Modeling Interbasin Groundwater Flow in Karst Areas: Model Development, Application, and Calibration Strategy

Van Tam Nguyen[a,*], Jörg Dietrich[a], Bhumika Uniyal[a]

[a]*Institute of Hydrology and Water Resources Management, Leibniz Universität Hannover, Appelstraße 9A, 30167 Hannover, Germany*

Abstract

Karstification is considered as one of the most common reasons for interbasin groundwater flow (IGF). IGF in some karst areas could be significant such that it must be accounted for in hydrologic modelling. In this study, the Soil and Water Assessment Tool (SWAT) was modified to explicitly account for IGF in karst areas. The modified model uses two conceptual models to simulate hydrologic processes in karst and non-karst regions. The modified model was applied in the karst-dominated region in the southwest Harz Mountains, Germany. Multisite streamflow data and satellite-derived actual evapotranspiration (ETa) were used for model calibration. Results show that (1) the modified model can be satisfactorily calibrated and validated for streamflow and ETa (2) the model performance for ETa and streamflow at some gauging stations are highly correlated, and (3) the use of satellite-derived ETa does not affect the model performance.

*Corresponding author

Email addresses: nguyen@iww.uni-hannover.de (Van Tam Nguyen), dietrich@iww.uni-hannover.de (Jörg Dietrich), uniyal@iww.uni-hannover.de (Bhumika Uniyal)

Keywords: Modified SWAT, Karst, Interbasin groundwater flow, Calibration, Satellite-derived evapotranspiration.

Software availability

Name of software: SWAT_IGF

Developer and contact address: Van Tam Nguyen (nguyen@iww.uni-hannover.de), Institute of Hydrology and Water Resources Management, Leibniz Universität Hannover, Appelstraße 9A, 30167 Hannover, Germany

Year available: 2019

Availability and cost: the source code is freely available at `https://github.com/tamnva/SWAT_IGF`

Language: Fortran

1. Introduction

The term "karst" refers to a region with distinct landscape features (e.g., sinking streams, sinkholes, and springs) and underground features (e.g., underground conduits and caves). In some karst regions, the karst landscape features could be absent or subtle, but their aquifers could be heavily karstified (Ford and Williams, 2007). Karst aquifers are developed as a result of dissolution of karstifiable rocks (e.g., limestone, dolomite, gypsum, and rock salt), the so-called karstification (Ford and Williams, 2007; Bögli, 1980; Howard, 1963). Karst aquifers account for about 10% to 15% of the continental area and karst groundwater is one of the sources of drinking water for approximately a quarter of the world's population (Ford and Williams, 2007). However, karst groundwater is particularly vulnerable to contamination due to their distinct hydrogeologic characteristics (Goldscheider, 2005; Doerfliger et al., 1999; Drew and Hötzl, 1999). Therefore, understanding the hydrogeologic characteristics of karst aquifers plays an important role in water resources management in karst regions.

Hydrogeologic characteristics of karst aquifers are different from other aquifers (Bakalowicz, 2005). Karst aquifers often exhibit a duality of recharge, infiltration, porosity, flow and storage (Goldscheider and Drew, 2007; White, 2002; Gun, 1986). Karst aquifers also show a high degree of spatial heterogeneity in hydraulic properties (Bonacci et al., 2006). Especially, the surface drainage basin in karst aquifers usually do not coincide with the groundwater basin (Spangler, 2001; Dar et al., 2014). Karstification is considered as one of the most common causes of interbasin groundwater flow (IGF) (Le Moine et al., 2007). Water recharged to karst aquifers could flow through

an underground conduit system spanning over several basins and emerge at springs located at distant sites (e.g., Anderson et al., 2006; Belcher et al., 2006; Le Moine et al., 2008). It should be noted that IGF could also occur in porous aquifer in form of regional groundwater flow (Tóth, 1963; Nguyen and Dietrich, 2018; Danapour et al., 2019), however, in this study we focus on IGF in karst areas. The term IGF in this study could be also understood as regional groundwater flow across surface topographic divides. IGF in karst areas could significantly alter the water budget of a basin (e.g., Anderson et al., 2006; Le Moine et al., 2008). Considering the aforementioned facts, IGF in karst areas should be accounted for in hydrological modeling, especially in the context of transboundary or interbasin groundwater management.

Various models have been used to simulate IGF in karst aquifers with varying model complexity, ranging from physically based distributed to conceptual lumped models. Physically based distributed models simulate groundwater flow based on hydraulic head gradient, therefore, groundwater could flow across topographic divide units, which are normally considered as isolated groundwater units in surface hydrology. Conceptual models can simulate IGF by allowing the simulation (or routing) of groundwater flow between topographical basins. Some models of these types are the Modular Three-Dimensional Finite-Difference Ground-Water Flow Model (MODFLOW, Scanlon et al, 2003), the modified WetSpa model (Liu et al., 2005), the modified Soil and Water Assessment Tool (SWAT, Arnold et al., 1998; Nerantzaki et al., 2015; Malagó et al., 2016; Palanisamy and Workman, 2014), modèle du Génie Rural 4 paramètres Journalier (GR4J, Perrin and Michel, 2003; Le Moine et al., 2007, 2008), the tank model (Anaya and Wanakule,

1993), and the multi-cell aquifer model (Rozos and Koutsoyiannis, 2006; Barrett and Charbeneau, 1997). SWAT is one of the most widely-used models to simulate the effect of land use, agricultural management practices and climate change on water and chemical yields in non-karst areas (Arnold et al., 2005; Gassman et al., 2007; Krysanova and White, 2007; Molina-Navarro et al., 2017). Therefore, the modified SWAT versions which account for IGF in karst areas could potentially help to explore these effects in karst regions.

The aforementioned modified SWAT models, the so-called KarstSWAT (Palanisamy and Workman, 2014) and KSWAT (Nerantzaki et al., 2015; Malagó et al., 2016), simulate IGF in karst regions. The KarstSWAT model was specifically developed for watersheds dominated by sinkholes and springflow, which is mainly fed by the water from sinkholes (Palanisamy and Workman, 2014). The KSWAT model combines the *adapted SWAT model* (Fig. 3, Malagó et al., 2016) and the *karst-flow model* (Nikolaidis et al., 2013). The *adapted SWAT model* assumes that all water entering the soil profile is karst groundwater recharge (Fig. 3, Malagó et al., 2016). However, part of the infiltrated water could contribute to the streamflow as lateral flow and baseflow if the underlying aquifer of a subbasin is not entirely a karst aquifer (e.g., Palanisamy and Workman, 2014). The *adapted SWAT model* does not differentiate between concentrated recharge and diffuse recharge. The *karst-flow model* is the two-linear-storage reservoir model, which receives the recharge simulated from the *adapted SWAT model* (or from the original SWAT model, Nikolaidis et al., 2013) and routes it to spring. Outflows from the two reservoirs of the *karst-flow model* represent flow from wide conduits and narrow fractures (Kourgialas et al., 2010; Malagó et al., 2016). Because of the lumped

feature of deep recharge from the *adapted SWAT model*, the KSWAT model does not explicitly differentiate between (1) the diffuse recharge and concentrated recharge, (2) between matrix storage and conduit storage. This is important because these recharges and storages are different in term of travel time and storage. In addition to the aforementioned disadvantages, the recharge area of the karst aquifer in the KarstSWAT and KSWAT models follows the subbasin delineation of SWAT.

In addition to the model development, parameter identification in karst regions is also subject to higher uncertainty compared to other regions (Brenner et al., 2018; Hartmann et al., 2017, 2013). This is because the karst aquifer is highly heterogeneous and the upper flux (actual evapotranspiration, ETa) and the lower flux (karst groundwater recharge) are usually unknown. In order to develop a robust model and to minimize the parameter uncertainty, especially in karst regions, multi-variable calibration is suggested. ETa is one of the main components of the hydrologic cycle. About 60% of the annual precipitation on the global land surface returns to the atmosphere as evapotranspiration (Jung et al., 2010; Oki and Kanae, 2006). Considering the aforementioned facts, observed ETa should be used for calibrating the model. However, a direct observation of ETa is very scarce.

In non-karst areas, many studies have used satellite-derived ETa for model calibration (e.g., Rajib et al., 2018; Franco and Bonumá, 2017; Vervoort et al., 2014; Rientjes et al., 2013; Droogers et al., 2010; Zhang et al., 2009; Muthuwatta et al., 2009; Immerzeel and Droogers, 2008). In these studies, satellite-derived ETa was either used as an independent calibration data set or as input data. Results showed that the model performance for streamflow

could decrease when constraining model calibration with satellite-derived ETa as an additional variable (Vervoort et al., 2014). However, the above-mentioned studies showed that using satellite-derived ETa in combination with observed streamflow for calibrating a hydrologic model could (1) better reproduce the catchment's water balance, (2) reduce the parameter uncertainty, (3) increase the model robustness, and (4) detect the structural model issues. In karst areas, the use of satellite-derived ETa as an additional calibration variable has not been given enough attention.

In this study, we developed a conceptual model which is able to (1) simulate surface and subsurface flows in both karst and non-karst areas, (2) apply for a region where the karst aquifer boundaries do not coincide with the surface subbasin boundaries, and (3) represent different recharges (diffuse recharge and concentrated recharge) and storages (matrix storage and conduit storage) in karst areas. The proposed concept was implemented into the SWAT model. The modified SWAT model was tested in the karst-dominated area in Lower Saxony, Germany. The effects of using satellite-derived ETa for model calibration on the model performance was examined in detail. The Moderate Resolution Imaging Spectroradiometer (MOD16 ETa, Mu et al., 2013) was used for the model calibration.

2. Methodology

2.1. The original SWAT model

In SWAT, a basin can be divided into subbasins, which are further divided into Hydrologic Response Units (HRUs). HRUs are created by lumping all areas having the same combination of land use, soil type and slope within

a subbasin. The HRU concept is computationally efficient while incorporating the aforementioned landscape properties. SWAT simulates two phases of the hydrologic cycle, the land phase and the routing phase. The land phase includes HRU-related processes such as surface processes (e.g., evapotranspiration, surface runoff, vegetation-related processes) and subsurface processes (e.g., percolation, lateral flow, groundwater recharge, return flow) (Fig. 1A). The routing phase includes stream-related processes (e.g., flood routing, nutrient transport) and reservoir routing. In SWAT, groundwater recharge is partitioned into shallow and deep aquifer recharge. Recharge into the shallow aquifer ultimately returns to stream as baseflow while recharge into the deep aquifer is considered as a loss. SWAT is not capable of simulating groundwater flow between HRUs (or subbasins) due to the non-spatial characteristic of the HRU concept. A more detailed description of the SWAT model is given by Neitsch et al. (2011).

2.2. The modified SWAT model for IGF

In this section, after a summary of the general hydrogeologic characteristics of karst areas, the modified SWAT for karst areas is presented. The modified SWAT model for modeling IGF, hereafter referred to as the SWAT_IGF model, is comprised of two conceptual models. The original conceptual model of SWAT is applied for non-karst areas (Fig. 1A) while modified conceptual model of SWAT is applied for karst areas (Fig. 1B). The two conceptual models were combined into a single program, resulting in a single executable file. An aquifer classification map is used as an additional criterion for the delineation of HRUs (Fig. 3C). This aquifer classification map contains information about the aquifer type and the extended recharge area of each

spring. Then, the SWAT_IGF will assign the appropriate conceptual model for the karst and non-karst HRUs automatically (Fig. 3C) and recharge from the extended karst area will be routed to the corresponding spring. The user needs to assign the amount of recharge to each spring (in case multiple springs are fed by the same recharge area).

