Distributed conceptual modelling in a Swedish lowland catchment: a multi-criteria model assessment

Sebastian Wrede, Jan Seibert and Stefan Uhlenbrook

ABSTRACT

Operational management and prediction of water quantity and quality often requires a spatially meaningful simulation of environmental flows and storages at the catchment scale. In this study, the performance of a fully distributed conceptual hydrologic model was evaluated based on the HBV (Hydrologiska Byråns Vattenbalansavdelning) and TAC^D (Tracer Aided Catchment model – Distributed) model concept in the meso-scale Fyrisån catchment in the Central Swedish lowlands. For a more spatially explicit representation of runoff generation processes of small landscape elements such as wetlands, a new sub-grid parameterization scheme was implemented in the model. In addition, a simple flow distribution and lake retention routine was introduced to better conceptualize the flow routing. During intensive model evaluation and comparison the model underwent conventional split-sample and proxy-basin tests. In this process, shortcomings of the model in the transferability of parameter sets and in the spatial representation of runoff generating processes were found. It was also demonstrated how a detailed comparison with a lumped benchmark model and the additional use of synoptic stream flow measurements allowed further insights into the model performance. It could be concluded that such a thorough model assessment can help to detect shortcomings in the spatial representation of the model and help facilitate model development.

Key words | conceptual hydrological catchment models, distributed modelling, process-oriented modelling, sub-grid variability

Sebastian Wrede (corresponding author)
Stefan Uhlenbrook
Faculty of Civil Engineering and Geosciences,
Water Resources Section,
Delft University of Technology,
NL-2600 GA Delft,
The Netherlands
E-mail: *wrede@uni-trier.de*

Sebastian Wrede
Centre de Recherche Public – Gabriel Lippmann,
Department of Environment and Agro-Biotechnologies, L-4422 Belvaux,
Luxembourg; and
Department of Hydrology,
University of Trier, D-54286 Trier,
Germany

Jan Seibert
Department of Geography,
University of Zurich, CH-8057 Zurich,
Switzerland;
Department of Physical Geography and Quaternary Geology, University of Stockholm,
SE-10691 Stockholm,
Sweden; and
Department of Earth Sciences,
Uppsala University, SE-75236 Uppsala,
Sweden

Stefan Uhlenbrook
Department of Water Sciences and Engineering,
UNESCO-IHE,
NL-2601 DA Delft,
The Netherlands

INTRODUCTION

Increasing environmental threats to water resources and aquatic ecosystems call for further development of hydrologic models that are able to better quantify environmental flows at the catchment scale. However, this development has been often associated with an increase in model complexity along with the lack of observational data and appropriate diagnostic tools to further constrain and evaluate model states and outputs (e.g. Wagener *et al.* 2001; Gupta *et al.* 2008). Thus, appropriate model complexity needs to be balanced with (i) the modelling purpose, (ii) the characteristics of the hydrological system and (iii) the data available (Wagener *et al.* 2001). In addition, powerful rigorous diagnostic tests are needed to evaluate to what degree a realistic representation of the natural system has been achieved and how the given model concept can be improved (Gupta *et al.* 2008). In this work a rigorous multi-criteria model assessment was performed to evaluate a process-oriented, distributed hydrologic model that was developed to better describe the spatial variability of hydrological states and fluxes under boreal conditions in order to serve as a basis for coupled solute transport applications (e.g. Lindgren *et al.* 2007; Exbrayat *et al.* 2010).

Boreal landscapes are dominated by forest and open land with distinct small scale landscape elements such as lakes and wetlands that have a great influence on runoff generation and solute transport (e.g. Arheimer & Wittgren

doi: 10.2166/nh.2012.056

1994, 2002; Gren 1995). This mosaic of alternating landscape patches with individual characteristics needs to be addressed by the chosen model concept. However, in most cases computational limitations hinder fully resolving spatial heterogeneity. Even distributed models are to some degree spatially lumped and spatial elements are parameterized using effective parameters. These parameters are assumed to take into account spatial heterogeneity of landscape characteristics and hydrological processes within a single model element, but might not be capable of reproducing the hydrological behaviour for an element as a sum of its sub-elements. This is especially true when differences in the functioning of the different sub-element units are significant. Typical measures to account for this so-called 'subgrid variability' are commonly used in macro-scale applications (e.g. Blöschl & Sivapalan 1995). These can consist of a statistical distribution function within a model element or a process adequate areal discretization by subdividing model elements into different sub-entities (Becker & Braun 1999).

An additional element of distributed hydrologic modelling is the lateral routing of water along flow pathways (surface and subsurface) and stream flow for which different methods with varying complexity and data demand are available (e.g. Singh 1995). Boreal environments are often characterized by stream networks intersected by numerous lakes that often lack detailed geometric descriptions. Consequently, only very simple routing or flow distribution functions are employed in conceptual modelling, such as the triangular weighting function of the HBV (Hydrologiska Byråns Vattenbalansavdelning) model (Bergström 1995). Nevertheless, such simple approaches are not feasible in distributed model concepts, where more explicit weighting functions accounting for spatial routing of flow through the stream network are required.

Successful conceptual model applications depend on accurate parameterization. This is usually achieved by comparing observed and modelled stream flow at the basin outlet. This approach might be sufficient for simple lumped models, but is not a rigorous enough criterion for distributed model evaluation in order to ensure a correct representation of internal state variables (e.g. Mroczkowski *et al.* 1997). Additional information, such as groundwater levels or soil moisture measurements, is required for a sufficient multi-criteria calibration procedure, but availability of suitable data is generally poor or lacking in most real-world applications. However, multi-scale validation by including runoff series from different sub-catchments enables an advanced parameter estimation and may lead to a subsequent improvement of model consistency and performance (Sooroshian & Gupta 1995).

To date, the most popular hydrological catchment model in Scandinavia is the conceptual lumped (or semi-distributed) HBV model (Bergström 1975, 1995). It dates back to the early 1970s and since then has been subject to continuous improvements and developments resulting in different model versions, such as HBV-96 (Lindström *et al.* 1997), HBV-IWS (e.g. Hundecha & Bardossy 2004), a Nordic HBV model version (Saelthun 1996) or a version for nutrient modelling (HBV-N: Arheimer & Brandt 1998). For larger areas, such as the whole of Sweden, the HBV model has been used in nested configurations. More recently, raster based versions of the HBV model (e.g. Beldring *et al.* 2003) were also developed. In such applications model complexity needs to be balanced with input data availability (Wagener *et al.* 2001), since more complex conceptualizations naturally coincide with increased parameter and model uncertainties (e.g. Beven 2006).

