An attempt of process-oriented rainfall-runoff modeling using multiple-response data in an alpine catchment, Loehnersbach, Austria

M. Johst, S. Uhlenbrook, N. Tilch, B. Zillgens, J. Didszun and R. Kirnbauer

ABSTRACT

The development of process-oriented hydrological models, which are able to simulate hydrological processes distributed in space and time, is crucial for optimal management of water resources. The model TACD (tracer aided catchment model, distributed) was modified and applied to the mountainous Loehnersbach catchment (16 km²), Kitzbueheler Alps, Austria, with the aim of simulating the dominant hydrological processes in a distributed way. It can be seen as a further developed, fully distributed version of the HBV-model with a more process-based runoff generation routine, which uses a spatial delineation of hydrological response units (HRUs). Good overall runoff simulations could be obtained for the whole catchment. Additional data, i.e. discharge from sub-catchments, snow height measurements and dissolved silica concentrations, enabled to some extent the evaluation of the simulation of single processes. Certain periods, e.g. short-term runoff fluctuations during snow melt periods, could not be simulated well even when different model modifications were executed. This indicates model shortcomings because of incomplete process understanding and the necessity for further experimental research as well as for new concepts of model structure. In particular, the understanding and mathematical description of subsurface storm flows has to be improved. The impact of different HRU delineations on discharge simulations at the catchment outlet was relatively low, as long as the direct runoff producing units remained constant. However, the impact on runoff predictions at sub-catchment scale was significant. This indicates an ‘averaging out’ effect for peculiarities and errors of runoff predictions at larger scales.

Key words | HBV model, model validation, mountain hydrology, multi-response data, process-oriented catchment modeling, TACD model

INTRODUCTION

The ability to reliably model hydrological processes at the catchment scale is essential for optimal management of water resources in mountainous catchments. This includes not only the simulation of, for instance, the daily discharge at the outlet of a catchment, but also adequate process simulation at various scales. The latter is needed to simulate the effects of environmental influences at different spatial and temporal scales. A reasonable simulation of the daily discharge can often be accomplished with a relatively simple model calibrated to observed data if the input data is adequate (e.g. Beven 2001). However, if more detailed spatio-temporally distributed simulations are required, for instance, for extensive environmental planning, more complex distributed hydrological models are required. Although nowadays a number of so-called physically based models exist (Singh 1995; Beven 2001), it is problematic to apply them to mountainous catchments with limited data. Physically based models are particularly
problematic to apply if sub-surface storm flow dominates catchment response, since our understanding of the governing processes is still incomplete (e.g. Beven 2005; Uhlenbrook 2006) and detailed observations for process-based simulations are generally lacking in mountainous environments.

Catchment models often have significant prediction uncertainties which are due to problems and errors of input data, the model's structure and parameters (e.g. Beven 2001). This causes different models or parameter sets to yield equally good simulation results (e.g. Beven & Binley 1992), and simulations of periods outside the calibration period or neighboring catchments to be very unreliable (e.g. Seibert 1999; Uhlenbrook et al. 1999). A way forward is to force the model to simulate several observed responses of the modeled system simultaneously (multi-criteria calibration; e.g. Beldring 2002). This can be achieved by calibrating the model using additional data, e.g. hydrochemical data (e.g. Mroczkowski et al. 1997), groundwater levels (e.g. Lamb et al. 1998), environmental isotopes (Seibert & McDonnell 2002) or the distribution of saturated areas (e.g. Güntner et al. 1999).

In the summer of 1987 several strong thunderstorms caused severe flood damage in the Kitzbüheler Alps. For this reason, the meso-scale Loehnersbach catchment was chosen as a research site to investigate the causes of flood generation in this region. Since 1991 hydrological and meteorological measurements as well as detailed mapping of physiographic catchment characteristics have been carried out (Kimbauer et al. 2001). In addition, sprinkling experiments (Markart & Kohl 1995) and combined tracer and geophysical investigations (Tilch et al. 2003, 2006) were carried out. The latter showed that floods are generated not only by surface flow from saturated and impervious areas but also to a large extent by subsurface flows.

The process-oriented catchment model TAC^D (tracer aided catchment model, distributed) was originally developed for the mountainous Black Forest Mountains, Germany (Ott & Uhlenbrook 2004; Uhlenbrook et al. 2004). It is a raster-based, modular catchment model, which at its core has a process-oriented runoff generation routine based on experimental results including tracer studies. This routine is based on the spatial delineation of units with the same dominating runoff generation processes which define the model structure. The model can be seen as a modified, fully distributed version of the HBV-model with a more process-based runoff generation routine.

The objectives of the presented work were, first, to apply and modify the catchment model TAC^D for the Loehnersbach catchment in such a way that is able to consider different dominant hydrological processes separately, and to assess its suitability for process-oriented modeling during various seasons. The second objective is to investigate different model structures, i.e. different spatial delineations of units with the same dominating runoff generation processes, and its effects on discharge predictions. The third objective is to evaluate additional experimental data to check the simulation of internal processes ('test of process basis of model predictions'). Therefore, the incorporation of measured runoff from sub-catchments (i.e. multi-scale data) is compared with tracer data measured at the catchment outlet as well as snow heights at a nearby meteorological station (i.e. multi-response data).