Recharge into the karst aquifer could either be classified as (1) autogenic or allogenic recharge or (2) concentrated or diffuse recharge (Gun, 1986; Taylor and Greene, 2008; Ford and Williams, 2007). Autogenic recharge originates from precipitation falling on the karst areas while allogenic recharge originates from runoff on non-karst areas. Concentrated recharge can occur via sinkholes, losing streams, closed depressions, and well-developed fissures. Diffuse recharge is areal recharge through the unsaturated soil zone. Recharge into the karst aquifer is often drained by a well-developed solution-conduit system and discharged via one or several springs. Flow in the conduit is often fast and turbulent while flow in the rock matrix is slow and laminar (White, 2002; Hartmann et al., 2014). However, the majority of karst groundwater is stored in the rock matrix. Due to the fast flow and small storage of the conduit system compared to that of the rock matrix, the response of discharge to recharge from the conduit system is often faster than that from the matrix storage.

In this study, the SWAT_IGF is proposed for the cases where (1) the recharge area and discharge points (springs) are located in different subbasins and (2) the discharge points are located in one subbasin. Further modifications could be done for other cases. A two-reservoir model is proposed to represent the duality of and storage and discharge of the karst

area (Fig. 1B). The first reservoir, hereinafter referred to as the matrix storage reservoir, represents groundwater storage in the rock matrix. The matrix storage reservoir receives diffuse recharge from the overlaying zone. The second reservoir, hereinafter referred to as the conduit storage reservoir, represents groundwater storage in the conduit system. The conduit storage reservoir receives (1) concentrated recharge from closed depressions, infiltration losses from streams, fractures and dolines and (2) diffuse discharge from the matrix storage reservoir. It should be noted that there could be flow from the conduit storage reservoir to the matrix reservoir (e.g., Screaton et al, 2004), however, it is not explicitly considered in this study. We consider flow from the matrix to the conduit as net flow, which already takes into account flow from the conduit to the rock matrix.

Diffuse recharge from the bottom of the soil profile to the matrix storage reservoir on day i, taking into account the delay time in the unsaturated zone, is calculated using the exponential decay weighting function (Venetis, 1969; Sangrey et al, 1984):

$$w_{rd,i} = (1 - e^{-1/\delta_{gw}}) \cdot \beta \cdot w_{seep,i} + e^{-1/\delta_{gw}} \cdot w_{rd,i-1} \tag{1}$$

where $w_{rd,i}$ and $w_{rd,i-1}$ (mm H_2O) is the amount of diffuse recharge to the matrix reservoir on day i and $i-1$, respectively, δ_{gw} (days) is the delay time for infiltrated water to reach the matrix storage reservoir, β (-) is the recharge separation factor, ranging from 0 to 1, w_{seep} (mm H_2O) is the total amount of water exiting the bottom of the soil profile on day i.

Outflow from the matrix storage reservoir is simulated using the linear storage-discharge relationship (e.g., Nikolaidis et al., 2013; Neitsch et al.,

2011):

$$Q_{matrix,i} = e^{-\alpha_{matrix}\cdot\Delta t}\cdot Q_{matrix,i-1} + (1-e^{-\alpha_{matrix}\cdot\Delta t})\cdot \sum_{j=1}^{nhrus} w_{rd,i,j}\cdot a_j \cdot 10^{-3} \quad (2)$$

where $Q_{matrix,i}$ and $Q_{matrix,i-1}$ (m^3 H_2O) are the outflows from the matrix storage reservoirs on day i and $i-1$, respectively, α_{matrix} (1/day) is the recession constant of the matrix storage reservoir, respectively, Δt is the time step (Δt = 1 day), $w_{rd,i,j}$ (mm H_2O) and a_j (m^2) are the diffuse recharge and area of the hydrologic response unit j, respectively, 10^{-3} is the unit conversion factor (from mm H_2O to m H_2O), nhrus is the number of HRUs in the recharge area.

Concentrated recharge from closed depressions, fractures, and sinkholes to the conduit storage reservoir on day i, $w_{rc,i}$ (mm H_2O), is calculated as follows:

$$w_{rc,i} = (1-\beta)\cdot w_{seep,i} \quad (3)$$

The total amount of recharge to the conduit storage reservoir on day i, $W_{rconduit,i}$ (m^3 H_2O), is expressed as follows:

$$W_{rconduit,i} = \sum_{j=1}^{nhrus} w_{rc,i,j}\cdot a_j \cdot 10^{-3} + rttlc_i + Q_{matrix,i} \quad (4)$$

where $rttlc_i$ (m^3 H_2O) is the mount of recharge from losing streams on day i. Outflow from the conduit storage reservoir is simulated using the linear storage-discharge relationship:

$$Q_{conduit,i} = e^{-\alpha_{conduit}\cdot\Delta t}\cdot Q_{conduit,i-1} + (1-e^{-\alpha_{conduit}\cdot\Delta t})\cdot W_{rconduit,i} \quad (5)$$

where $Q_{conduit,i}$ and $Q_{conduit,i-1}$ (m^3 H_2O) are outflows from the conduit storage reservoir on day i and $i-1$, respectively, $\alpha_{conduit}$ (1/day) is the recession constant of the conduit storage reservoir.

The total runoff of a basin where the springs are located, $Q_{river,i}$ (m^3 H_2O), is calculated as follows:

$$Q_{river,i} = Q_{conduit,i} + Q_{direct,i} \tag{6}$$

where $Q_{direct,i}$ (m^3 H_2O) is the direct runoff (the sum of surface runoff and lateral flow) from the basin where the spring is located.

It should be noted that the conduit and matrix reservoirs proposed in this study correspond to the upper and lower reservoirs of the *karst-flow model* (Nikolaidis et al., 2013), respectively. The conduit and the matrix reservoirs are arranged in series while the upper and lower reservoirs are arranged in parallel. The lower reservoir receives recharge from the upper reservoir while the conduit receives recharge from the matrix reservoir. Springflow in the *karst-flow model* is directly fed by the upper and lower reservoirs while it is only directly fed by the conduit reservoir in the SWAT_IGF model. Outflows from both reservoirs in both models are simulated using a linear storage-discharge relationship.

3. Case Study

3.1. Study area and data

The study area is located in the southwest Harz Mountains (non-karst area) and the southern Harz rim (karst-dominated area) in Northern Germany with a drainage basin of about 384 km^2 (Fig. 2). The study area

has two outlets located at the Rhume spring and Lindau gauging stations. The study area receives inflow from the Odertalsperre reservoir. The Digital Elevation Model (DEM) obtained from the Niedersächsische Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz (NLWKN) shows that the elevation of the study area varies from 142 m to 929 m above mean sea level (a.m.s.l). Land use/land cover (LULC) map was taken from the Copernicus Land Monitoring Service. The soil map (BÜK 200) and soil profile data were obtained from the Bundesanstalt für Geowissenschaften und Rohstoffe (BGR) (Fig. 3). Initial soil hydraulic conductivity and soil available water content were derived by using the pedotransfer functions/tables (Wessolek et al., 2009). The dominant land use/land cover classes are forest and agricultural, accounting for about 55% and 31% of the study area, respectively. The most dominant soil type in the southwest Harz Mountains is spodic Cambisols from acid igneous and metamorphic rocks, covering 46% of the study area. In the southern Harz rim, most of the soils were developed from gypsum with low water-holding capacity (Schnug et al, 2004). Observed groundwater level data at three wells located within and nearby the Pöhlder Becken were collected from the NLWKN (Fig. 2).

Daily weather data (precipitation, wind speed, temperature, solar radiation, and relative humidity) from 1997-2010 were obtained from Deutscher Wetterdienst (DWD). Weather data from observed stations were interpolated for all subbasins using the inverse distance weighing (IDW) method. The study area has an average annual precipitation of 1242 mm/yr with high spatial variability. The annual precipitation is up to 1619 mm/yr in the southwest Harz Mountains, whereas that in the southern Harz rim is 862

mm/year. Temperature in the study area decreases with an increase in elevation. Daily observed streamflow and reservoir outflow were obtained from the NLWKN and the Harzwasserwerke (HWW). The MOD16 ETa at 8-day time step and 1 km^2 spatial resolution was downloaded using the MODISTools (Tuck et al., 2014).

3.2. Geology

The study area consists of two distinct geologic areas, the southwest Harz Mountains and the southern Harz rim (Grimmelmann, 1992). The Harz Mountains were part of the European Variscan fold belt formed by the collision of Africa, Baltica, Laurentia and other microplates in the early Paleozoic Era (Tait et al., 1997; Haggett, 2002). The Harz Mountains were later eroded and a large part of it was inundated by the Zechstein Sea (Haggett, 2002; Koster, 2005). Under hot and dry climatic conditions of the late Permian period, a large amount of evaporites was formed in the inundated area after several evaporation cycles (Taylor, 1998; Schnug et al, 2004; Kramm and Wedepohl, 1991; Böttcher, 1999; Tucker, 1991).

After other geologic processes, the underlying geology of the southern Harz Mountains nowadays mainly consists of Palaeozoic greywacke, shale, and conglomerate (Fig. 4) while in the southern Harz rim, the Permian Zechstein (dolomite, gypsum, anhydrite) was exposed to the surface and subjected to the karstification process (Voigt et al., 2008; Schnug et al, 2004; Böttcher, 1999; Paul and Vladi, 2001). There is a 2- to 6-km-wide strip of exposed Permian Zechstein in the southern Harz rim with various karst features such as sinking streams, sink holes, caves, and springs (Liersch, 1987). The karst area in this region is subjected to a continuous karstification

process. About 7092 tons of sulphur bound to gypsum are washed from this karst-dominated area each year (Schnug and Haneklaus, 1998; Herrmann, 1969). Gelogical cross-sections in the area show that the Permian Zechstein rocks are exposed to the surface near the southern Harz rim and overlaid by non-karstifiable rocks in the south. At the Oder and Sieber rivers, it was overlaid by a Quaternary fluvial deposit layer originated from the Harz Mountains (Fig. 4). Detailed geologic maps and geologic cross-sections of the study area can be found in Herrmann (1969), Grimmelmann (1992), Liersch (1987), Voigt et al. (2008), and NIBIS®Kartenserver (`http://nibis.lbeg.de/cardomap3/?TH=647`).

3.3. Hydrogeology

The main Rhume spring outlet is located in a NW-SE trending fault, where flow in the underground conduit of the Zechstein deposits is blocked by a low permeability Lower Buntsandstein stratum (Herrmann, 1969; LaMoreaux and Tanner, 2001). Besides the main outlet with a diameter of about 20 m, there are about 360 small outlets located nearby (Herrmann, 1969). They altogether release an average discharge of about 2.2 m^3/s via a small stream with a minimum of 1.5 m^3/s during low flow periods. This indicates that there could be a relatively big subsurface matrix storage in the area compared to the Rhume spring subbasin. The sum of discharge from the main Rhume spring outlet and its neighboring outlets is hereafter referred to as the Rhume spring discharge. Many studies have been conducted to explain the origin of the water from the Rhume spring discharge since early 20th century.

