sources of Swiss Watersheds. *J. Hydrol.* 329, 207–223.