**STUDY SITE AND FIELD INVESTIGATIONS**

The study has been performed in the mountainous, meso-scale Loehnersbach catchment (16 km², elevation range: 1100–2249 m a.s.l.), located in the Kitzbueheler Alps, western Austria (Figure 1). The mean annual precipitation amounts to 1500 mm (proportion of snow about 50%), generating a mean annual discharge of approximately 1100 mm. The main river Loehnersbach is a torrent with a nival runoff regime, and it is part of the Danube drainage system. The study site is located in the Northern Greywacke region, which is characterized by a mixture of shale and sandstones which are highly susceptible to weathering. Due to previous tectonic processes and the glacial history deep fissure systems are present, which are visible in the upper parts of the catchment. The lower parts of the slopes and valley floors are covered by a thick mantle of drift, debris and boulders. The land use is dominated by forests (37%, on the shady side up to the natural tree line), pasture (mainly on the sunny side), dwarf-shrubs (in the upper regions) as well as bedrock outcrops and boulder trains.

Precipitation data was collected at 5-minute intervals during the summer at four stations (Figure 1); one of them
also measured air temperature. Additional meteorological data (see below) was available from two climate stations which are located about 3 km away. Discharge is observed at the gauging station Rammern all year round with a weir. In summer daily discharge measurements are carried out at up to eight locations using the salt dilution method.

MODEL DESCRIPTION AND INPUT DATA

General

The model TAC$^D$ (tracer aided catchment model, distributed) can be described as a modified, fully distributed version of the HBV model (Bergström 1992) with a more process-based runoff generation routine. It is a grid based, conceptual rainfall runoff model with a modular structure; it has been designed to simulate runoff, different runoff components and solute concentrations in a more process-oriented way (Uhlenbrook et al. 2004). Processes incorporated into the model include snowmelt, interception, evapotranspiration, surface storage, overland flow, different interflow processes, groundwater flow and channel routing. Compared to other conceptual runoff models TAC$^D$ focuses more on describing the processes of runoff generation. Therefore hydrological functional units have to be delineated spatially and they define the model structure.

TAC$^D$ is coded in an environmental modeling language which is part of the GIS PC-Raster (Karssenberg et al. 2001). The spatial discretization is limited by the resolution of the digital elevation model but also restricted by catchment size and computing power. In this study, it was applied using a cell size of 50 $\times$ 50 m$^2$ and hourly time steps which should be appropriate to capture the dominant runoff generation and concentration processes at catchment scale. However, small scale processes, e.g. short-term rainfall intensity variations, local overland flow generation and re-infiltration, some interception processes or detailed flood routing, cannot be represented in a fully process-based way.

Snow module, interception and soil module

The snow module and the soil module have been taken from the HBV model (Bergström 1992). To simulate snow cover and snowmelt a temperature index method is used. Precipitation is modeled as snow if the air temperature is below a certain threshold temperature $TT$ (0°C for forest, –0.2°C for other land uses). The amount of snow
is multiplied by a constant snowfall correction factor \( SFCF \) (1.05) which should account for the systematic measuring error during snowfall. Snowmelt occurs if the temperature is above \( TT \) and the amount depends on the degree-hour factor \( CFMAX \) (0.08 mm h\(^{-1}\)C\(^{-1}\)). According to Bergström (1992) a maximum of 10% of the melt water can be stored in the snow cover, and the refreezing factor is 5% of \( CFMAX \). These parameters were not changed in this study. For the actual evaporation rates of cells with snow cover constant values were assessed according to Rachner and Matthaüs (1990): 0.008 mm h\(^{-1}\) (Mar - Sept), 0.006 mm h\(^{-1}\) (Feb), 0.004 mm h\(^{-1}\) (Oct), 0.002 mm h\(^{-1}\) (Nov, Jan) and 0.0004 mm h\(^{-1}\) (Dec).

The interception module is based on the Hoyningen-Huene approach (DVWK 1996). Here, the maximum interception storage capacity \( SImax \) [mm] is a function of the leaf area index \( LAI \) [m\(^2\) m\(^{-2}\)] and vegetation coverage \( COV \) [-], which are both estimated for each plant community separately. The intercepted precipitation depends on \( COV \) and \( SImax \) in a non-linear manner, and evaporates with time depending on the potential evapotranspiration rate.

The soil module simulates percolation and storage of water as well as actual evapotranspiration (depending on soil water availability) for units of the same soil type (not for saturated areas). Depending on the parameter \( FC \) [mm], which can be related to the field capacity and a macro-porosity parameter \( BETA \) [-], actual soil moisture \( Ssm \) [mm] and recharge is calculated using a non-linear function (Bergström 1992):

\[
\text{rec} \left\{ \frac{\text{precipitation}}{\text{runoff}} \right\} = \left( \frac{Ssm}{FC} \right)^{BETA}
\]

This relatively simple concept accounts for the dominant processes at catchment scale and incorporates the fact that percolation occurs even before field capacity is reached. For each grid cell potential evapotranspiration (see chapter ‘meteorological input data’) is reduced linearly, if actual soil moisture is lower than 60% of \( FC \) (comparable to a fixed LP parameter in the HBV model).

**Runoff generation and routing module**

The concept of the runoff generation routine is based on the delineation of the catchment into units with the same dominant runoff generation process, i.e. hydrological response units (HRUs). The units are represented by different types of linear reservoirs (Figure 2). The lateral outflow \( Q \) [mm h\(^{-1}\)] depends on the storage coefficient \( k \) [h\(^{-1}\)], the actual storage content \( S \) [mm], the slope \( \beta \) [\(^\circ\)] of the specific grid cell, and the mean slope \( \beta \) [\(^\circ\)] of the respective HRU:

\[
Q = k \cdot S \cdot (1 + \tan \beta / \tan \bar{\beta})
\]

The slope term was added to Equation (2) to consider the variability of flows within each HRU. In mountainous catchments the range of slopes for some units can be relatively large. A linear dependency of the outflow of a cell on the slope (c.f. hydraulic gradient in Darcy’s law) was not
feasible, since some cells are also located in flat areas ($\tan \beta$ gets close to zero).