Thürnau (1913) conducted tracer tests with Uranine and found that the infiltrated tracers from the area in the southern Harz rim, which were later known as the Pöhlder Becken, reappears at the Rhume spring (Fig. 2). Thürnau (1913) was also able to determine the main losing streams (Fig. 4) in the Pöhlder Becken as well as the travel time of tracers from the infiltration points to the Rhume spring. Haase et al. (1970) analyzed the water balance in the study area and found that there are significant infiltration losses in the Sieber and Oder rivers. In 1981, another tracer tests with about 12 kg of Uranine were carried out at sinkholes near Herzberg (Liersch, 1987). The injected tracers were detected at the Rhume spring about 78 hours after the injection and were almost undetectable after 25 days. From this experiment, a flow path of about 7500 m and a horizontal groundwater flow velocity of over 100 m/h were estimated (LaMoreaux and Tanner, 2001). A three-reservoir storage model was proposed to explain the breakthrough curve of tracer concentration at the Rhume spring (Liersch, 1987). Rienäcker (1987) found that the time-lag between peak discharges of the Sieber (at Hattorf gauging station), of the Oder (at Scharzfeld gauging station) and the Rhume spring varies between 24 to 72 hours, depending on the existing groundwater reservoir storage level. Results from various geophysical and tracer experiments showed that infiltrated water from the Pöhlder Becken, hereinafter referred to as the recharge area of the Rhume spring (Fig. 2), and transmission losses of the rivers located in this area are the main sources of the Rhume spring discharge (Goldmann, 1986; Liersch, 1987).

The recharge area of the Rhume spring receives allogenic recharge from upstream subbasins via a connected river network in the area. In addition, it

also receives groundwater inflow from southwest Harz Mountains. However, the estimated amount is negligible, < 0.03 m^3/s (Grimmelmann, 1992). The estimated contribution of flow from the Rhume basin (with an area of 8 km^2) is about 4% of the Rhume spring discharge. About 96% of the Rhume spring discharge is from IGF, of which about 60% originates from the infiltration loss of the Oder and Sieber rivers (Goldmann, 1986; Liersch, 1987; LaMoreaux and Tanner, 2001). Therefore, the original SWAT_IGF should be used instead of the original SWAT to explain 96% of the flow volume at the Rhume spring.

4. Model setup, calibration and validation

4.1. Model setup

The study area was divided into 26 subbasins and 1094 HRUs based on LULC, soil, DEM and aquifer map (Fig. 3). The thresholds for defining HRUs were set to zero to include all of the basin landscape. The SWAT_IGF model uses the conceptual model presented in Fig. 1A for the southwest Harz Mountains and the conceptual model presented in Fig. 1B for the southern Harz rim (Fig. 3C). Infiltration losses (w_{seep}) and river transmission losses ($rttlc$) from the karst area located outside the recharge area of the Rhume spring were considered as losses from the hydrologic system. The model was set to run for the period of 14 years (from 1997 to 2010) with 3 years of warm-up (1997-1999), 6 years of calibration (2000-2005), and 5 years of validation (2006-2010) at a daily time step. In order to have a comparable result with MOD16 ETa, the Penman-Monteith method (Monteith, 1965; Allen, 1986; Allen et al., 1989) (which was used for deriving MOD16 ETa) was used for calculating evapotranspiration in SWAT_IGF.

4.2. Calibration and validation strategy

In this study, the Sequential Uncertainty Fitting (SUFI-2) in the SWAT-Calibration and Uncertainty Programs (SWAT-CUP) was used for parameter sensitivity, model calibration, validation and uncertainty analysis (Abbaspour, 2013; Abbaspour et al, 2007, 2004). The selected parameters and their initial ranges (Tab. 2) were chosen based on local expertise and literature review (Arnold et al., 2012; White and Chaubey, 2005; Lam et al., 2012; Maier and Dietrich, 2016; Uniyal et al., 2017; Nguyen and Dietrich, 2018; Rajib et al., 2018). Global sensitivity analysis was used to identify the important influencing factors and to reduce the number of parameters for model calibration. SUFI-2 uses multiple regression and t-test to identify the relative sensitivity of each parameter. Within this approach, a higher absolute value of *t-stat* and a smaller *p-value* indicate a higher sensitivity of the parameter (Abbaspour et al, 2018).

Several multi-criteria objective functions were proposed and tested. The following form of the multi-criteria objective function was found to be appropriate for this study:

$$OF = max\Big(\frac{w_1 \cdot \sum_{i=1}^{5} NSE_{Q_i} + w_2 \cdot NSE_{Q_{Lindau}} + w_3 \cdot NSE_{Q_{Rhumespring}} + w_4 \cdot NSE_{ETa}}{5 \cdot w_1 + w_2 + w_3 + w_4}\Big) \tag{7}$$

where OF is the multi-criteria objective function, NSE_{Q_i} is the Nash-Sutcliffe efficiency (Eq. 8, Nash and Sutcliffe, 1970) for streamflow at five streamgauging stations inside the catchment (Hattorf, Scharzfeld, Herzberg, Kupferhütte, and Pionierbrücke), $NSE_{Q_{Lindau}}$, $NSE_{Q_{Rhumespring}}$ and NSE_{ETa} are the NSE for streamflow at the catchment outlets (Rhume spring and Lindau gauging

stations) and the NSE for ETa, respectively, w is the weight. For sensitivity analysis, the weights in the objective function were assigned as follows: $w_1 = 1$, $w_2 = 5$, $w_3 = 5$, $w_4 = 5$. Therefore, the model performances for streamflow at the Lindau, Rhume spring, five aforementioned gauging stations inside the catchment, and for ET are considered equally important in the objective function.

Three calibrations scenarios were carried out with an increase in the number of calibrated variables from calibration scenarios S1 to S3 (Tab. 1). If a variable is not calibrated, its corresponding weight in the objective function is set to zero (Tab. 1). The objective of these calibration scenarios is to examine the effects of using multi-site streamflow and MOD16 ETa for model calibration on the model performance. For model calibration, 1000 parameter sets were generated using Latin hybercube sampling. These parameter sets were used for all three calibration scenarios.

Although only the NSE was considered in the objective function, the Kling-Gupta efficiency (KGE, Gupta et al., 2009) and percent bias (PBIAS) was also calculated for the best simulation as follows:

$$NSE = 1 - \frac{\sum_{i=1}^{n}(x_i^{obs} - x_i^{sim})^2}{\sum_{i=1}^{n}(x_i^{obs} - \bar{x}^{obs})^2} \tag{8}$$

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \tag{9}$$

$$PBIAS(\%) = 100 \cdot \frac{\sum_{i=1}^{n}(x_i^{obs} - x_i^{sim})}{\sum_{i=1}^{n} x_i^{obs}} \tag{10}$$

where x_i^{obs} and x_i^{sim} are the observed and simulated values, respectively, at time step i, $\bar{x}^{obs}$ is the mean of observed values, n is the number of simulated

values, r is the linear regression coefficient between observed and simulated values, α (β) is the ratio of standard deviation (mean) of observed over standard deviation (mean) of simulated values.

In SUFI-2, parameter uncertainty, which is represented as a uniform distribution, integrates all types of uncertainties (e.g., uncertainty in input data, model concept, model parameter, and measured variables). All of these uncertainties ultimately propagate into the model output uncertainty, which is expressed by the 95% prediction uncertainty band (95PPU). The *p-factor* (the percentage of measured data bracketed by the 95PPU band) and *r-factor* (the average thickness of the 95PPU band divided by the standard deviation of the measured data) are used to characterize the 95PPU band (Abbaspour et al, 2018).

5. Results and discussion

5.1. Sensitivity analysis and best calibrated parameter set

Tab. 2 shows the results of global sensitivity analysis for 21 model parameters. Parameter sensitivity ranking was based on the values of *t-stat* and *p-value*. It is seen that CN2 is the most sensitive parameter. This indicates that streamflow, karst groundwater recharge, and evapotranspiration are strongly affected by the surface runoff generation process. The parameter CH_K2 (riverbed hydraulic conductivity) is listed among the most sensitive parameters. This is because river transmission losses in the karst area could infiltrate into the conduit network and formulate interbasin groundwater flow, ultimately affect the catchment water balance. The high sensitivity

ranking of ESCO is because this parameter controls the amount of evaporation from the soil.

It is seen that the parameter which controls the amount of deep groundwater recharge (RCHRG_DP) was found insignificant. This is because this parameter only exists in the conceptual model for the non-karst area. The non-karst area in this case is the Harz Mountains with high topographic gradient. In this area, the runoff coefficient is expected to be high, therefore, the amount of deep groundwater recharge is expected to be minor compared to surface runoff. The newly introduced parameters for the karst area (β, $\alpha_{conduit}$, α_{matrix}) are not identified as sensitive parameters. This could be due to the fact that these parameters only affect the Rhume spring discharge, which plays a minor role in the objective function (Eq. 7 and Tab. 1). However, one-at-time sensitivity analysis shows that these parameters significantly affect the dynamic of the simulated Rhume spring hydrograph and they should be taken into account for a successful model calibration.

Based on the result of sensitivity analyses and the process-based evaluation as aforementioned, the seven most sensitive parameters and the three parameters of the karst model were selected for model calibration. The best parameter values obtained from automatic calibration were shown in Tab. 2.

5.2. The role of using MOD16 ETa and multi-site streamflow data and for model calibration

Calibration results show that the calibration scenarios S2 and S3 have the same best parameter values (Tab. 2) and the same number of behavioral simulations (71 behavioral simulations with a behavioral threshold of 0.5). As a result, the model performance statistics between the calibration scenarios S2 and S3 are identical (Tab. 4). This indicates that using MOD16 ETa for model calibration does not effect the model performance in this case study. A detailed examination of the results shows that simulated ETa from the calibration scenario S2 fits well with MOD16 ETa despite MOD16 ETa was not used for model calibration (Fig. 5 and Tab. 4). In addition, the model performance for ETa tends to be improved with improvement of the model performance for streamflow at the Lindau, Scharzfeld, and Kupferhütte gauging stations. This was shown by a strong positive correlation ($r \geq 0.78$) between NSE_{ETa} and NSE_Q at these gauging stations in the calibration scenario S2 (Fig. 6). As a result, the best model performance for streamflow in these gauging stations is likely to be among the "best" model performances for ETa and the use of MOD16 ETa for model calibration might not have any effect (or only minor effects) on the model performance. The results indicate that if there is a strong positive correlation in the model performances between two different variables in a multi-variable calibration, one variable can be dropped out of the objective function without having much influence on the model performance. For multi-site calibration, the selected stream gauges should be located in different rivers unless there are some major changes in the river segment.

The aforementioned results, however, should be considered along with the weights used in the objective function (Tab. 1). It should be noted that differences between the calibration results of scenarios S2 and S3 occur if the weight for NSE_{ETa} accounts for more than 70% of all weights in the objective function, $w_4 \geq 0.7 \cdot (5 \cdot w_1 + w_2 + w_3 + w_4)$. It means that improving the model performance for ETa is the main objective, which is not the objective in this study.

It is seen from the Tab. 4 that the model performance for streamflow at the Rhume spring was reduced, from NSE = 0.75 (scenario S1) to NSE = 0.69 (scenario S2), when streamflow data at additional stream gauges were used for model calibration. However, the model prediction uncertainty was reduced and the model robustness was increased. This is shown by a decrease in the *r-factor* (from 1.10 to 1.01) and a decrease in the difference of NSE between the calibration and validation periods (from 0.27 to 0.07, Tab. 4). In the calibration scenario S2, the model performance for streamflow at all gauging stations (except at the Rhume spring) and for ETa are improved compared to that in the calibration scenario S1. The results indicate that in a karst-dominated region, multi-gauge calibration should be done in order to have a better model performance. Therefore, only results from the calibration scenario S2 were discussed in detail in the remaining sections.