Here, a fully distributed, process-oriented catchment model based on the HBV concept was evaluated for water quality applications in a meso-scale lowland catchment in Sweden with mixed land use. Research questions were: (i) how to develop an efficient method to account for the subgrid variability of land use parameters within a raster-based hydrological model; and (ii) how to integrate a simple runoff routing routine for surface water bodies (channel network and lakes) in a lowland catchment? For evaluation of the revised model concept an intensive multi-criteria model assessment including synoptic runoff data and a comparison with the standard lumped HBV model was carried out.

THE FYRISÅN CATCHMENT

For model assessment in this study the meso-scale Fyrisån catchment (Figure 1) was used. Several investigations have

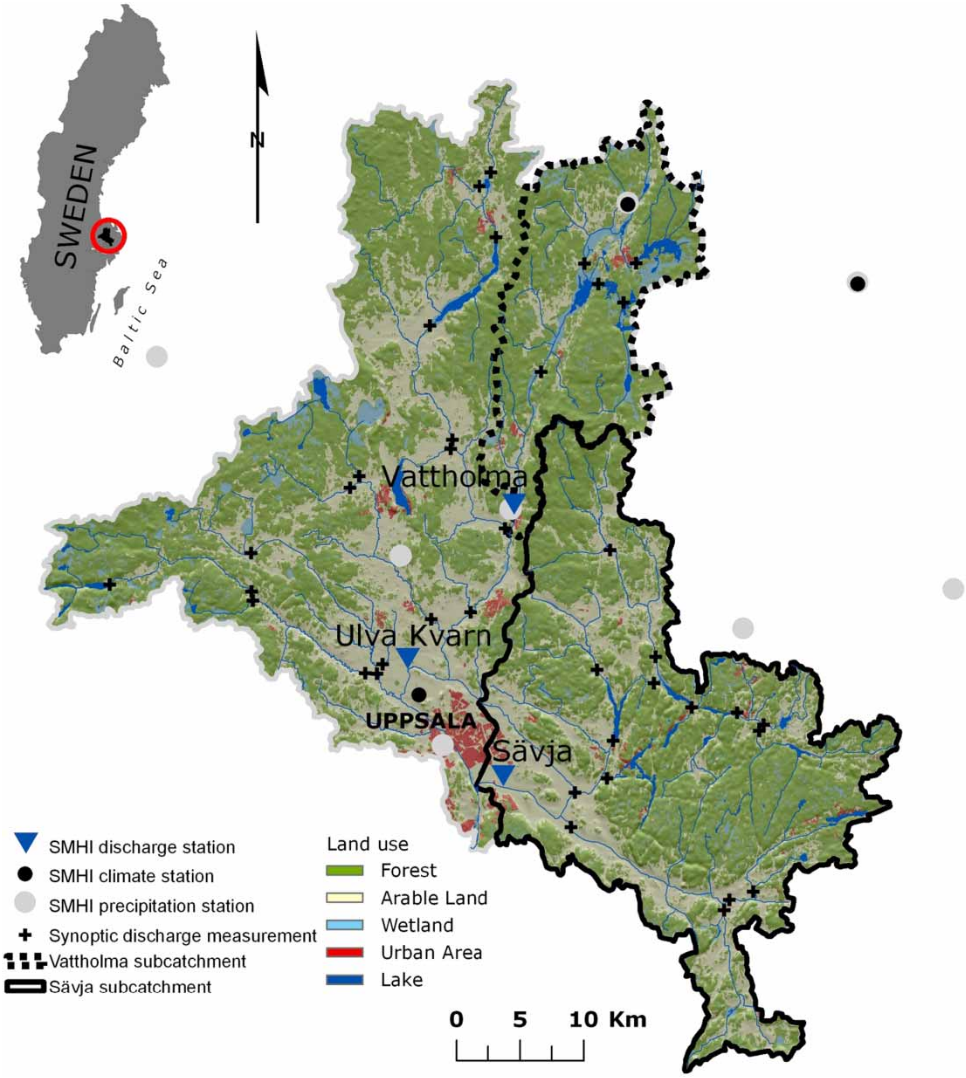

Figure 1 | Fyrisån catchment and instrumentation network.

been carried out in this research area previously, including the fundamental work of Hjulström (1993) on the morphological activity of rivers. More recent studies dealt with fluvial sediment transport and nutrient transport modelling (e.g. Darracq *et al.* 2005; Lindgren *et al.* 2007) and the application of various hydrologic catchment models (e.g. Seibert 1997, 1999; Motovilov *et al.* 1999; Xu 1999). The Fyrisån catchment is situated in the eastern part of the central Swedish lowlands, 60 km north of Stockholm. It belongs to the Mälaren–Norrström drainage basin and covers an area of approximately 2,000 km² before it discharges into Lake Ekoln. The landscape is topographically and morphologically characterized by the low-lying and flat Precambrian peneplain with most parts of the area ranging between 30 and 50 m above sea level and a highest point of 110 m. The area is mostly covered by forests (60%) of pine and spruce or a smaller fraction of mixed deciduous woodland, whereas the wide and flat river valleys in the south, particularly around the city of Uppsala, are mainly used for agriculture (32%). There are numerous wetlands of varying extent which cover about 4% of the area. Lakes constitute 2% of the landscape and are mainly small with a mean surface area of 0.4 km² (Gretener 1994). Besides the city of Uppsala, settlements (2%) are generally small in extent and scattered over the area, and only a small portion can be regarded as impervious. The distribution of predominant soil types can be roughly related to land use information. Clay soils constitute most parts of the farmland in the region, while till soils are generally covered by forest (Seibert 1999).

The research area is located within a region of relatively low annual precipitation with a corrected mean annual precipitation for the meteorological station Uppsala of 636 mm and an increase towards the west as well as the east somewhat above 700 mm. Maximum precipitation is observed in August and minimum in February and March. Snow constitutes about 20–30% to the total precipitation with an average duration of snow cover of 100–110 days per year (Seibert 1994). The mean annual temperature for the station Uppsala is +6 °C with values ranging from a maximum in June (+17 °C) to a minimum in February (–5 °C). The runoff regime shows a typical Baltic regime with a dominant primary snowmelt spring flood in April, a secondary rainfall peak flow in autumn and a low flow period during the summer months.