Table 1 shows the calibrated parameters for the soil and runoff generation routine. During the calibration process (see results section) a plausible grading of the parameter values was an important precondition. That means that material that features the highest hydraulic conductivity should show the highest $k$ parameter value in the model, e.g. the value for the $k$ parameter for the highly conductive valley floors is much higher than for the consolidated debris cover. It has to be noted, that these parameters cannot be seen directly as soil hydraulic conductivity parameters but as effective parameters which reflect a spatial differentiation of runoff generation processes.

The sequentially connected reservoirs can overflow depending on the maximum storage capacity $H$ [mm]. Vertical water flux is controlled by a percolation parameter $T$ [mm h$^{-1}$]. The different reservoirs are connected in reservoir cascades to the stream network (Figure 2). The lateral flow is simulated using the PC-Raster GIS-function ‘accufraction flux’ which means that in one time step the outflow of the higher lying cells is added to the reservoir content at the lower lying cell. Lateral drainage direction is determined by the single-flow direction algorithm (D8), which is adequate for mountainous catchments where water fluxes follow mainly the steepest gradient. To simulate runoff routing in the channel network the kinematic wave approach is used (iteration loop, 6 min time steps). The assumption that the energy slope mostly equals the riverbed slope should be appropriate for the mountainous Loehnersbach catchment with its mainly steep channels. Further model details are discussed in Uhlenbrook et al. (2004).

### Spatial delineation of the Loehnersbach catchment

The characteristics of topography, land use, soils, geomorphology and geology control lateral and vertical flow processes. To capture the spatial variability of runoff generation at the catchment scale, units with the same dominating flow processes i.e. hydrological response units (HRUs) have to be delineated in a preprocessing. This is usually done by a hierarchical overlay of different spatial data sets using GIS functions, whereas of course the final HRU map is always reflecting the kind of information as well as the spatial resolution of the input data.

For the Loehnersbach catchment three different approaches to delineate such units of the same dominating runoff generation process were followed and applied within the

<table>
<thead>
<tr>
<th>Runoff generation unit (HRU)</th>
<th>FC [mm]</th>
<th>BETA [-]</th>
<th>$k$ [h$^{-1}$]</th>
<th>$H$ [mm]</th>
<th>$T$ [mm h$^{-1}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) saturated area</td>
<td>–</td>
<td>–</td>
<td>0.3</td>
<td>10</td>
<td>–</td>
</tr>
<tr>
<td>(2) bedrock outcrops</td>
<td>20</td>
<td>0.5</td>
<td>0.3</td>
<td>20</td>
<td>0.4</td>
</tr>
<tr>
<td>(3) valley floor</td>
<td>50</td>
<td>0.5</td>
<td>0.2</td>
<td>1200</td>
<td>–</td>
</tr>
<tr>
<td>(4) non-consolid. debris</td>
<td>60*</td>
<td>4.5*</td>
<td>0.05</td>
<td>1200</td>
<td>0.3</td>
</tr>
<tr>
<td>(5) consolidated debris</td>
<td>130*</td>
<td>4.0*</td>
<td>0.005</td>
<td>1200</td>
<td>0.1</td>
</tr>
<tr>
<td>(6) fractured bedrock</td>
<td>–</td>
<td>–</td>
<td>0.008</td>
<td>2000</td>
<td>–</td>
</tr>
<tr>
<td>(7) less fractured bedrock</td>
<td>–</td>
<td>–</td>
<td>0.0005</td>
<td>1800</td>
<td>–</td>
</tr>
<tr>
<td>forest</td>
<td>195</td>
<td>2.0</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>pasture</td>
<td>115</td>
<td>2.8</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

*only valid for dwarf-shrubs.
model TACD (Figure 3): (i) an empirical, (ii) a more objective, and (iii) a partly randomly distributed one. In the empirical approach detailed mappings of lithology and hydrogeology (Pirkl 1990), soil types (Markart & Kohl 1993), plant communities (Burgstaller & Schiffer 1993) and saturated areas (Pirkl 1990) were used to delineate the different HRUs (Figure 3a). In contrast, in the more objective, GIS-based approach only generally available data sets were applied to delineate the HRUs (Figure 3b) using, for instance, remote sensing data to identify valley floor sediments and bedrock outcrops. The aim in developing this more objective approach was to find a method which allows performing a spatial delineation for the whole alpine Greywacke region, where only general available data sets are available. The partly randomly distributed spatial delineation (Figure 3c) was created to assess the model uncertainty concerning the location of the different hydrological units as well as their spatial extent. Here the location of saturated areas and valley floors remained unchanged, since this would obviously make the results arbitrary.

Seven types of HRUs were distinguished for both the empirical and the more objective approach according to Tilch et al. 2006 (simplified sketch in Figure 2), whereas the delineation of saturated areas and the two bedrock types was the same for both approaches. The HRUs for the Loehnersbach catchment were defined as:

1) Saturated areas: Saturation excess overland flow occurs on permanently wet areas (riparian zones, holes, areas nearby springs). The spatial extent of these areas was surveyed in the field under different moisture status. These maps were compared to each other as well as to a map with soil and vegetation.