5.3. Simulated streamflow

Fig. 7A-G presents the observed and simulated streamflow hydrographs and their respective flow duration curves during the calibration period with the best calibrated parameters. It is seen that the SWAT_IGF tends to underestimate high flows (Fig. 7A-G) and low flows (Fig. 7D, E and F). The

underestimation of high flows and low flows is inherited from the original SWAT (e.g., Uniyal et al., 2017; Nguyen and Dietrich, 2018; Nguyen et al., 2018). This could be a reason for the small *p-factor* observed from the model calibration outputs (Tab. 4). The good fit between simulated low flows at the Lindau and Sharzfeld gauging stations with observed data (Fig. 7A-G) is due to the effect of using observed outflow from the Oder dam (Odertalsperre, Fig. 2) as input data to the model. At the Hattorf gauging station, low flows were overestimated by the model (Fig. 7B). This is due to a non-linear relationship between discharge and transmission losses of the Sieber river, which cannot be represented in the current SWAT_IGF model. In this river, transmission losses are reported to be higher (more than 70% of the river discharge) with smaller discharges (Thürnau, 1913). At the Rhume spring gauging station, the observed flow duration curve is well reproduced by the model and the 95PPU band covers most of the observed values (*p-factor* = 96). Simulated results show that runoff generated from the Rhume spring basin accounts for about 4% of the Rhume spring discharge, whereas the remainder (96%) is from IGF. The results match well with the ones reported by Goldmann (1986). Simulated results from the SWAT_IGF also show that annual transmission losses from the Sieber and Oder river systems contribute about 59% of the Rhume spring discharge, which is similar to the previously estimated value of 60% (LaMoreaux and Tanner, 2001).

Due to a significant contribution of IGF to the Rhume spring as aforementioned, the original SWAT model failed to simulate flow at this gauging station (Fig. 7G). It should be noted that simulated streamflows in the karst area (Lindau, Hattorf, and Scharzfeld gauging stations) from the original

SWAT could be better than the SWAT_IGF. This is because parameters of the SWAT_IGF model in the karst region are further constrained to match the simulated streamflow at the Rhume spring with observed data. Therefore, we did not compare the simulated streamflow from the original SWAT and the SWAT_IGF at these gauging stations. In the validation period (2006-2010), similar results were also observed (Fig. 8A-G).

5.4. Simulated karst groundwater storage variation

Fig. 9A-C shows 1) the variations of simulated karst groundwater storage (the total groundwater storage in the matrix and conduit storage reservoirs) in the recharge area of the Rhume spring and 2) changes in the observed groundwater levels in three wells (Fig. 2). It is expected that changes in the groundwater levels reflect the variations in karst groundwater storage. In three wells, it is seen that the annual variations in the simulated karst groundwater storage agree well with the observed groundwater levels. Especially with well 1, a high correlation coefficient ($r = 0.93$) between the simulated groundwater storage and the observed groundwater levels was found (Fig. 9A). At wells 2 and 3 (Fig. 9B-C), lower correlation coefficients ($r = 0.73$ and $r = 0.47$, respectively) were found. The simulated karst groundwater storage varies from 35 to 67 million m^3 with an average value of about 48 million m^3.

6. Conclusions and recommendations

Interbasin groundwater flow (IGF), especially in karst areas, could significantly alter the water budget of a region. In this study, the original SWAT model was modified for simulating IGF in karst areas, resulting in

the SWAT_IGF model. A two-linear-reservoir model was proposed to represent the duality of recharge, infiltration, storage, and discharge in the karst area. The study area is located in a karst-dominated region in the southwest Harz Mountains, Germany. The model was successfully calibrated at the Rhume spring and at multiple sites for streamflow, and for ETa by using MOD16 ETa.

Calibration results show that multi-site calibration is necessary to achieve a good model performance. Simulated ETa from the SWAT_IGF model matches well with MOD16 ETa despite MOD16 ETa was not used for model calibration. The use of MOD16 ETa as additional calibration variable does not affect the model performance. This is because the model performance for ETa tends to be improved with an improvement of the model performance for streamflow at some gauging stations. The conclusion regarding the use of MOD16 ETa for model calibration, however, should not be generalized to other satellite remote sensing products and to studies in other areas.

The SWAT_IGF model was demonstrated as a robust model by further validating the model outputs with other data. The SWAT_IGF is also highly flexible. It could be applied in both karst and non-karst areas where the surface subbasin boundaries do not coincide with the subsurface subbasin boundaries. The model uses a parsimonious approach for modelling IGF in karst systems while explicitly representing the duality of recharge, discharge, and storage in karst regions.

The SWAT_IGF introduced in this study, however, has not been developed for modelling solute transport. Different solute transport models could be incorporated into the SWAT_IGF model due to its flexible structure. For

example, future studies could apply a well-mixed model for modelling solute transport in the conduit because flow in the conduit storage is fast and turbulent. For solute transport in the soil matrix, the catchment scale formulation of transport based on travel time distributions appears to be a promising tool (Botter et al., 2011; Benettin et al., 2013). The concept of travel time based formulation of transport could be used to simulate (1) the delay between input and output solute concentration signals and (2) different selection schemes for outflow from the rock matrix. In addition, the recharge separation factor (β) was assumed to be constant regardless of the rainfall event characteristics. Future studies could use different recharge separation factors depending on different rainfall event characteristics (Hartmann et al., 2015b).

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. We thank the Editor and three anonymous reviewers for their constructive comments, which helped significantly improve quality of the manuscript. We also thank the DWD, HWW, BGR, and NLWKN for providing data.

References

Abbaspour, K.C., Vaghefi, S.A., Srinivasan, R. A., (2018). A guideline for successful calibration and uncertainty analysis for Soil and Water Assessment: A review of papers from the 2016 International SWAT Conference. Water, 2018, 10, 6.

Abbaspour, K.C., 2013. SWAT-CUP 2012. SWAT Calibration and Uncertainty ProgramA User Manual.

Abbaspour, K.C., Yang, J., Maximov, I., Siber, R., Bogner, K., Mieleitner, J., Zobrist, J., Srinivasan, R. (2007). Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. J. Hydrol., 333(2-4), 413-430.

Abbaspour, K.C., Johnson, C.A., Van Genuchten, M.T., 2004. Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. Vadose Zone Journal, 3(4), 1340–1352.

Allen, R.G., 1986. A Penman for all seasons. J. Irrig. Drain Eng. ASCE, 112 (4), 348–368.

Allen, R.G., Jensen, M.E., Wright, J.L., Burman, R.D., 1989. Operational estimates of reference evapotranspiration. Agron. J., 81 (4), 650–662.

Anaya, R., Wanakule N., 1993. A lumped parameter model for the Edwards Aquifer. Texas Water Resources Institute, Texas A&M University.

Anderson, K., Nelson, S., Mayo, A., Tingey, D., 2006. Interbasin flow revis-

ited: the contribution of local recharge to high-discharge springs, Death Valley, CA. J. Hydrol. 323, 276–302.

Arnold, J. G., Fohrer, N., 2005. SWAT2000: Current capabilities and research opportunities in applied watershed modelling. Hydrol. Process., 19 (3), 563–572.

Arnold, J.G., Moriasi, D.N., Gassman, P.W., Abbaspour, K.C., White, M.J., Srinivasan, R., Santhi, C., van Harmel, R.D., Van Griensven, A., Van Liew, M.W., Kannan, N., Jha, M.K., 2012. SWAT: model use, calibration, and validation. Trans. ASABE 55 (4), 1491–1508.

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, 73–89.

Bakalowicz, M., 2005. Karst groundwater: a challenge for new resources. Hydrogeol. J. 13 (1), 148–160.

Barrett, M.E., Charbeneau, R.J., 1997. A parsimonious model for simulating flow in a karst aquifer. J. Hydrol. 196 (1–4), 47–65.

Benettin, P., Velde, Y., Zee, S. E. A. T. M., Rinaldo, A., and Botter, G., 2013. Chloride circulation in a lowland catchment and the formulation of transport by travel time distributions, Water Resour. Res. 49, 4619 4632

Belcher, W.R., Bedinger, M.S., Back, J.T., Sweetkind, D.S., 2009. Interbasin flow in the great basin with special reference to the southern Funeral Mountains and the source of Furnace Creek springs, Death Valley, California, US. J. Hydrol. 369 (1–2), 30–43.

638 Bonacci, O., Ljubenkov, I., Roje-Bonacci T., 2006. Karst flash floods: an
639 example from the Dinaric karst (Croatia). Nat. Hazards Earth Syst. Sci.
640 6, 195–203.

641 Botter, G., Bertuzzo, E., Rinaldo, A., 2011. Catchment residence and travel
642 time distributions: The master equation. Geophys. Res. Lett., 38, L11403.

643 Bögli, A., 1980. Karst Hydrology and Physical Speleology. Springer, Berlin.

644 Böttcher, M.E., 1999. The stable isotopic geochemistry of the sulfur and
645 carbon cycles in a modern karst environment, Isot. Environ. Health Stud.
646 35 (1–2), 39–61.

647 Brenner, S., Coxon, G., Howden, N. J. K., Freer, J., Hartmann, A. (2018).
648 Process-based modelling to evaluate simulated groundwater levels and fre-
649 quencies in a Chalk catchment in south-western England. Nat. Hazards
650 Earth Syst. Sci., 18, 445–46.1

651 Danapour, M., Højberg, A.L., Jensen, K.H., Stisen, S. (2019). Assessment
652 of regional inter-basin groundwater flow using both simple and highly pa-
653 rameterized optimization schemes. Hydrogeol. J. 27, 1929–1947.

654 Dar, F. A., Perrin, J., Ahmed, S., Narayana, A. C., 2014. Review: carbonate
655 aquifers and future perspectives of karst hydrogeology in India. Hydrogeol.
656 J. 22 (7), 1493–1506.

657 Doerfliger, N., Jeannin, P.-Y., Zwahlen, F., 1999. Water vulnerability assess-
658 ment in karst environments: a new method of defining protection areas
659 using a multi-attribute approach and GIS tools (EPIK method). Environ.
660 Geol. 39 (2), 165–176.

661 Drew, D., Hötzl, H., 1999. Karst Hydrogeology and Human Activities. Im-
662 pacts, Consequences and Implications, Balkema, Rotterdam.

663 Droogers, P., Immerzeel, W.W., Lorite, I.J., 2010. Estimating actual ir-
664 rigation application by remotely sensed evapotranspiration observations.
665 Agric. Water Manage. 97 (9), 1351–1359.

666 Ford, D.C., Williams, P.W, 2007. Karst Hydrogeology and Geomorphology.
667 John Wiley & Sons, Chichester.

668 Franco, A.C.L., Bonumá, N.B., 2017. Multi-variable SWAT model calibration
669 with remotely sensed evapotranspiration and observed flow. Brazilian J.
670 Water Resour. 22, e35.

671 Gassman, P. W., Reyes, M., Green, C. H., Arnold, J. G., 2007. The Soil and
672 Water Assessment Tool: historical development, applications, and future
673 research directions. Trans. ASABE, 50 (4), 1211–1250.

674 Goldmann, A., 1986. Die Wasserbilanz der Rhumequelle. Bericht Wasser-
675 wirtschaftsamt Göttingen.

676 Goldscheider, N., 2005. Karst groundwater vulnerability mapping: applica-
677 tion of a new method in the Swabian Alb, Germany. Hydrogeol. J. 13 (4),
678 555–564.