MODEL DEVELOPMENT

Lumped HBV model

The conceptual lumped runoff model used in this study was the HBV model (Bergström 1975, 1995). Daily discharge is simulated by using daily precipitation and temperature data as well as monthly estimates of potential evaporation as driving variables. The conceptualization of hydrological processes employs sequentially linked routines and functions representing the major processes of the land phase hydrological cycle. These routines include a snow module, simulating snowmelt with the degree-day method, a soil routine where groundwater recharge and actual evaporation are functions of actual water storage in a soil box, and a runoff generation routine where runoff from the groundwater storage is represented by linear storage equations. Channel routing is simulated via a triangular weighting function. Detailed descriptions of the basic HBV model including the governing equations can be found elsewhere (e.g. Bergström 1995; Seibert 1997).

Distributed HBV model

In this work the lumped HBV model has been extended to a fully distributed (grid based) version with a spatial resolution of $250 \times 250\,\mathrm{m}^2$. While the general model structure from the lumped HBV model remained unchanged, key elements for the spatial distribution of the model were adopted from the distributed and process-oriented catchment model TAC^D (Tracer Aided Catchment model – Distributed) (Uhlenbrook *et al.* 2004; Uhlenbrook & Sieber 2005). These elements contain the modular model structure with adapted runoff generation routines as well as the integration into the geographical information system PCRaster (Karssenberg *et al.* 2001). PCRaster offers a dynamic modelling language for raster based applications and enables lateral cell to cell routing with a single-flow direction algorithm (D8) (O'Callaghan & Mark 1984). This distributed and revised version of the HBV model, including its modifications, aims at providing a more process-oriented and spatially more explicit representation of the hydrological system under boreal conditions. It further serves as a hydrological basis for the solute transport model HBV-ND that has been used in several studies (Lindgren *et al.* 2007; Exbrayat *et al.* 2010).

In the following paragraphs two major new model elements, which were added to the previous model versions of HBV or TAC^D and were tested in the lowland Fyrisån catchment, are described in more detail. These are an approach to consider sub-grid variability and a distributed flow routing method.

Sub-grid variability

The new sub-grid variability scheme using fractions of land-use classes rather than the usual approach to assign the major class to the entire grid cell has several advantages. In particular, it is effective when the model grid cell resolution exceeds the resolution of comparatively smaller scale land use patterns or landscape features, as often found in boreal landscapes, so that their accurate representation is no longer feasible. A conventional procedure to counteract this problem consists of an increased spatial resolution by decreasing the grid cell size. However, this is usually not a practical solution as a better representation of small scale landscape features with such a high spatial resolution results in a considerable rise of model computation time. Strasser & Etchevers (2005) provided an example that necessitated an increase in grid cell resolution of elevation and meteorological input data by a factor of 64 to improve variability of snowmelt in model simulations. They argued that sub-grid parameterization methods are beneficial in order to achieve reasonable computation time and to fulfil data requirements. As such, a sub-grid parameterization scheme certainly has clear advantages. Computation time is comparably smaller than for higher spatial resolutions, as it is dependent only on the number of different land use classes or landscape features. Moreover, sub-grid parameterization is area accurate, meaning that the grid resolution has no effect on the correct representation of surface area fractions.

Similar to the approach used in TAC^D, runoff generation was parameterized separately for distinct landscape elements, such as land use types. However, this traditional approach of attributing the major land use type in a grid cell to the entire cell becomes problematic when larger grid cell sizes are used, as smaller landscape elements might not be correctly represented or even neglected. Therefore, a new sub-grid parameterization scheme was introduced, where different land use types within a single grid cell are represented by their fraction of the entire cell. In the present model, five land use classes (forest, agriculture, urban area, lakes and wetlands) are distinguished, each representing a specific runoff generation routine. The relative areas of these five classes are then used to weight flow and storage amounts and to compute mean flow and storage conditions for the entire grid cell (Figure 2). Lateral cell to cell processes include flow from the groundwater storage to neighbouring cells as well as channel routing.

In the case of the Fyrisån catchment, conventional grid aggregation led to a substantial decrease of small scale landscape features in the model representation, as illustrated by the wetland area which was underestimated by almost 50% (Table 1). This dramatic decline of the estimated wetland area can be explained by the patchy and small scale character of wetlands in this area that cannot be captured by the coarse raster structure and even results in an overestimation of catchment and forest area during grid aggregation.

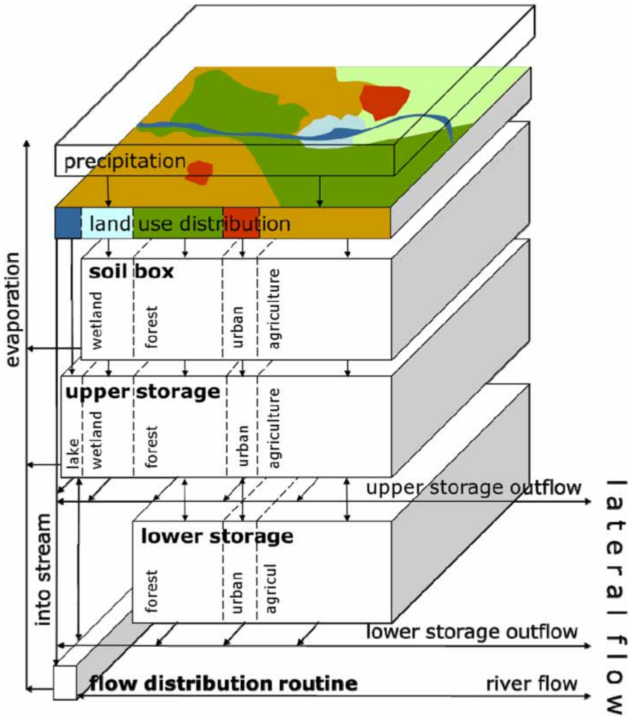

Figure 2 | Schematic representation of the sub-grid parameterization scheme in a single model grid cell. All water flows and storages within the model routines are weighted according to their designated land use class. Vertical flows indicate fluxes between different model storage compartments. Lateral flows indicate fluxes to neighbouring draining grid cells.