Figure 3 | Spatial delineation of units with the same dominant runoff generation mechanisms (underlying bedrock units are not shown): a) empirical approach, b) more objective approach, and c) units partly randomly distributed.
types, indicating wet areas. The proportion of saturated areas varies between 2.6% and 7.2%. In this study for both the empirical and objective approach a survey with 3.7% saturated areas was used assuming that this reflects mean moisture state. Most saturated and impervious areas are directly connected to the stream network. For this unit type no infiltration and percolation to the underlying groundwater reservoir is designed (constantly exfiltrating areas).

(2) Bedrock outcrops: Infiltration excess overland flow dominates mainly in impervious areas such as bedrock outcrops or roads. These areas were derived from the vegetation type ‘pioneer vegetation’ (empirical approach) and from remote sensing data (objective approach) respectively. Since the weathering material is likely to be eroded or is at least very thin no shallow subsurface flow is considered for this spatial unit (i.e. no underlying debris cover), but deep groundwater recharge and flow is possible. The percentage in area of roads is low in the Loehnersbach catchment (0.2%). Nevertheless, to account for the direct runoff contribution from these areas some cells (usually near streams) were defined as cells with infiltration excess runoff.

(3) Valley floors: Fast subsurface flow occurs at the valley floors. These areas were derived hydrogeological map (empirical approach) and from remote sensing data (objective approach) respectively. Due to the high percentage of boulders and the frequently transported bed load a high hydraulic conductivity of the sediment can be assumed. Saturated areas located in the valley floors as well as runoff contributions to deeper groundwater were excluded for these units (groundwater discharge areas).

(4) Non-consolidated debris cover: Delayed subsurface flow is generated within non-consolidated scree and glacial drift covers mostly located at steep slopes. Here the finer material has been eroded and the soil cover is thin. Experimental investigations in a sub-catchment showed that subsurface storm flow generated in these areas contribute considerably to floods (Tilch et al. 2003, 2006). The degree of consolidation could be derived from the hydrogeological map (empirical approach). For the objective approach a specific critical slope angle of 27° could be determined which matched the empirical delineation of consolidated (<27°) and less consolidated (>27°) weathering material well.

(5) Consolidated debris cover: Slow subsurface flow occurs in consolidated debris and glacial drift covers, which are located on flatter areas e.g. on cirque valley floors.

(6) Fractured bedrock: Fast base flow is generated in highly fractured and strongly weathered bedrock below the soil and drift cover. These spatial units were delineated according to the hydrogeological mapping for the empirical approach as well as for the more objective approach.

(7) Less fractured bedrock: Slow base flow occurs within this unit. All base flow is discharging at valley floor units (see Figure 2).

Concerning the spatial percentage of the HRUs the differences between the empirical and the more objective map were low for the whole catchment. The spatial percentage (relating to all reservoir types) of the directly responding valley floors is higher for the objective (1.8%) than for the empirical delineation (0.7%), whereas the empirical map shows a higher percentage of bedrock outcrops (2.2%; objective: 1.5%).

This way of HRU delineation can be seen as a quasi 3-dimensional approach, considering that runoff generation processes at the surface can differ from those in the subsurface. The reservoirs of the runoff generation routine were arranged in two or three layers depending on the spatially distributed data sets (GIS overlay exercise; Figure 2). It has to be noted that the bottom and the middle reservoir layers both cover 100% of the area but that the upper layer is only defined for saturated areas.

**Hydrochemical simulations**

To check the temporal mixing pattern of runoff components in TAC³ concentrations of dissolved silica were simulated. This hydrochemical tracer has been used to examine runoff sources and flow pathways in different environments with silicate minerals (e.g. Rice & Hornberger 1998; Uhlenbrook & Hoeg 2003). The concentrations at a given point are calculated by assigning mean concentrations of dissolved silica to the different runoff components and using a linear
mixing approach. The mean concentrations were derived from field measurements at sites which are likely to generate mainly one specific runoff component.

Meteorological input data

The input data (hourly values) were collected at four rain gauges within the Loehnersbach catchment and two climate stations nearby (Saalbach and Schmittenhoehe). The rain gauges (one with temperature sensor) were recording in 5 minute intervals but were out of operation during winter. Therefore, for winter times daily values of precipitation were obtained from the two climate stations, located in the Saalach valley (station Saalbach; 1010 m a.m.s.l.) and on the mountain Schmittenhoehe (1973 m a.m.s.l., about 3 km away from the catchment). Since no information on the temporal patterns of precipitation events was available, the daily winter data were uniformly distributed over the day. This approach fails to represent convective storm events with highly variable precipitation intensities, but since those events are rare during winter and precipitation mostly falls as snow, the uncertainty of model results due to incorrect precipitation input during winter can be seen as low. Rainfall was corrected using an approach of Sevruk (reported in Schulla 1997) which accounts for the systematic error caused by wind. Daily values of mean wind velocity were observed at the climate stations and downscaled to hourly values. The correction of snowfall is described above. Daily snow height data was available from both climate stations.