679 Goldscheider, N., Drew, D., 2007. Methods in Karst Hydrogeology. Taylor &
680 Francis, London.

681 Grimmelmann, W., 1992. Hydrogeologisches Gutachten zur Bemessung und

682 Gliederung des Trinkwasserschutzgebiets Pöhlder Becken. Niederschsis-
683 chen Landesamtes für Bodenforschung, Hannover.

684 Gun, J., 1986. A conceptual model for conduit flow dominated karst aquifers.
685 In: Günay, G., Johnson, A.I. (Eds.), Proc. Ankara Symp. "Karst water
686 resources", July 1985, IAHS Publ. 161, 587–596.

687 Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition
688 of the mean squared error and NSE performance criteria: Implications for
689 improving hydrological modelling. J. Hydrol., 377(1-2), 80–91.

690 Haggett, P., 2002. Encyclopedia of World Geography. Vol. 12. Germany, Aus-
691 tria and Switzerland. Second ed. Marshall Cavendish Inc., New York.

692 Haase, H., Schmidt, M., Lenz, J., 1970. Der Wasserhaushalt des Westharzes-
693 Hydrologische Untersuchungen von 1941–1965. Veröff. Nieders. Inst. Lan-
694 deskunde Landesentwicklung Univ. Göttingen, A91, 1–96.

695 Hartmann, A., Barberá, J.A., Andreo, B., 2017. On the value of water quality
696 data and informative flow states in karst modelling. Hydrol. Earth Syst.
697 Sci. 21, 5971–5985.

698 Hartmann, A., Gleeson, T., Rosolem, R., Pianosi, F., Wada, Y., Wagener,
699 T., 2015. A large-scale simulation model to assess karstic groundwater
700 recharge over Europe and the Mediterranean. Geosci. Model Dev., 8 (6),
701 1729–1746.

702 Hartmann, A., Mudarra, M., Andreo, B., Marn, A., Wagener, T., and Lange,
703 J. (2014). Modeling spatiotemporal impacts of hydroclimatic extremes on

groundwater recharge at a Mediterranean karst aquifer. Water Resour. Res., 50, 6507–6521.

Hartmann, A., Wagener, T., Rimmer, A., Lange, J., Brielmann, H., Weiler, M., 2013. Testing the realism of model structures to identify karst system processes using water quality and quantity signatures. Water Resour. Res. 49, 3345–3358.

Hartmann, A., Goldscheider, N., Wagener, T., Lange, J. Weiler, M., 2014. Karst water resources in a changing world: review of hydrological modeling approaches. Rev. Geophys. 52, 218–242.

Herrmann, A., 1969. Die geologische und hydrogeologische Situation der Rhumequelle am Südharz. Jh. Karst-u. Höhlenkunde 9, 107–112.

Howard, A.D., 1963. The development of karst features. Bull. Nat. Spel. Soc. 25, 45–65.

Immerzeel, W.W., Droogers, P., 2008. Calibration of a distributed hydrological model based on satellite evapotranspiration. J. Hydrol. 349 (3–4), 411–424.

Jung, M., Reichstein, M., Ciais, P., Seneviratne, S.I., Sheffield, J., Goulden, M.L., Bonan, B., et al., 2010. Recent decline in the global land evapotranspiration trend due to limited moisture supply. Nature 467, 951–954.

Koster, E.A., 2005. The Physical Geography of Western Europe. Oxford Univ. Press, New York.

Kourgialas, N.N., Karatzas, G.P., Nikolaidis, N.P., 2010. An integrated framework for the hydrologic simulation of a complex geomorphological river basin. J. Hydrol., 381 (3–4), 308–321.

Kramm, U., Wedepohl, K.H., 1991. The isotopic composition of strontium and sulfur in seawater of Late Permian (Zechstein) age. Chem. Geol. 90 (3–4), 253–262.

Krysanova, V., White, M., 2007. Advances in water resources assessment with SWAT – an overview. Hydrol. Sci. J., 60 (5), 771–783.

Lam, Q.D., Schmalz, B., Fohrer, N., 2012. Assessing the spatial and temporal variations of water quality in lowland areas, Northern Germany. J. Hydrol. 438–439, 137–147.

LaMoreaux, P.E., Tanner, J.T., 2001. Springs and bottled waters of the world. Ancient history, source, occurrence, quality and use. Springer Verlag, Berlin

Le Moine, N., Andreassen, V., Perrin, C., Michel, C., 2007. How can rainfallrunoff models handle intercatchment groundwater flows? Theoretical study based on 1040 French catchments. Water Resour. Res. 43, W06428.

Le Moine, N., Andreassen, V., Mathevet, T., 2008. Confronting surface- and groundwater balances on the La Rochefoucauld-Touvre karstic system (Charente, France). Water Resour. Res. 44, W03403

Liersch, K. M., 1987. Zur Wasserbilanz der Rhumequelle und ihres Einzugsgebietes, des Pöhlder Becken. N. Arch. F. Nds. 36, 293–305.

747 Liu, Y.B., Batelaan, O., Smedt, F.D., Huong, N.T., Tam, V.T., 2005. Test
748 of a distributed modelling approach to predict flood flows in the karst
749 Suoimuoi catchment in Vietnam. Environ. Geol. 48 (7), 931–940.

750 Maier, N., Dietrich, J., 2016. Using SWAT for strategic planning of basin
751 scale irrigation control policies: a Case study from a humid region in north-
752 ern Germany. Water Resour. Manage. 30, 3285–3298.

753 Malagó, A., Efstathiou, D., Bouraoui, F., Nikolaidis, N.P., Franchini, M.,
754 Bidoglio, G., Kritsotakis, M., 2016. Regional scale hydrologic modeling of
755 a karst-dominant geomorphology: the case study of the Island of Crete. J.
756 Hydrol. 540, 64–81.

757 Molina-Navarro, E., Andersen, H. E., Nielsen, A., Thodsen, H., Trolle, D.,
758 2017. The impact of the objective function in multi-site and multi-variable
759 calibration of the SWAT model. Environ. Modell. Softw. 93, 255–267.

760 Monteith, J.L., 1965. Evaporation and environment. In: The State and Move-
761 ment of Water in Living Organisms, Proc. 15th Symposium Society for Ex-
762 perimental Biology, Swansea, Cambridge University Press, London, 205–
763 234.

764 Mu, Q., Zhao, M., Running, S.W., 2013. MODIS global terrestrial evapo-
765 transpiration (ET) product (NASA MOD16A2/A3), Algorithm Theoreti-
766 cal Basis Document, Collection 5. NASA, Washington.

767 Muthuwatta, L.P., Booij, M.J., Reintjes, T.H., Bos, M.G., Geiske, A.S., Ah-
768 mad, M. 2009. Calibration of a semi-distributed hydrological model using
769 discharge and remote sensing data. In: Yilmaz, K.K. et al. (Eds.), New

770 Approaches to Hydrologiocal Prediction in Data-Sparse Regions. IAHS,
771 Hydrabad, 52–58

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

774 Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2011. Soil and Wa-
775 ter Assessment Tool theoretical documentation version 2009. Grassland,
776 Soil and Water Research Laboratory, Agricultural Research Service and
777 Blackland Research Center, Texas Agricultural Experiment Station, Tem-
778 ple, Texas.

779 Nerantzaki, S.D., Giannakis, G.V., Efstathiou, D., Nikolaidis, N.P., Si-
780 betheros, I.., Karatzas, G.P., Zacharias, I., 2015. Modeling suspended sed-
781 iment transport and assessing the impacts of climate change in a karstic
782 Mediterranean watershed. Sci Total Environ., 538, 288–297.

783 Nguyen, V.T., Dietrich, J., 2018. Modification of the SWAT model to simu-
784 late regional groundwater flow using a multicell aquifer. Hydrol. Process.
785 32(7), 939–953.

786 Nguyen, V.T., Dietrich, J., Uniyal, B., Tran, D.A., 2018. Verification and
787 correction of the hydrologic routing in the soil and water assessment tool.
788 Water 10, 1419.

789 Nikolaidis, N.P., Bouraoui, F., Bidoglio, G., 2013. Hydrologic and geochem-
790 ical modeling of a karstic Mediterranean watershed. J. Hydrol. 477, 129–
791 138.

792 Oki, T., Kanae, S., 2006. Global hydrological cycles and world water re-
793 sources. Science 313, 1068–1072.

794 Palanisamy, B., Workman, S.R., 2014. Hydrologic modeling of flow through
795 sinkholes located in streambeds of Cane Run stream, Kentucky. J. Hydrol.
796 Eng. 20 (5), 04014066–1–12.

797 Paul, J., Vladi, F., 2001. Zur Geologie der Einhornhhle bei Scharzfeld am
798 sdwestlichen Harzrand. Ber. Naturhist. Ges. Hannover 143, 109–131.

799 Perrin, C., Michel, C., Andréassian, V., 2003. Improvement of a parsimonious
800 model for streamflow simulation, J. Hydrol., 279, 275–289.

801 Rajib, A., Evenson, G.R., Golden, H.E., Lane, C.R., 2018). Hydrologic model
802 predictability improves with spatially explicit calibration using remotely
803 sensed evapotranspiration and biophysical parameters. J. Hydrol., 567,
804 668–683.

805 Rientjes, T.H.M., Muthuwatta, L.P., Bos, M.G., Booij, M.J., Bhatti, H.a.,
806 2013. Multi-variable calibration of a semi-distributed hydrological model
807 using streamflow data and satellite-based evapotranspiration. J. Hydrol.
808 505, 276–290.

809 Rienäcker, I., 1987. Bericht zum Forschungsprogramm “Pöhlder Becken und
810 Rhumequelle”: “Hydrochemische und hydrologische Untersuchungen an
811 der Rhumequelle und in ihrem Einzugsgebiet Ein Beitrag zur Nutzung
812 des Wassers für die Trinkwasserversorgung”. Bericht EEW Duderstadt.

813 Rozos, E., Koutsoyiannis, D., 2006. A multicell karstic aquifer model with
814 alternative flow equations. J. Hydrol. 325, 340–355.

Sangrey, D.A., Harrop-Williams, K.O., Klaiber, J.A., 1984. Predicting ground-water response to precipitation. J. Geotech. Eng. 110 (7). 957–975.

Scanlon, B.R., Mace, R.E., Barett, M.E., Smith, B., 2003. Can we simulate regional groundwater flow in a karst system using equivalent porous media models? Case study, Barton Springs Edwards Aquifer, USA. J. Hydrol., 276, 137–158.

Schnug, E., Ernst., W.H.O., Kraztz S., Knolle F., Haneklaus S., 2004. Aspects of ecotoxicology of sulphur in the Harz region - a guided excursion Landbauforschung Völkenrode, 54 (3), 129–143.

Schnug E., Haneklaus S., 1998. Diagnosis of sulphur nutrition. In: Schnug, E. (Ed.), Sulphur in Agroecosysterns. Kluwer Academic Press, Dordrecht, 1–38.

Screaton, E., Martin, J.B., Ginn, B., Smith, L., 2004. Conduit properties and karstification in the unconfined Floridan Aquifer. Groundwater, 42(3), 338–346.

Spangler, L.E., 2001. Delineation of recharge areas areas for karst springs in Logan Canyon, Bear River Range, northern Utah. In: Kuniansky, E.L. (Ed.), U.S. Geological Survey Karst Interest Group Proceedings, Water Resour. Invest. Rep. 01–4011, 186–193.