Table 1 | Land use distribution depending on grid resolution for the Fyrisån catchment. For the conventional aggregation the major land use class was assigned to the entire cell

	Original data set (25 × 25 m^2) and sub-grid parameterization (250 × 250 m^2)		Conventional aggregation (250 × 250 m^2)	
Land use	**km²**	**%**	**km²**	**% of original**
Catchment size	2,006	100	2,063	102.9
Forest	1,202	59.9	1,303	108.4
Agriculture	646	32.2	638	98.8
Wetlands	91	4.5	57	62.6
Settlements	34	1.7	35	102.9
Lakes	32	1.6	31	96.9

Simulation of channel routing and lakes

In distributed modelling, flow routing through channel networks and lakes plays an essential role in larger scale model applications. In the Fyrisån catchment conventional routing methods, such as the kinematic wave approach applied in the TAC^D model (Uhlenbrook *et al.* 2004), were not applicable as detailed information on channel and lake properties was lacking. Also, the application of much simpler techniques such as the MAXBAS weighting function of the HBV model was neglected in favour of a spatially more explicit representation. A simple method accounting both for spatial and temporal distribution of water flow within the channel network was needed. A relatively simple routing module was developed that computes a downstream distribution of water content for each stream grid cell, according to a parameterized triangular weighting function. Water fractions from a stream grid cell are distributed over its adjacent downstream cells (Figure 3). The form of the weighting function is dependent on two parameters DMAX [-] and DPEAK [-]. DMAX specifies the number of downstream grid cells over which the water content of the initial grid cell spreads, while DPEAK is a measure for the location of the maximum peak of the triangular weighting function, as shown in Figure 3. This conceptualization allows a more flexible parameterization of the flow routing in complex channel networks by allowing symmetric (comparable to MAXBAS in the traditional HBV model) and non-symmetric flow weighting and distribution. A daily time step in the implementation of the spatial distribution function appeared adequate to capture the temporal propagation of water flow. However, this channel flow is interrupted by numerous lakes in which flow retention occurs. Thus, these lakes are explicitly accounted for as separate storage reservoirs in the model structure and the daily stream discharge at lake inlets is attributed to these lake storages in conjunction with lateral inflows of adjacent land use cells. At the respective lake outlets, outflow is computed on the basis of a nonlinear power function depending on inflow and water storage. Afterwards, the outflow is added to the next stream section, where the downstream distribution of water continues. This implementation of flow routing in combination with the spatially explicit representation of lakes throughout the catchment allows a more process realistic conceptualization of lake retention and storage with only two parameters.

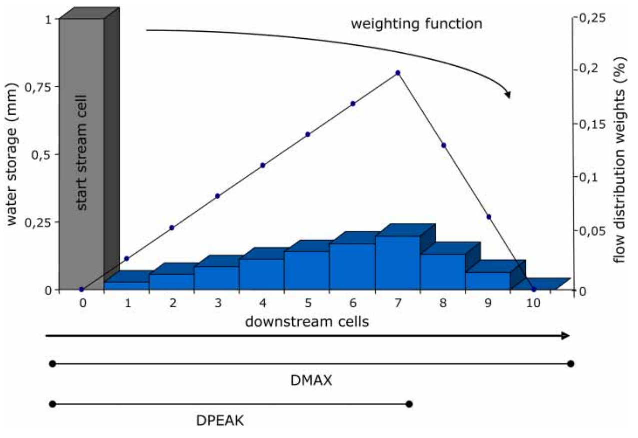

Figure 3 | Downstream distribution of stream water according to the flow distribution parameters DMAX and DPEAK of the flow routing routine.

DATA BASE AND MODELLING PROCEDURE

Spatial data

The low topographic gradient throughout the lowland Fyrisån watershed was one main concern during data preparation. For the correct delineation of sub-catchments and the local drainage network a high resolution digital elevation model (DEM) would have been beneficial, but was unfortunately not available. Therefore, digital elevation data, derived by the Shuttle Radar Topography Mission (SRTM) at a 3 arc-seconds 90×90 m resolution (SRTM-3) was used (e.g. Rabus *et al.* 2003) and resampled to the 250×250 m^2 resolution of the model domain.

The local drainage network delineation according to the D8 routing algorithm (O'Callaghan & Mark 1984) was based on the SRTM DEM data set. Due to the low topographical relief in this part of Sweden it was necessary to consider auxiliary information of a predefined stream network to force the computed drainage network through this observed stream network (Hutchinson 1989).

Information about land use, including stream network and lakes, was available as a vector data set from respective topographical maps (Lantmäteriet 2003). Unfortunately, no detailed soil map was available for the research area. However, the major land use classes can be seen as a surrogate for certain dominating soil classes; in forested areas till soils are common whereas agricultural areas correspond to clay soils (Seibert 1999). Five characteristic hydrological response units (HRUs) were differentiated (forest, agriculture, wetland, settlement, and lake) based on land use information. These five different HRUs were derived by reclassification and aggregation of 23 available land use classes and parameterized individually with the help of typical literature values about the prevailing soil characteristics, to allow for the better characterization of different major runoff generation processes in each HRU.

Meteorological and hydrological data

Meteorological and hydrological time series for the investigation period from 1992 to 2005 were obtained from the standard observation network of the Swedish Meteorological and Hydrological Institute (SMHI). Daily uncorrected precipitation data at eight stations situated within the vicinity of the research area and northwards from Uppsala were available. Moreover, three climate stations provided daily mean temperature data. Inverse distance weighting was used for regionalization of meteorological input data. Eriksson (1981) published monthly potential evaporation estimates for the whole of Sweden from which suitable values for the River Fyris were selected. The hydrological discharge observation network from SMHI contains three regular stations within the drainage basin (Figure 1), but no outlet station capturing the total drainage basin was available. The sub-catchments Vattholma (281 km^2) and Sävja (717 km^2) had continuous records over the whole application period while for the station of Ulva Kvarn (950 km^2) (including the smaller Vattholma sub-catchment) discharge measurements exist only until the year 2000. To support the model application additional data was collected by a synoptic discharge measurement campaign that was carried out during low flow at 26 locations within the Fyrisån drainage basin on 29 June 2005.

Model parameterization and calibration

In this study an automated model calibration procedure was applied by coupling the hydrologic model to the parameter estimation program PEST (Parameter ESTimation). PEST is a model-independent nonlinear parameter estimation and optimization package, frequently used for model calibration in different research fields (Doherty & Johnston 2003; Doherty 2005). It is based on the implementation of the Gauss–Marquardt–Levenberg algorithm, which combines the advantages of the inverse Hessian method and the steepest gradient method to allow a fast and efficient convergence towards the objective function minimum. The best model parameter set is selected within a specified range of parameter values by minimizing the discrepancies between model results and simulated or predefined values in a weighted least square sense.