To regionalize the precipitation the inverse distance weighting method (IDW) was combined with an elevation regression method. The latter accounts for the fact that longer events rainfall patterns are influenced by topography. The IDW method, a widely used interpolation method, computes precipitation at a specific grid cell by calculating a weighted average of the observed values within a defined radius (weight depends on distance of the station). For the mean annual precipitation at the two climate stations a mean elevation gradient of 22.5 mm per 100 m was found. Consequently a temporally constant elevation factor \( f \) was calculated according to Equation (3) which varied between 0.93 (lowest point) and 1.12 (highest point). Here the mean annual precipitation \( P \) of each cell \( x_i \) has been determined by using the mean elevation gradient starting from the mean annual precipitation at one climate station. \( P_{\text{mean}} \) represents the mean areal precipitation for the whole catchment (i.e. mean of the annual precipitation at all climate stations). Finally, 20% of the interpolated precipitation was multiplied with the elevation factor and added to 80% of the IDW-interpolated precipitation.

\[
f(x_i) = 1 + \frac{P(x_i) - P_{\text{mean}}}{P_{\text{mean}}}
\]

Since 5 minute temperature data was observed only at one rain gauge within the catchment and only in summer daily means from the surrounding climate stations needed to be taken. To downscale those daily values in hourly values a sine shaped curve with a minimum at 5:00 AM and a maximum at 1:00 PM was designed. Its amplitude was modified depending on the sunshine duration at the respective day: the longer the sunshine duration, the higher the insolation at daytime and the higher the terrestrial radiation at night, thus the larger the amplitude. Thus, the values of this sine curve varied between 3 and \(-3\) (sunshine duration over 5 hours per day) and 1 and \(-1\) (sunshine duration lower than 5 hours per day). The sine curve was added to the daily temperature data of the climate stations resulting in hourly values oscillating around the daily means. The disaggregated temperature data was finally regionalized using a temporally variable elevation gradient. If the temperature gradient was positive (higher temperature at the mountains tops than in the valley) an inversion occurred, and as no further information on these situations (about 40 days per year, mainly in December and January) was derivable the temperature value of the higher climate station was then uniformly distributed over the catchment. Furthermore, to account for shading effects hourly temperature was modified according to the elevation, exposition and surrounding topography at the specific grid cell. This was done by using the GIS based model POTRAD (van Dam 2000) which calculates the potential radiation depending on the latitude, the date and an elevation model. The disaggregated and regionalized values were verified by comparing them to the available hourly data observed within the catchment which was not used before to generate hourly temperature data.
The potential evaporation was estimated depending on temperature, relative sunshine duration and global radiation according to Turc-Wendling (DVWK 1996). Daily sunshine duration, recorded at the climate stations, was transformed to hourly relative values (range: 0 to 1) considering the length of the day and time of precipitation, and finally regionalized using the IDW method. The global radiation was calculated with the model POTRAD (van Dam 2000; see above). The actual evapotranspiration is then determined in the snow, interception and soil module depending on water availability (see above).

Hydrological data and model evaluation

Runoff at the outlet Rammern is observed at 15-minutes intervals at the same gauging station by the Technical University Vienna (TU) and the Hydrographical Service Austria (HD). As the TU data were more reliable they were used in this study. Instantaneous discharge was measured once per day during summer periods at five stream sections (sub-catchment areas: 0.8–4.6 km²; Figure 1) using the salt dilution method. Water samples were collected in one to eight hours intervals at the outlet of five sub-catchments in August 2002, to observe the change of hydrochemical parameters during floods and, thus, to investigate the changing contributions of different runoff components. The samples were analyzed for major anions (Cl, NO₃ and SO₄) and major cations (Na, K, Ca and Mg), dissolved silica and the environmental isotopes O¹⁸.

The model efficiency, $R_{eff}(Q)$ [$\bullet$ – $\circ$], (Nash & Sutcliffe 1970) and the model efficiency using logarithmic runoff values, $R_{eff}(logQ)$ [$\bullet$ – $\circ$], were used to evaluate the agreement between simulated and observed discharge. Efficiency values can be between $-\infty$ and 1.0; 1.0 indicates a perfect agreement between simulated and observed discharge. In addition, the coefficient of determination, $R^{2}$ [$\bullet$ – $\circ$], as well as the volume error [mm a⁻¹] were used as statistical measures to assess the simulations.

RESULTS

Classical model calibration and model validation

The model was applied to four hydrological years (Oct – Sep) whereby the year 1999/2000 was taken as the calibration period since it had the best input data quality. The initial parameter set was estimated according to catchment characteristics, literature values or was derived from previous TACD-applications to catchments in the southern Black Forest, Germany. To get reasonable initial storage volumes (model initialization) the year before the calibration period was modeled several times with changing parameters until the simulated runoff fits the initial runoff of the calibration period well. The corresponding spatially distributed storage volumes were taken as initial values for all calibration model runs. Since one model run takes about 30 min (computing capacity: 2.4 GHz) neither a parameter fine-tuning nor an investigation of the equifinality problem (using Monte Carlo simulations) was targeted.

The calibration was executed manually by comparing the simulated and observed discharge at gauging station Rammern visually and maximizing described statistical measures of goodness. The runoff generation parameters were calibrated based on the understanding of runoff generation mechanisms during different hydrologic conditions. For instance, the parameters of the groundwater reservoirs were calibrated by considering low flow conditions whereas the storage parameters for fast runoff components were optimized to fit high flow periods. After about 110 calibration runs a good agreement between simulated and observed discharge could be reached (Figure 4). The model was validated for three hydrologic years (1996/97, 2000/01 and 2001/02) for which during the snowmelt periods only HD discharge data was available.

The statistical measures of goodness ($R_{eff}(Q)$, $R_{eff}(logQ)$, and $R^{2}$) demonstrate good model results for the calibration and validation periods (Table 2). The lower
efficiency for the 1996/97 validation period is due to an overestimation of the simulated discharge after a high flood in July. Since, thereafter, a constant offset was measured whereas, the flow dynamics were very well simulated (as indicated by a high $R^2$ value) it seems that this flood might have caused technical problems at the gauging station which resulted in the observed offset. In the years 2000/01 and 2001/02 the number of observation values is only about a fourth of those in the calibration period due to lack of data in winter and spring. Therefore the statistics for those years represent just the goodness for the simulations in summer.