Tait, J.A., Bachtadse, V., Franke, W., Soffel, H.C., 1997. Geodynamic evolution of the European Variscan fold belt: paleomagnetic and geological constraints. Geol. Rundsch., 86, 585–98.

837 Taylor, J.C.M., 1998. Upper Permian-Zechstein, In: Glennie, K.W. (Ed.),
838 Petroleum Geology of the North Seam Basic Concepts and Recent Ad-
839 vances, 4th ed. Blackwell Scientific Publications, Oxford, 174–211.

840 Taylor, C.J., Greene, E.A., 2008. Hydrogeologic characterization and method
841 used in the investigation of karst hydrology. In: Rosenberry, D.O.,
842 LaBaugh, J.W. (Eds.), Field Techniques for Estimating Water Fluxes be-
843 tween Surface Water and Groundwater, chap. 3. Techniques and Methods
844 4-D2. U.S. Department of the Interior, U.S. Geological Survey, 75–114.

845 Thürnau, K., 1913. Der Zusammenhang der Rhumequelle mit der Oder und
846 Sieber. Jb. Gewässerkunde Norddeutschlands, Bes. Mitt. 2 (6), 1–25.

847 Tóth, J., 1963.. A theoretical analysis of groundwater flow in small drainage
848 basins. J. Geophys. Res. 68 (16), 4795–4812.

849 Tuck, S.L., Phillips, H.R.P., Hintzen, R.E., Scharlemann, J.P.W., Purvis, A.,
850 Hudson, L.N., 2014. MODISTools - downloading and processing MODIS
851 remotely sensed data in R. Ecol. Evol. 4, 4658–4668.

852 Tucker, M.E., 1991. Sequence stratigraphy of carbonate evaporite basins: the
853 Upper Permian (Zechstein) of northeast England and adjoining North Sea.
854 J. Geol. Soc. 148, 1019–1036.

855 Uniyal, B., Dietrich, J., Vasilakos, C., Tzoraki, O., 2017. Evaluation of SWAT
856 simulated soil moisture at catchment scale byfield measurements and Land-
857 sat derived indices. Agric. Water Manage. 193, 55–70.

858 Venetis, C., 1969. A study on the recession of unconfined acquifers. Interna-
859 tional Association of Scientific Hydrology. Bull. Int. Assoc. Sci. Hydrol. 14
860 (4), 119–125.

861 Vervoort, R.W., Miechels, S.F., van Ogtrop, F.F., Guillaume, J.H.A., 2014.
862 Remotely sensed evapotranspiration to calibrate a lumped conceptual
863 model: Pitfalls and opportunities. J. Hydrol. 519, 3223–3236.

864 Voigt, R., Grüger, E., Baier, J., Meischner, D., 2008. Seasonal variability of
865 Holocene climate: a palaeolimnological study on varved sediments in Lake
866 Jues (Harz Mountains, Germany). J. Paleolimnol. 40, 1021–1052.

867 Wessolek, G., Kaupenjohann, M., Renger, M., 2009. Bodenphysikalische
868 Kennwerte und Berechnungsverfahren für die Praxis. Bodenökologie und
869 Bodengenese (40). University of Technology, Berlin, Germany.

870 White, W.B., 2002. Karst hydrology: recent developments and open ques-
871 tions. Eng. Geol. 65, 85–105.

872 White, K.L., Chaubey, I., 2005. Sensitivity analysis, calibration, and valida-
873 tions for a multisite and multivariable SWAT model. J. Am. Water Resour.
874 Assoc. 41 (5), 1077–1089.

875 Zhang, Y., Chiew, F.H.S., Zhang, L., Li, H., 2009. Use of remotely sensed
876 actual evapotranspiration to improve rainfallrunoff modeling in southeast
877 Australia. J. Hydrometeorol. 10 (4), 969–980.

Table 1: List of calibration scenarios and the corresponding weights in the objective function

Scenario	Calibrated variable	Weight values in the objective function
S1	Only $Q_{Rhumespring}$	$w_1 = 0$, $w_2 = 0$, $w_3 = 5$, $w_4 = 0$
S2	All Q	$w_1 = 1$, $w_2 = 5$, $w_3 = 5$, $w_4 = 0$
S3	All Q and ETa	$w_1 = 1$, $w_2 = 5$, $w_3 = 5$, $w_4 = 5$

Table 2: Selected parameter for sensitivity analysis and sensitivity ranking

Parameter	Initial range	Description	Ranking
Surface runoff and channel processes			
1) CN2	[-0.25, 0.25]	SCS runoff curve number	1
2) SURLAG	[0.05, 10]	Surface runoff lag time (days)	15
3) SOL_K	[-0.2, 0.2]	Soil hydraulic conductivity (mm/h)	18
4) SOL_AWC	[-0.2, 0.2]	Soil available water capacity	6
5) CH_K2($sub^{4-6,19,21,26}$)	[1, 15]	Riverbed hydraulic conductivity (mm/h)	3
6) CH_K2($sub^{9,11,13}$)	[10, 40]		2
Evapotranspiration and plant water uptake			
7) ESCO	[0, 1]	Soil evaporation compensation factor	7
8) EPCO	[0, 1]	Plant uptake compensation factor	13
9) REVAPMN	[0, 500]	Threshold for groundwater revap to occur	5
Snow fall and snow melt			
10) SFTMP	[-1.5, 1]	Snowfall temperature (T °C)	9
11) SMTMP	[0, 3]	Snowmelt base temperature (T °C)	8
12) TIMP	[0, 1]	Snowpack temperature lag factor	4
Groundwater and karst processes			
13) GW_DELAY	[1, 9]	Groundwater delay (days)	21
14) GWQMN	[0, 1000]	Threshold for return flow to occur	12
15) ALPHA_BF	[0, 1]	Baseflow recession constant	10
16) RCHRG_DP(sub^{15})	[0, 1]		15
17) RCHRG_DP(sub^{24})	[0, 1]	Deep aquifer percolation factor	17
18) RCHRG_DP($sub^{10,12,14}$)	[0, 1]		11
19) β	[0.7, 0.9]		19
20) $\alpha_{conduit}$	[0.05, 0.015]	Karst parameters	16
21) α_{matrix}	[0.002, 0.003]		20

CN2, SOL_K, and SOL_AWC are changed by relative change, all other parameters are changed by replacing.

All parameters are changed at the basin scale except otherwise mentioned (e.g., $sub^{9,11,13}$ means changes are only applied to subbasins 9, 11, and 13.

Table 3: Selected parameters for calibration and the best parameter values

Parameter	Scenario S1	Scenarios S2 and S3
CN2	0.06	-0.03
CH_K2($sub^{9,11,13}$)	26.06	25.65
CH_K2($sub^{4-6,19,21,26}$)	14.35	14.01
TIMP	0.48	0.89
REVAPMN	247.25	140.75
SOL_AWC	-0.11	0.05
ESCO	0.87	0.27
β	0.81	0.77
$\alpha_{conduit}$	0.0136	0.0084
α_{matrix}	0.0021	0.0023

Table 4: Model performance statistics and characteristics of the 95PPU band.

Variable	NSE	PBIAS	KGE	*p-factor*	*r-factor*
Calibration scenario S1					
$Q_{Rhumespring}$	0.75 (0.48)	-0.2 (-2.6)	0.83(0.76)	0.96	1.10
Calibration scenarios S2 and S3					
Q_{Lindau}	0.75 (0.74)	0.1 (-3.6)	0.75(0.76)	0.45	0.24
$Q_{Hattorf}$	0.58 (0.70)	-7.5 (1.3)	0.68(0.78)	0.29	0.21
$Q_{Scharzfeld}$	0.91 (0.91)	3.3 (5.0)	0.90(0.82)	0.74	0.16
$Q_{Herzberg}$	0.61 (0.67)	2.6 (5.8)	0.76(0.78)	0.36	0.18
$Q_{Kupferhütte}$	0.60 (0.70)	9.9 (8.4)	0.72(0.72)	0.39	0.25
$Q_{Pionierbrücke}$	0.54 (0.60)	4.8 (5.4)	0.73(0.77)	0.32	0.19
$Q_{Rhumespring}$	0.69 (0.62)	0.5 (-0.9)	0.79(0.80)	0.96	1.01
ETa	0.82 (0.79)	-1.0 (2.12)	0.91(0.89)	0.58	0.36

Numbers outside parentheses indicate values of the calibration period while numbers inside parentheses indicate values of the validation period.

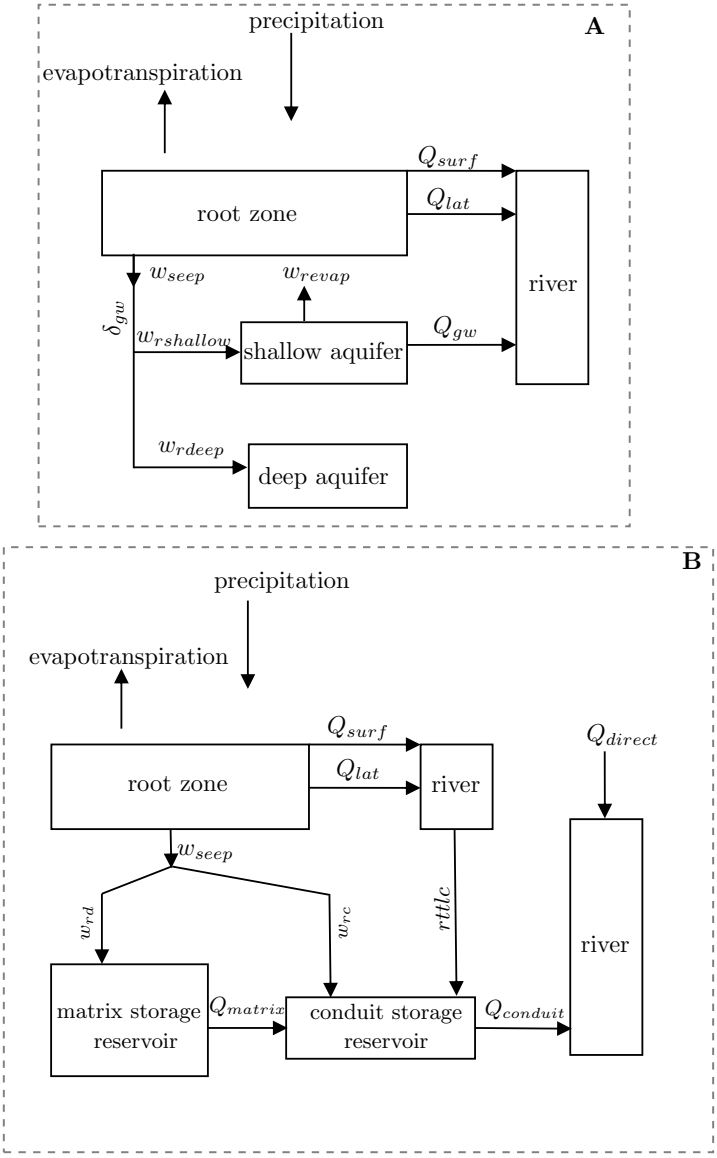

Figure 1: Conceptual models of the SWAT_IGF model. (A) the conceptual model for the non-karst area (the original conceptual model of SWAT), (B) the conceptual model for the karst area (modified from SWAT). Q_{surf} is the surface runoff, Q_{lat} is the lateral flow, w_{revap} is the groundwater revap, $w_{rshallow}$ and w_{rdeep} are the shallow and deep groundwater recharge, respectively, other variables were described in text.