For the automatic model calibration, 19 out of 28 model parameters were selected. The remaining parameters were fixed according to literature values or tied with an equal ratio to preceding parameters that were subject of the calibration process in order to reduce the overall parameter space and enable a fast and successful calibration result (Table 2). The relatively large amount of parameters results from the separate parameterization of runoff generation for each different land use class. A detailed sensitivity and uncertainty analysis could not be achieved in the framework of this paper as it was computationally too demanding. However, in a similar model concept with even higher parameterization, Sieber & Uhlenbrook (2005) conducted a sensitivity analysis of model parameters by using the GLUE (Generalized Likelihood Uncertainty Estimation) approach and demonstrated the plausibility of the model structure and process conceptualizations. This study, in combination with earlier HBV model applications (e.g. Bergström 1990; Seibert 1997, 1999; Uhlenbrook *et al.* 1999) as well as TAC^D model applications (e.g. Ott & Uhlenbrook 2004; Uhlenbrook *et al.* 2004; Wissmeier & Uhlenbrook 2007; Johst *et al.* 2008), guided the selection of upper and lower bounds for each parameter of the distributed HBV model. Initial parameter values were selected on the basis of previous best manual calibration trials in order to start with the optimal known parameter set. By choosing these initial parameter sets from different ranges within the parameter space, the capability of PEST to find the global minimum over local minima of the objective function by the calibration algorithm was determined.

To evaluate the performance of the obtained parameter sets in the course of the automatic calibration process, different objective functions were used as they judge the model performance by different aspects (Seibert 1999). In addition to model efficiency R_{eff} (Nash & Sutcliffe 1970) and volume error V_{E} criterions, the R_{V} criterion proposed by Lindström *et al.* (1997) were computed (Table 3). The latter was finally chosen for automatic model calibration, as it adequately

Table 2 | Model parameters with ranges and initial values used for the PEST calibration of the hydrological model

Parameter	**Explanation**	**Unit**	**Initial**	**Minimum**	**Maximum**	**Estimate**
Snow routine						
TT	Threshold temperature	°C	0	–2.5	2.5	Calibrated
TT_{diff}	TT for forest	°C	0	–2.5	2.5	Calibrated
SFCF	Snowfall correction factor	–	0.6	0.4	1	Calibrated
$SFCF_{diff}$	SFCF for forest	–	0	0.4	1	Calibrated
CFMAX	Degree-day factor	mm °C^{-1} d^{-1}	2	1	8	Calibrated
CWH	Water holding capacity	–	0.1	–	–	Fixed[a]
CFR	Refreezing coefficient	–	0.05	–	–	Fixed[a]
Soil routine						
LP	Reduction of evaporation	–	0.6	0.3	1	Calibrated
FC_{forest}	Field capacity for forest	mm	300	50	500	Calibrated
$FC_{agricul}$	Field capacity for agriculture	mm	200	50	500	Calibrated
$FC_{wetland}$	Field capacity for wetland	mm	100	–	–	Tied
FC_{urban}	Field capacity for urban	mm	150	–	–	Tied
$BETA_{forest}$	Shape coefficient for forest	–	4	1	6	Calibrated
$BETA_{agricul}$	Shape coefficient for agriculture	–	4	–	–	Tied
$BETA_{wetland}$	Shape coefficient for wetland	–	4	–	–	Tied
$BETA_{urban}$	Shape coefficient for urban	–	4	–	–	Tied
Runoff generation routine						
UrbanSplit	Portion of sealed urban areas	d^{-1}	0.5	–	–	Fixed
$K_{US\ forest}$	Upper recession coefficient for forest	d^{-1}	0.25	0.01	0.4	Calibrated
$K_{LS\ forest}$	Lower recession coefficient for forest	d^{-1}	0.005	0.001	0.15	Calibrated
$PERC_{forest}$	Percolation from upper to lower box forest	mm d^{-1}	0.05	0.001	3	Calibrated
$H_{US\ forest}$	Maximal storage capacity upper box forest	mm	350	1	1,000	Calibrated
$H_{LS\ forest}$	Minimal storage capacity lower box forest	mm	1,000	–	–	Fixed
$K_{US\ agricul}$	Upper recession coefficient for agriculture	d^{-1}	0.35	0.01	0.4	Calibrated
$K_{LS\ agricul}$	Lower recession coefficient for agriculture	d^{-1}	0.007	0.001	0.15	Calibrated
$PERC_{agricul}$	Percolation from upper to lower box agriculture	mm d^{-1}	0.008	0.001	3	Calibrated
$H_{US\ agricul}$	Maximal storage capacity upper box agriculture	mm	250	1	1,000	Calibrated
$H_{LS\ agricul}$	Minimal storage capacity lower box agriculture	mm	1,000	–	–	Fixed
$K_{US\ wetland}$	Upper recession coefficient for wetland	d^{-1}	0.05	0.01	0.4	Calibrated
$H_{US\ wetland}$	Maximal storage capacity upper box wetland	mm	150	–	–	Calibrated
$K_{US\ urban}$	Upper recession coefficient for urban	d^{-1}	0.5	0.01	0.4	Calibrated
$K_{LS\ urban}$	Lower recession coefficient for urban	d^{-1}	0.003	0.001	0.15	Calibrated
$PERC_{urban}$	Percolation from upper to lower box urban	mm d^{-1}	0.01	0.001	3	Calibrated
$H_{US\ urban}$	Maximal storage capacity upper box urban	mm	100	1	1,000	Calibrated
$H_{LS\ urban}$	Minimal storage capacity lower box urban	mm	1,000	–	–	Fixed
Lake and flow distribution routine						
K_{lake}	Recession coefficient lake	d^{-1}	1	0.001	1	Calibrated
$ALPHA_{lake}$	Nonlinear weighting coefficient	–	1	0.001	1	Calibrated
DMAX	Flow distribution length	–	109.1	3	160	Calibrated
PEAK	Flow distribution peak location	–	81.83	–	–	Tied

[a] Bergström (1995).

Table 3 | Objective functions

Objective function	Symbol	Definition	Unit	Value for 'perfect' fit
Efficiency[a]	R_{eff}	$1-\frac{\sum_{i=1}^{n}\left(Q_{i,\mathrm{obs}}-Q_{i,\mathrm{sim}}\right)^2}{\sum_{i=1}^{n}\left(Q_{i,\mathrm{obs}}-\overline{Q_{\mathrm{obs}}}\right)^2}$	–	1
Relative volume error	V_{E}	$\frac{\sum_{i=1}^{n}\left(Q_{i,\mathrm{obs}}-Q_{i,\mathrm{sim}}\right)}{\sum_{i=1}^{n}Q_{i,\mathrm{obs}}}$	–	0
R_{V} criterion[b]	R_{v}	$R_{\mathrm{eff}}-w\lvert V_E\rvert$	–	1

[a] Nash & Sutcliffe (1970).

[b] Lindström *et al.* (1997) with weight: $w = 0.1$.

balanced model efficiency R_{eff} and relative volume error V_{E} of the model.