In general, the simulated discharge fits the observed hydrograph very well during low flow conditions. Also the recessions, in particular those caused by subsequent flood waves in the sub-catchments were well modeled (c.f. ‘secondary flood peaks’; Kirnbauer et al. 2001). In agreement with the results of tracer-hydrological studies (Tilch et al. 2003) such delayed recessions are the result of two differently responding subsurface flow reservoirs, which were nicely reproduced by the respective runoff components (Figure 5). Higher floods and particularly the snowmelt induced daily discharge fluctuations could not be simulated satisfactorily. However, one has to keep in mind that the uncertainty of input data is high during these periods.

Discharge simulations in sub-catchments

In order to evaluate if “… the model works well but also for the right reasons” (Klemes 1986) it is often claimed that the simulations be checked with additional data. These are e.g. measurements of snow depth, groundwater level or soil moisture (point data) as well as the observed discharge of sub-catchments and solute concentrations (more integral data). Therefore, the instantaneous discharge measurements at several stream sections were compared to the model simulations. The statistical measures of goodness calculated for four summer periods by using the empirical spatial delineation vary between 0.10 and 0.77 (Table 3).

Figure 5 shows the model result for the sub-catchment Klambach. It is demonstrated that the simulation of the flow dynamics, especially for the long recessions are satisfactory, keeping in mind that the model has not been calibrated to the runoff from sub-catchments. Concerning the flow volume as well as the peak value, it is important to note that the two floods in August 2002 were extreme (peak discharge 4.9 m³/s at Rammern) and very large flooding occurred in the whole region. No measurements could have been done for these flood peaks.

In general, for all sub-catchments the modeled discharge is often too high during low flow conditions. This might be due to an overestimation of groundwater discharge at the respective stream section. In the model all groundwater runoff at a specific stream cell is added to the stream discharge (no bypassing groundwater). This seems to be appropriate for the gauging station at the outlet Rammern, but it is not necessarily the case at the stream sections of the sub-catchments where the discharge measurements took place.

### Table 2

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>n (Qobs)</td>
<td>5850</td>
<td>6606</td>
<td>1604</td>
<td>1842</td>
</tr>
<tr>
<td>$R_{eff}(Q)$</td>
<td>0.61</td>
<td>0.72</td>
<td>0.74</td>
<td>0.83</td>
</tr>
<tr>
<td>$R_{eff}(\log Q)$</td>
<td>0.77</td>
<td>0.83</td>
<td>0.82</td>
<td>0.83</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.87</td>
<td>0.84</td>
<td>0.79</td>
<td>0.85</td>
</tr>
</tbody>
</table>

*calibration period.
Simulation of snow melt periods

The modular structure of TAC$^D$ allows an evaluation of the simulation of different processes during snow melt periods. The observed snow depth at the climate station Schmittenhoehe is compared to the simulated water equivalent of the snow cover at the respective altitude and exposition (Figure 6). Keeping in mind that the snow density is lower during snow accumulation and higher during melting periods the snow water equivalent and snow cover dynamics seem to be reasonably well simulated.

However, the highly fluctuating discharge during snowmelt periods could not be modeled satisfactorily (Figure 7). The comparison of the simulated snowmelt to the observed discharge clearly points out, that the daily fluctuations in discharge are mainly caused by snowmelt water. The melting water is likely to contribute directly to stream flow dynamics, for instance as surface runoff, fast interflow near the surface or through groundwater displacement mechanism (piston-flow) which are generally not well understood in hillslope hydrology (c.f. Uhlenbrook 2006). However, despite the daily fluctuations of the simulated snowmelt water the simulated discharge responses at the outlet Rammern are too smoothed. At the beginning of May the simulated discharge even develops opposed to the observed trend. This clear model failing in spring is caused by too long a storage of snowmelt water in the basin. The melting water has to go through the runoff generation routine (reservoir cascades) before it flows into the stream. Thus, by calibrating the snow parameters or applying different spatial delineations no significant improvement could be reached.

Simulation of dissolved silica as a tracer

To test the temporal mixing pattern runoff components the simulated and observed concentrations of dissolved silica were compared at five sub-catchments for a period in August 2002 (Figure 8). Two small floods caused by convective rainfall events occurred during a long flow recession after the two extreme floods in early August. The model simulates a distinct concentration decrease for both flood peaks while the measurements show no concentration decrease at all for the first peak. This seems to be due to a too highly simulated discharge at the gauging station Rammern and all sub-catchments, caused by an overestimation of the areal precipitation of this event. But

Table 3 | Statistical measures of goodness for the discharge simulations of sub-catchments during the summers 1997, 2000, 2001 and 2002 (mean for all summers)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_{df}(Q)$</td>
<td>0.73</td>
<td>0.65</td>
<td>0.30</td>
<td>0.43</td>
<td>0.32</td>
</tr>
<tr>
<td>$R_{df}(\log(Q))$</td>
<td>0.61</td>
<td>0.39</td>
<td>0.29</td>
<td>0.10</td>
<td>0.35</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.77</td>
<td>0.75</td>
<td>0.63</td>
<td>0.46</td>
<td>0.36</td>
</tr>
</tbody>
</table>

Figure 6 | Observed snow depths compared to the simulated water equivalent of the snow cover.