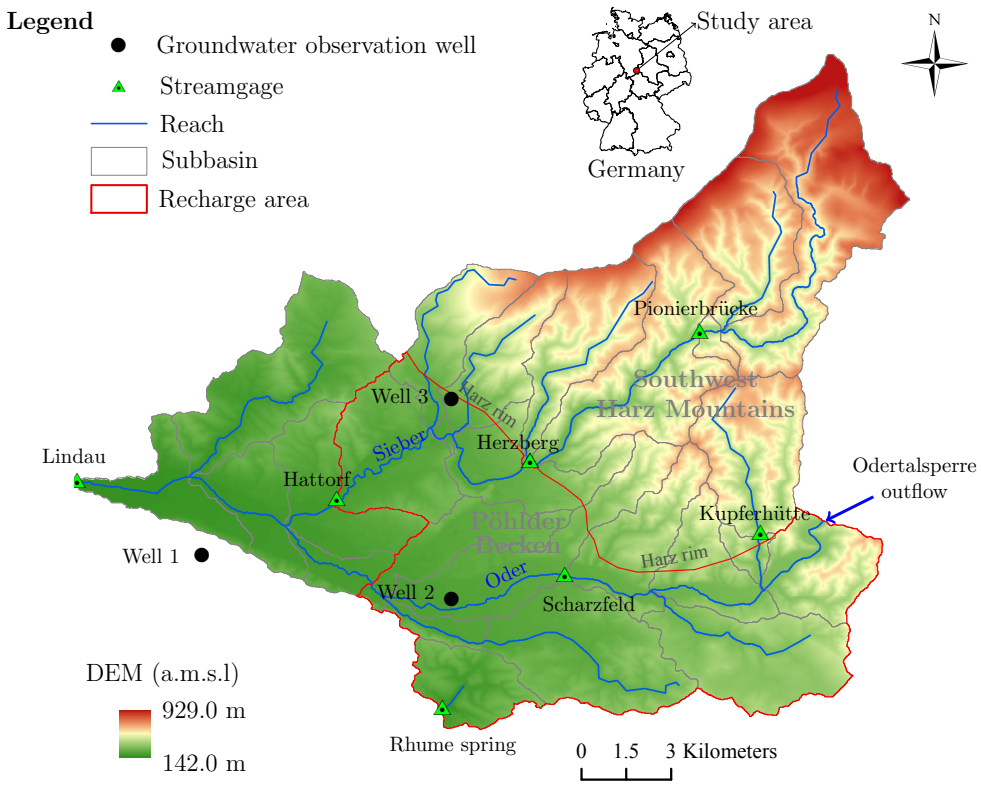

Figure 2: The study area with the Digital Elevation Model.

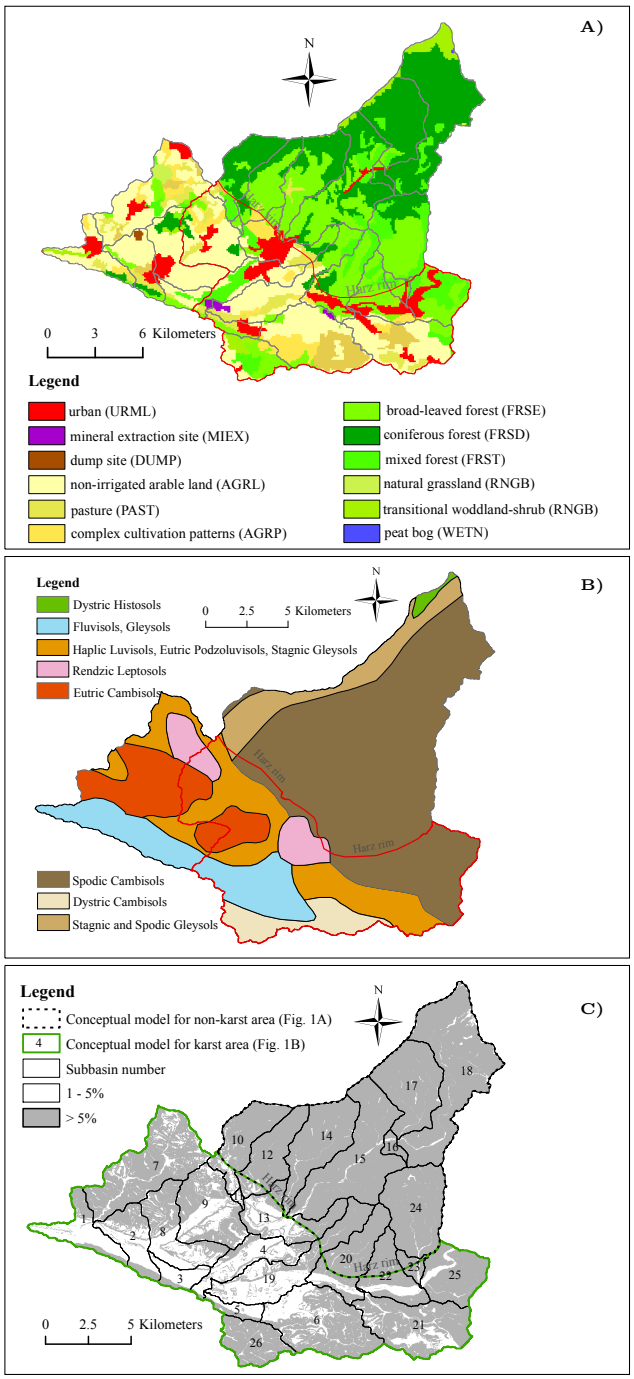

Figure 3: Distribution of (A) land use/land cover, (B) soil, and (C) slope with subbasin numbers in the study area.

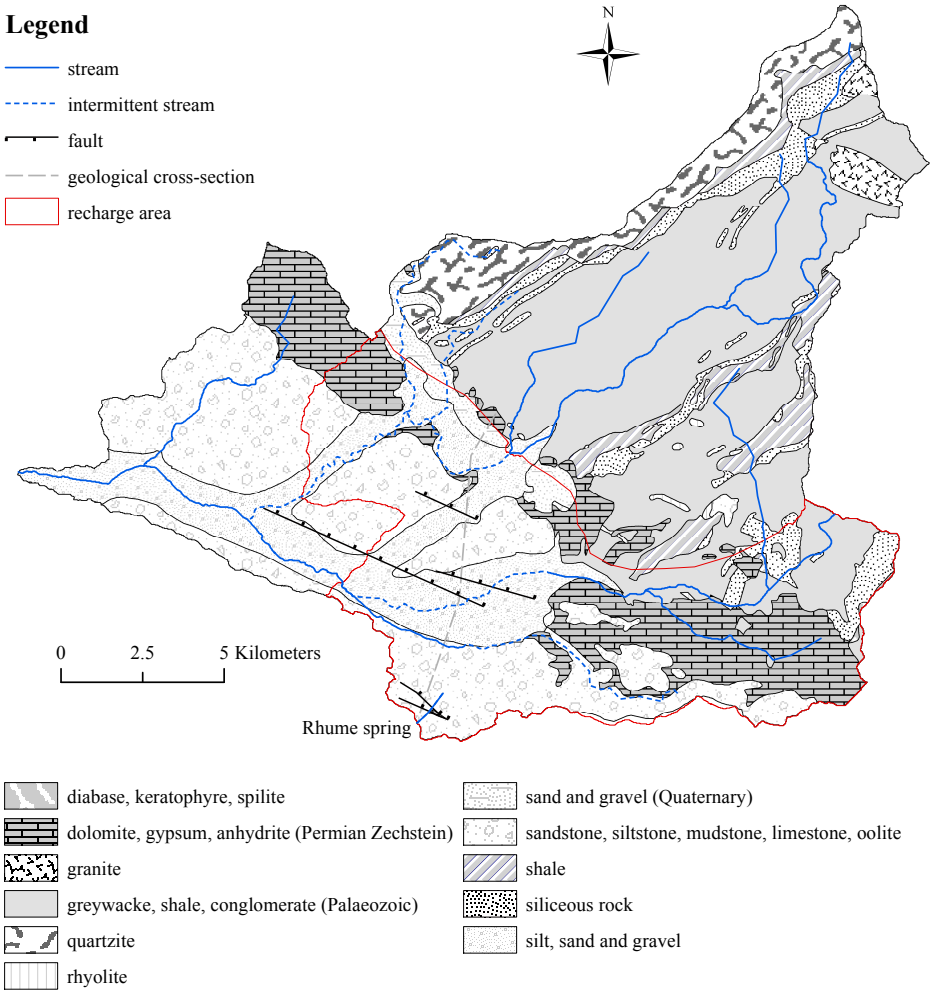

Figure 4: Geological map of the study area (BGR). Location of the faults and different types of streams were identified according to Thürnau (1913) and Grimmelmann (1992). More information about the geological cross-section could be found in Grimmelmann (1992).

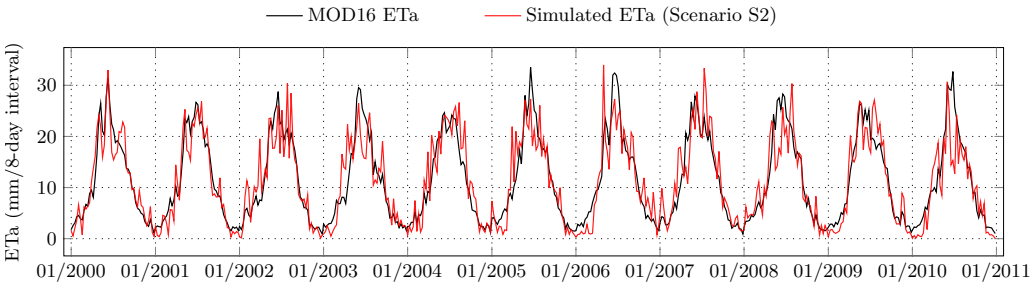

Figure 5: Time series plot of MOD16 ETa and simulated ETa from the calibration scenario S2.

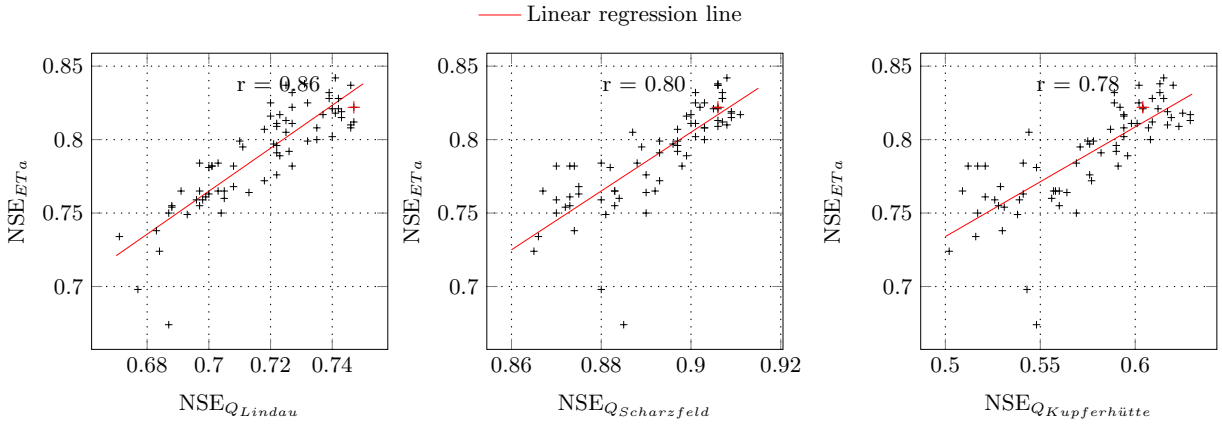

Figure 6: Scatter plots of NSE$_{ETa}$ versus NSE$_{Q_{Lindau}}$, NSE$_{Q_{Scharzfeld}}$ and NSE$_{Q_{Kupferhütte}}$ for behavioral simulations in the calibration scenario S2 (from 2000-2005). The red cross indicates the simulation corresponding to the best parameter set.