Model calibration against daily runoff was carried out for two sub-catchments, Vattholma and Sävja, with the inclusion of the runoff station Ulva Kvarn for validation purposes. The simulation period of this study ranged from October 1994 to June 2005, preceded by a 2 year warming-up period starting in October 1992. This period was divided into two sub-periods of 5 years, starting with the calibration period from October 1994 to September 1999 and followed by a validation period from October 1999 to June 2005.

Model validation was conducted step-wise according to the hierarchical scheme for systematic testing of hydrological simulation models proposed by Klemes (1986): In a first step, the model was calibrated for both sub-catchments individually. Afterwards, the best derived parameter set for each catchment was exchanged and used for a new model simulation in the adjacent catchment, employing the proxy-basin test procedure. In a third step, the model was calibrated on both catchments in order to determine an optimal joined parameter set and, in a last step, all derived parameter sets underwent a traditional split-sample test procedure.

RESULTS

Split-sample test results

The results of the individual catchment calibration showed satisfactory fits between measured and simulated discharge for the calibration period with R_{V} values ranging from 0.85 to 0.90. Noticeable is a strong decline during the validation period from 0.90 to 0.73 (R_{V}) for Vattholma compared to 0.85 to 0.77 (R_{V}) for Sävja (Table 4).

A similar model performance was observed for the validation period using the individual optimized parameter sets in each of the other watersheds. While R_{V} values for the first 5 years remained at a reasonable level of 0.73 and 0.76, the following 5 years showed significantly poorer statistical measures with the strongest declines for the Vattholma sub-catchment. In addition to this decrease in efficiency (R_{V}) for the second half of the time period the relative volume error increased and revealed, along with R_{V}, a systematic underestimation of the flow dynamics and volume by the model. These apparent errors of the model simulation became evident for single years 1996 and 1997 of the calibration period, but also dominated throughout the second half of the Vattholma runoff record, especially in 2000 and 2004, where years with multiple spring melt peaks prevailed. Figure 4 demonstrates the lack of the model to capture the variable runoff situation for this period with almost opposite simulations of the flow dynamics. In contrast to this, the model performances for the remaining years are satisfactory and are shown for the hydrological year 1998 in Figure 5. In this case the model is able to capture quite accurately the entire runoff dynamics on the basis of the individual Vattholma parameter set.

Proxy-basin test results

It is interesting to further evaluate the hydrologic differences between the two catchments. Therefore, efficiency values (R_{eff}) were computed as a benchmark for comparing (i) the observed specific runoff records from Sävja to Vattholma and (ii) the simulated specific runoff records from these basins as it was proposed by Seibert *et al.* (2000). Results are listed in Table 5 and reveal greater similarities (higher R_{eff}) between the model generated time series than between the measured specific discharges (lower R_{eff}). It is also apparent that the model performed better (Table 4) than the benchmark of simply transferring the specific discharge between the two catchments (Table 5).

The aspect of model parameter dependency on individual catchments was further tested with simultaneous calibration on both catchments to derive a joined parameter

Table 4 | Overview of statistical performance measures for split-sample and proxy-basin test results. Simulation performance for data used in the calibration period is shown in **bold**, for the validation period in regular font: **calibration period (01.10.94–30.09.99)**; validation period (01.10.99–30.06.05)

		Vattholma		**Sävja**		**Ulva Kvarn**
Catchment used for calibration	**Objective function**	**1994–1999**	**1999–2005**	**1994–1999**	**1999–2005**	**1994–1999**
Vattholma (281 km^2)	R_V	**0.90**	0.73	0.73	0.63	0.79
	R_{eff}	**0.90**	0.73	0.75	0.65	0.81
	V_E	**0.00**	0.03	0.17	0.16	0.18
Sävja (717 km^2)	R_V	0.76	0.54	**0.85**	0.77	0.83
	R_{eff}	0.78	0.57	**0.85**	0.78	0.83
	V_E	0.19	0.25	**0.02**	0.06	0.01
Vattholma & Sävja (joined parameter set)	R_V	**0.83**	0.63	**0.83**	0.77	0.82
	R_{eff}	**0.84**	0.64	**0.84**	0.78	0.83
	V_E	**0.09**	0.13	**0.10**	0.07	0.11

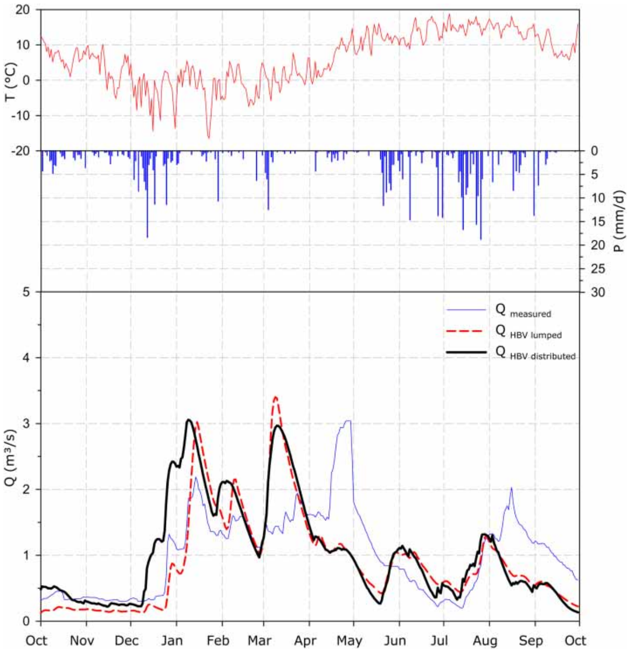

Figure 4 | Hydrological year 2000, which was the year with the worst model fit for Vattholma.

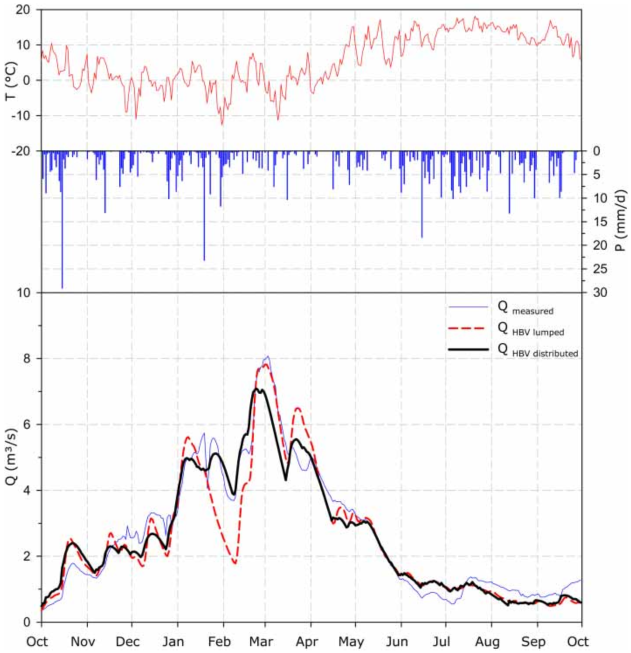

Figure 5 | Hydrological year 1998, which was the year with the best model fit for Vattholma.

set. This was achieved by assigning the same weights (50%) to the objective functions for each catchment. It led to a general decrease of model performance for the calibration phase compared to the results obtained by calibration on individual catchments alone, but provided an increase of the overall model performance for the whole application period. Validation also included the runoff station Ulva Kvarn with independent runoff records of the calibration phase. When comparing the statistical performance measures for all these catchments in Table 4, it becomes evident that the overall best fit could be achieved with the joined parameter set. This was the best parameter set that could be derived in this study for a model application covering the whole Fyrisån basin.