Figure 7 | Simulated and observed discharge at the gauging station Rammern compared to the simulated snowmelt (output of the snow module) of the whole basin.
the difference might also be due to the fact that the first peak is mainly generated by pre-event water while in the model TACD also event water caused the flow increase (possible model concept shortcoming). A further difference between the modeled and observed concentrations is the behavior of the sub-catchments compared to each other. The absolute silica concentrations differ much more among the sub-catchments than was simulated. This indicates that the data base for estimating the silica concentrations at the different units is not sufficient to capture the hydrochemical heterogeneity at catchment scale. However, the relative concentration changes during the second event were modeled reasonable well. This demonstrates that the temporal mixing pattern of different runoff components during events is not generally wrong, if good input data for an adequate runoff simulation is available.

Simulations using different spatial delineations

Three different approaches to delineate the hydrological units were executed and applied without re-calibrating the model: an empirical, a more objective and a partly randomly distributed approach (Figure 3). The simulated discharges at gauging station Rammern were not as different as expected. However, the differences for some sub-catchments were much more distinct resulting from larger differences in the areal percentage of the hydrological units in the respective sub-catchment. For instance, the sub-catchment Klammbach responses vary differently for the three spatial delineations (Figure 9). Table 4 shows the areal percentages of the units in the Klammbach sub-catchment and the respective statistical measures of goodness for the runoff simulation. The objective as well as the partly random spatial delineation yielded the best simulation results for summer 2002. These results show that for the whole catchment the differences concerning the location of the hydrological units seems to be averaged out whereas the differences can be high for the sub-catchments (see also Figure 6).
DISCUSSION

In general, the discharge at the catchment outlet (gauging station Rammern) could be simulated successfully. Here, all statistical measures of goodness were around 0.8 (Table 2), which demonstrates the general suitability of TACD for this alpine environment. It is clear that similar statistical measures for the runoff simulation could be also obtained with a simpler model approach (e.g. Beven 2001), but showing this was not the aim of this study. The model efficiency was worse for the simulated discharge at the sub-catchments since the model was calibrated concerning the discharge at the main gauging station Rammern. The goodness of the runoff simulations varies for various seasons and different hydrological conditions. The flow dynamics during winter, summer and autumn in particular during low flows, flow recessions and the subsequent flood waves could be simulated well, but highly short-term fluctuations of discharge (snowmelt, high floods) could not be modeled adequately with the given model concept. Searching for the reason why the runoff reaction during floods and snowmelt is smoothed the model has been modified in four ways. Firstly, the percentage of directly contributing areas was maximized, which means the saturated areas were delineated using a soil map and defining all gley soils as areas with saturation overland flow (percentage in area: 7.2%). Secondly, the spatial delineation of HRUs was modified aiming at taking the numerous erosion channels into account which are common in the upper regions of the Loehnersbach catchment. Runoff that is generated on these steep areas flows very quickly to the next stream. Therefore, all grid cells between impervious areas (mainly bedrock outcrops in the upper regions) and the next downhill stream cell were also defined as impervious (storage type 2) and so they were able to contribute directly to flood runoff. These two attempts to extend fast runoff generating areas led to a more accurate discharge simulation for large events but otherwise to an overestimation of surface flow during small floods. Furthermore, in doing so the measured daily discharge fluctuations during snowmelt periods were not captured. A third approach to accelerate fast responding processes was the increase of the direct flow controlling model parameters. The storage coefficients \( k \) for the non consolidated debris cover was raised up to \( 0.1 \text{ h}^{-1} \) and for the saturated and impervious areas up to \( 1 \text{ h}^{-1} \). Thus the daily fluctuations were more pronounced but the recessions became much worse. Finally, as a fourth modification the lateral subsurface flow velocity of units with non consolidated debris cover was increased. Besides the parameter values the flow velocity in the GIS-based model TACD generally depends on the spatial and temporal resolution in the used GIS environment. A grid size of \( 50 \times 50 \text{ m}^2 \) and hourly time steps allows for a maximum lateral flow velocity of \( 0.02 \text{ m s}^{-1} \) (assuming diagonal flow through the cell), which seems too slow for surface or near-surface runoff in steep areas in this mountainous catchment. Increasing the lateral flow velocity by enabling the passage of up to four grid cells within one model time step (maximum flow velocity of \( 0.08 \text{ m s}^{-1} \)) did not lead to drastic improvements for the simulation of snow melt periods.

The model results for different spatial delineations of HRUs did not vary significantly for the whole catchment since the percentages of the different units were relatively similar. Even a different location of moderately or slowly responding HRUs did not effect the discharge prediction widely, which is likely due to an averaging-out-effect. However, at the sub-catchment scale (Figure 9) the differences for the various spatial delineations were much higher analogically to the higher differences in areal percentages of the HRUs. The location of the directly flood producing areas (i.e. areas with overland flow and fast subsurface storm flow) is always important, as if they become disconnected from the channel network a very different runoff

<table>
<thead>
<tr>
<th>Spatial delineation</th>
<th>Empirical</th>
<th>Partly random</th>
<th>Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>More consolidated debris</td>
<td>8.9%</td>
<td>8.9%</td>
<td>41.7%</td>
</tr>
<tr>
<td>Less consolidated debris</td>
<td>88.6%</td>
<td>88.6%</td>
<td>56.8%</td>
</tr>
<tr>
<td>Bedrock outcrops</td>
<td>1.8%</td>
<td>1.8%</td>
<td>1.0%</td>
</tr>
<tr>
<td>Valley floor</td>
<td>0.7%</td>
<td>0.7%</td>
<td>0.5%</td>
</tr>
<tr>
<td>( R_{eff}(Q) )</td>
<td>0.63</td>
<td>0.91</td>
<td>0.90</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.83</td>
<td>0.91</td>
<td>0.92</td>
</tr>
</tbody>
</table>
response can be expected. It could also be shown, that besides the calibration of the model parameters an adjustment of the spatial distribution of hydrological units may yield better process-oriented model results, in other words, more realistic simulation of the physical processes. In doing so, additional data like runoff data of sub-catchments or tracer data were helpful to evaluate the model results as the simulation of internal variables could be proved.