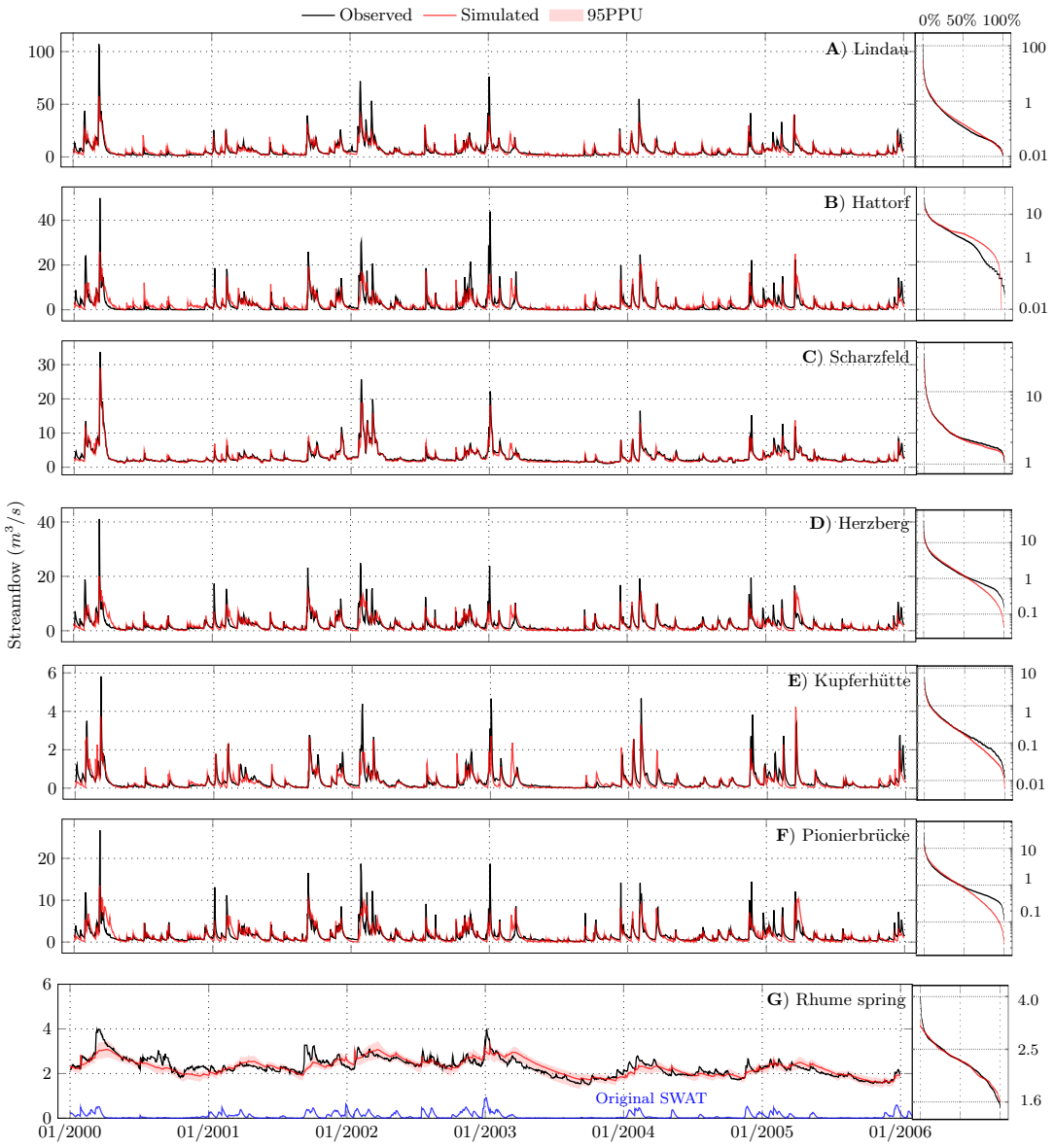

Figure 7: Times series plots of streamflows and flow duration curves (attached to the right of the respective time series plot) of the observed and the simulated streamflow from SWAT_IGF during the calibration period (2000-2005). The simulated streamflow at the Rhume spring (blue line) using the original SWAT was added in (G). For a better visualization, only the 95PPU band for the Rhume spring was shown.

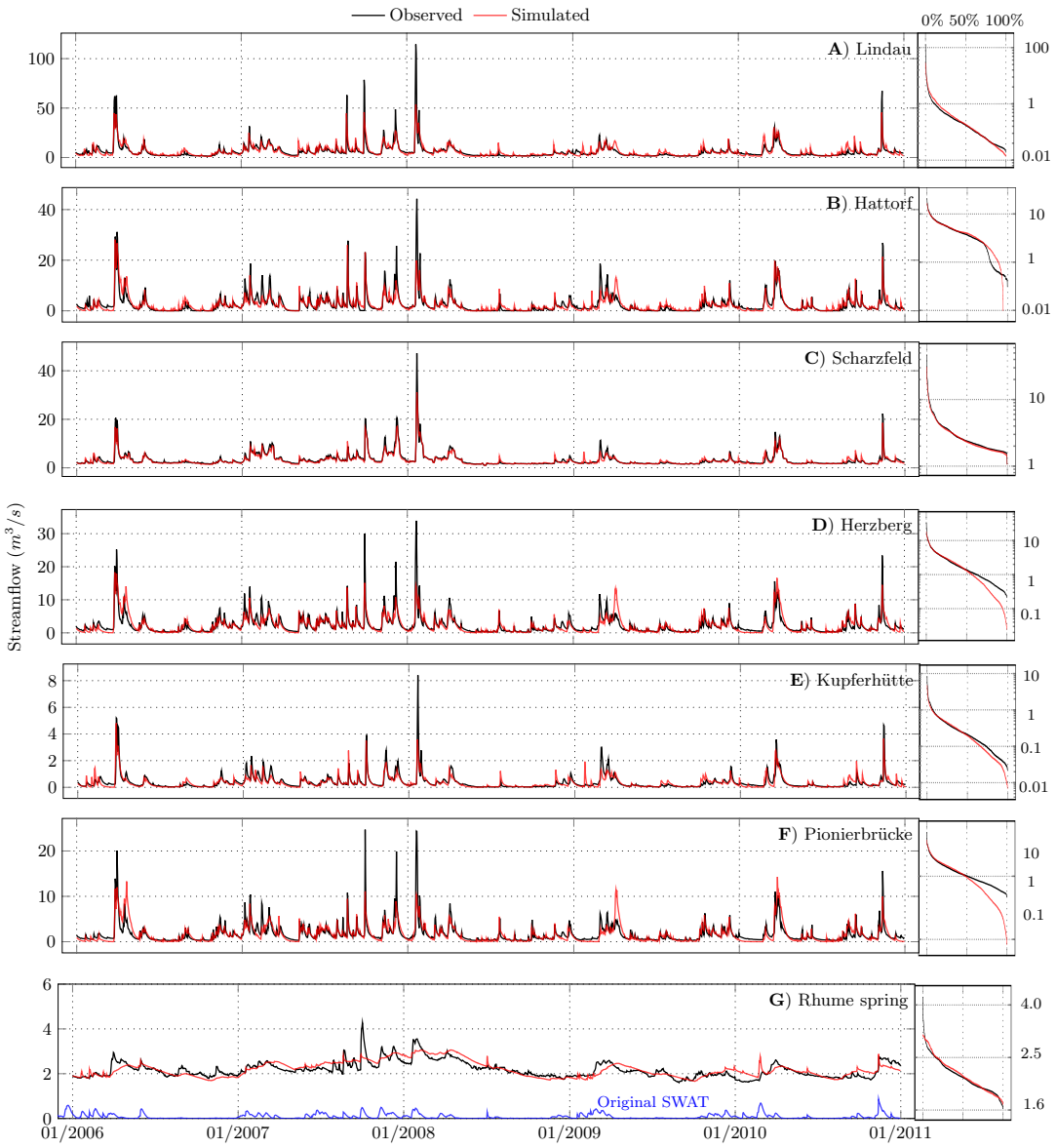

Figure 8: Times series plots of streamflows and the flow duration curves (attached to the right of the respective time series plot) of the observed and the simulated streamflow from SWAT_IGF during the validation period (2006-2010). The simulated streamflow at the Rhume spring (blue line) using the original SWAT was added in (G).

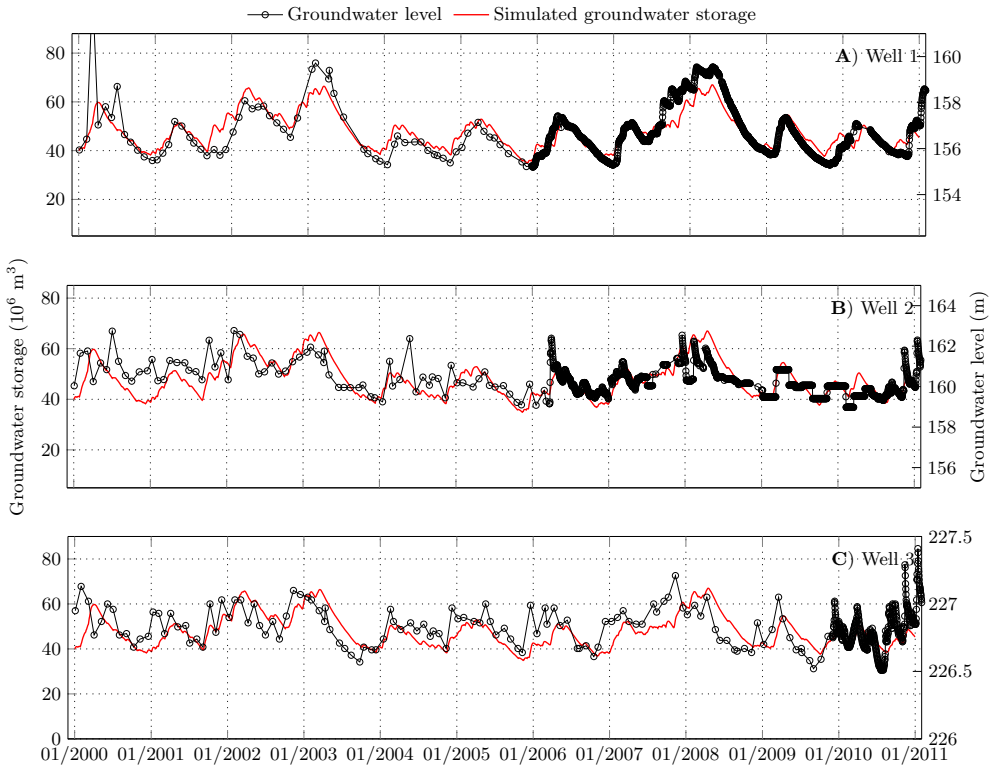

Figure 9: Variations of the simulated karst groundwater storage (the matrix and conduit storage reservoirs) and the observed groundwater levels.