Table 5 | Comparison of efficiencies (R_{eff}) derived from runoff records from Vattholma and Sävja

Calculation of efficiency between Vattholma and Sävja based on…	**Resulting efficiency (R_{eff})[a]**
Observed specific runoff records	0.63
Simulated specific runoff records	0.84

[a] Nash & Sutcliffe (1970).

Model comparison

The performance of the distributed model was compared to the lumped HBV model using the same regionalized input data (precipitation and temperature as well as monthly potential evaporation estimates) for both model applications. In this case it becomes evident that the distributed, highly parameterized model was not able to outperform the simpler, less parameterized lumped HBV model in terms of runoff related efficiency measures (Table 6). For most years similar model behaviour can be observed resulting in model errors for the same years throughout all catchments (Figures 4 and 5).

Synoptic runoff measurements

Synoptic runoff measurements from a field campaign in the River Fyris allowed the distributed model simulations to be evaluated at several grid cells along the stream network during low flow conditions. These synoptic runoff measurements were used in two ways. In a first step observed and simulated runoff volumes were compared. Figure 6 illustrates that, in general, flow volumes could be reproduced reasonably well by the distributed model. The second evaluation compared observations and simulations as specific runoff values, i.e. runoff divided by the respective subcatchment areas. This latter evaluation removed the influence of catchment size and resulted in large scatter in Figure 7, indicating shortcomings of the distributed model in the simulations of spatial variations. The additional inclusion of these synoptic measurements in the calibration only partly improved the performance (Figure 7).

DISCUSSION

Split-sample test results

In general, the model performance was acceptable for calibration and validation periods for the catchments for which the model had been calibrated. Model efficiencies were smaller for the validation periods, which can partly be explained by a clear change in runoff regime from one dominated by a single spring flow to a more erratic unstable flow regime with multiple snow-melt runoff events. Unstable flow regimes have been noted during former model applications in this region (Motovilov *et al.* 1999) and have been further investigated by Krasovskaia & Gottschalk (1992) for Scandinavian countries.

Proxy-basin test results

The proxy-basin test revealed additional shortcomings in the spatial representation of the model. Despite the inclusion of the spatially explicit, land use dependent runoff generation routine and the distributed flow and lake routing, the model was not able to capture major changes in runoff for different catchments on the basis of the individually calibrated parameter set (Table 4). This was reflected in low efficiencies (R_V) for the exchanged parameter sets, especially for the second part of the application period. Once again differences in runoff regime might come into play, but computed efficiencies (R_V) of the measured as well as simulated specific

Table 6 | Comparison of distributed vs. lumped model results obtained with joined parameter set. Results from distributed model (**bold**); results from lumped model (regular)

Objective function	**Vattholma**				**Sävja**				**Ulva Kvarn**	
	1994–1999		**1999–2005**		**1994–1999**		**1999–2005**		**1994–1999**	
R_V	**0.83**	0.86	**0.63**	0.74	**0.84**	0.86	**0.78**	0.79	**0.82**	0.84
R_{eff}	**0.84**	0.86	**0.64**	0.76	**0.84**	0.86	**0.78**	0.79	**0.83**	0.84
V_E	**0.09**	0.04	**0.13**	0.01	**0.1**	0.04	**0.07**	0.01	**0.11**	0.05

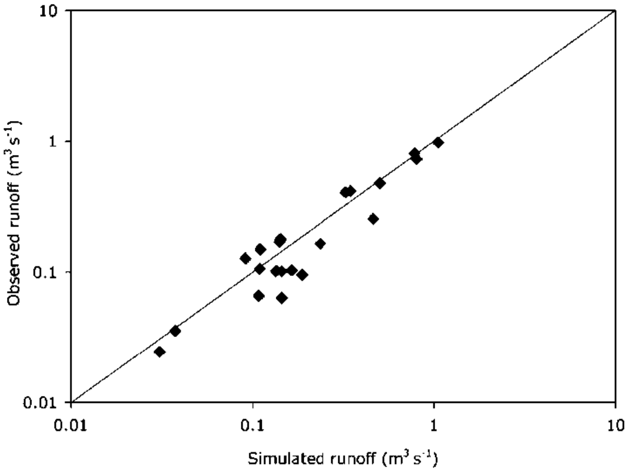

Figure 6 | Validation of the model performance by comparison of synoptic discharge measurements with simulated discharge ($r = 0.97$).

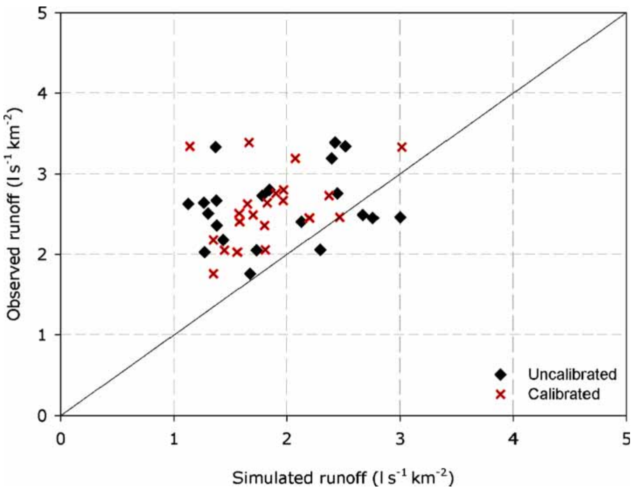

Figure 7 | Validation of the model performance by comparison of synoptic specific discharge measurements with simulated synoptic specific discharge (uncalibrated: $r = 0.22$; calibrated on synoptic measurements: $r = 0.36$).

discharges of each catchment indicate a more basic problem: while efficiencies (R_V) for the transferred parameter sets are low, signifying considerably different runoff behaviour of the two sub-catchments, the calibrated parameter sets provide much higher efficiencies (R_V) and thereby show that the diverse runoff character was not sufficiently met by the model structure. This is in line with a similar proxy-basin test study of TOPMODEL, where Donnelly-Makowecki & Moore (1999) also found to their surprise that their model that explicitly accounted for topography was not superior to a lumped model when transposing it to another catchment.