However, the application of the TACD model to the Loehnersbach catchment demonstrated also the limitations of using additional data sets for an extensive validation of the model results. Data was only available as point measurements (i.e. snow height) which bears the problems of the representativeness and the regionalization of the measurements, or the data is available at locations where a measurement was possible only for short periods because of technical reasons (i.e. sub-catchment discharge and hydrochemical measurement). For the period with the tracer measurements, it was a little unlucky that the hydro-meteorological input measurements seem to be quite uncertain in respect of the whole catchment in particular for the first event, which can be characterized as a short and intense convective rain storm event. In general, there is a lack of measurements during snowmelt periods and larger floods, which is due to logistical problems in such an environment. Although the Loehnersbach catchment is a quite well investigated area and a lot of process studies have been carried out at the micro-scale, only a small amount of data were suitable to evaluate the model simulations at catchment and sub-catchment scale. Therefore, in future investigations field studies and model developments should be carried out simultaneously. This warrants that data be gathered which helps directly in the model parameterization and during model evaluation.

CONCLUSIONS

The application of the GIS based model TACD to the alpine Loehnersbach catchment yielded a deeper insight into the model functioning concerning its strengths and its shortcomings. The simulation results were very good for the discharge at the catchment outlet but they varied for different system states and the five sub-catchments. Any additional data such as snow depth and tracer concentrations were helpful to evaluate the model performance concerning the simulation of internal system processes. Three different spatial delineations of HRUs have been applied to assess their sensitivity. In this case study it appears that the larger the catchment, the less important the spatial accuracy of the spatial delineation of the different HRUs whereas the location of fast responding areas were never changed.

Different model adaptations have been executed to reach a better simulation of fast runoff processes but did not lead to the targeted improvements. The maximization of fast responding HRUs, the incorporation of erosion channels as well as the increase of lateral flow velocities yielded better fits for single high floods but not for the general runoff behavior and the very high daily fluctuations during snowmelt. Such fast responding areas as well as the valley floors can be very important for geomorphological and biogeochemical processes (McClain et al. 2003) and can serve as runoff generation hot spots within a catchment (Uhlenbrook 2006). Thus, a better physical understanding and accordingly mathematical description of the processes dominating at these areas during flood formation is needed. This is linked to a necessary higher temporal detailed observation and modeling of these processes.

Furthermore, it seems that neither an incomplete calibration of the model parameters nor the uncertainty of input data is responsible for the model shortcoming concerning the simulation of fast runoff processes. For instance, the temperature input data is not responsible for the failing runoff simulation in spring since the simulated snowmelt and the measured runoff are highly correlated. Instead it is a hint that rather the retention effect in the runoff generation routine caused by the concept of reservoir cascades along the hillslope smoothes the impulses too much (absence of piston-flow type behavior). However, the physics of these flow processes are not understood completely (e.g. Beven 2005; Uhlenbrook 2006). Thus it was concluded that before implementing a new runoff generation routine in particular for snow melt periods deeper insights of dominating mechanisms at catchment scale are needed. This illustrates well the necessary loop of model development (discussed e.g. in Singh 1995; Beven 2001): field investigations, perceptual description, mathematical formulation, model application, critical model
evaluation, next iteration with new specific field investigation to improve the model where necessary and so on.

Given the fact that a catchment is not a static system e.g. soil hydraulic parameters and the drainage network change depending on the actual moisture state, future conceptual rainfall runoff models at catchment scale should allow for a variable parameterization or a switch-over to different model structures depending on the actual system state. This might help to improve the above illustrated shortcomings and would lead to a more accurate simulation of specific runoff processes.

ACKNOWLEDGEMENTS

We are all very grateful to prof. Dr Christian Leibundgut (University of Freiburg, Germany) who chaired a research programme on runoff generation processes and modeling (period: 2000–2004), under which the presented research results were initiated. The authors thank the German research foundation (Deutsche Forschungsgemeinschaft, DFG, Bonn, Germany) for financial support, grant no. Le 698/12-1. Special thanks to Dr. Bruno Merz (GeoForschungsZentrum Potsdam, Germany) for fruitful discussions during earlier stages of the model application. The empirical spatial delineation was widely based on the hydro-geological map of Dr. Herbert Pirkl (Geocological consulting, Vienna, Austria). Thanks to the two anonymous reviewers for providing helpful and very detailed comments.

REFERENCES


Klemes, V. 1986 Dilettantism in hydrology: transition or destiny. Wat. Resour. Res. 22(9), 177–188.


Ott, B. & Uhlenbrook, S. 2004 Quantifying the impact of land-use changes at the event and seasonal time scale using a process-oriented catchment model. Hydrol. Earth Syst. Sci. 8, 62–78.


Rice, K. C. & Hornberger, G. M. 1998 Comparison of hydrochemical tracers to estimate source contributions to peak

Schulla, J. 1997 *Hydrologische Modellierung von Flussgebieten zur Abschätzung der Folgen von Klimaänderungen*. Zürcher geographische Schriften, Heft 69, Geographisches Institut ETH, Zürich, Switzerland.


First received 26 May 2005; accepted in revised form 5 September 2007