The fact that parameter sets adapted to one catchment cannot be simply transferred to the adjacent catchments, although the most important spatial processes controlling the flow regime are included in the model structure, underlines the effective character of these parameter sets that still partly incorporate regional spatial heterogeneity characteristics of each catchment. Nevertheless, it should be also stated that proxy-basin tests often result in rather drastic performance reductions and are even failed by many established models (Refsgaard & Henriksen 2004). This might also be the reason why it has been rarely applied, despite its rather informative character (Andreassian *et al.* 2009).

The simultaneous calibration of the model to both catchments supported this finding with the presence of a joined parameter set that is able to adequately capture the entire runoff hydrographs for all sub-catchments. This joined parameter set accounts for regional runoff dynamics with a slightly reduced efficiency (R_V), but performs much better on an overall basis than both previous individual parameter sets (Table 4).

Besides the model application another often applied approach for runoff predictions in ungauged basins is the transfer of specific discharge from a nearby watershed scaled by the catchment size. This alternative reveals mostly convincing results for catchments with almost identical input data (e.g. Seibert *et al.* 2000) and was compared to the prior model outputs. Table 4 reveals significantly higher efficiencies (R_{eff}) of transferred model parameters against a considerably reduced R_{eff} of the simple transfer method (Table 5). As both basins are located in close vicinity, both are subject to almost identical meteorological conditions, so that the differences in land use and lake distribution are assumed to be the reason for the better model performance. This reflects the value of model applications compared to simpler alternatives, despite the aforementioned deficits to adequately account for spatial heterogeneity.

Besides the model structure, another point in the discussion that should be considered is the applied calibration method. Simple lumped models do not suffer from high computation times, so intensive calibration procedures which necessitate extensive model runs (e.g. Monte Carlo simulations or genetic algorithms) can be easily employed. On the other hand, distributed models mostly lack efficiency in computation time and exhibit, in most cases, even higher parameterizations due to complex spatial structures. Therefore, less model runs for calibration purposes are feasible and the optimal global parameter set is not inevitably achieved every time. This effect might be underestimated, but could be clearly verified on the basis of different start parameter sets for the coupled parameter estimator PEST. It was found that the variation of initial parameters resulted in different optimized parameter sets with varying model performances.

Model comparison

By further comparing the lumped HBV model with the distributed model, it was shown that the latter did not lead to a significant improvement of the discharge simulation performance. Both models were more or less equivalent in their success in simulating discharge at the different catchment outlets and performed equally well in split-sample and proxy-basin tests.

However, such equal model performance was not necessarily anticipated beforehand. Concerning the transfer of model parameters to adjacent catchments, it was expected that the distributed and more process-oriented model would outperform the lumped concept, based on its spatially more explicit representation of hydrological processes and its increased degrees of freedom resulting from the enlarged parameter space. Within the old debate about the value of distributed versus lumped modelling (e.g. Beven 1996; Refsgaard *et al.* 1996) this model comparison may thus be seen as another example for the supremacy of less parameterized lumped model concepts, if the objective is the best discharge fit at the catchments outlet, given a limited input data set that is explored to its full potential (e.g. Jakeman & Hornberger 1993; Lischeid & Uhlenbrook 2003; Carpenter & Georgakakos 2006; Breuer *et al.* 2009). This confirms, in accordance with earlier work (e.g. Grayson *et al.* 1992; Michaud & Sorooshian 1994; Refsgaard & Knudsen 1996; Beven 2001), that distributed models seldom demonstrate superiority over much simpler lumped or semi-distributed models, if tested only against runoff at the catchment outlet that constitutes lumped data, integrated over the whole catchment. In contrast, the main advantage of a distributed model is its capability of simulating additional internal state variables that can be subject to multi-criteria calibration, if additional data is available (e.g. groundwater data, synoptic runoff measurements). This is especially important in terms of evaluating model structure uncertainty (Fenicia *et al.* 2008) and refining hydrological process descriptions in order to prevent the 'the model is right for the wrong reasons' case (Klemes 1986), to which lumped model concepts may tend to.

Such spatially differentiated and more process-oriented simulations might be beneficial in integrated catchment modelling when the focus lies on the more accurate representation of spatial patterns. For the case of nutrient transport modelling, point sources and diffuse sources could be incorporated according to their spatial representation so that degradation and retention patterns can be simulated in a spatially more representative way (e.g. Lindgren *et al.* 2007). In situations where models with rather coarse grid resolutions need to consider small scale processes in an adequate manner with limited computation power, a sub-grid parameterization can be appropriate. This is frequently the case at larger scales such as regional climate models with land surface schemes (e.g. Kotlarski & Jacob 2005) or Soil-Vegetation-Atmosphere-Transfer (SVAT) models (e.g. Strasser & Etchevers 2005), but can also be transferred to meso-scale model applications like nutrient transport modelling (e.g. Exbrayat *et al.* 2010), where small lakes, wetlands or riparian zones can have a considerable impact on nutrient flows and distributions (Gren 1995; Carpenter *et al.* 1998; Hooper 2001).

Synoptic runoff measurements

The synoptic runoff measurements allowed an evaluation of the spatial representativity of the model simulation for distributed runoff predictions. Overall, the model was able to reproduce the spatial variations of runoff volumes, but not the variations in specific runoff. This multi-scale evaluation indicated that the model can capture scale induced runoff volumes, but was not capable of reproducing the spatial variations of runoff generation processes in a low flow situation, despite all calibration efforts. This might be partially explained by measurement uncertainties and shortcomings to correctly delineate sub-catchments in this lowland region. In addition, potential evaporation estimates were not differentiated for different land use classes which might further explain these shortcomings. However, using only one snapshot campaign during low flow conditions is certainly the most important limitation, but the poor agreement between simulated and predicted specific runoff gave an indication that a further evaluation of the spatial representation of runoff generating mechanisms in the model structure is needed. This corresponds to other authors that argue for spatially distributed data to fully evaluate the capabilities of a distributed model (Gupta *et al.* 1998; Grayson & Blöschl 2001). Despite such limitations, synoptic runoff measurements can provide useful insights for model evaluation or calibration purposes with a few additional observations (e.g. Perrin *et al.* 2007; Seibert & Beven 2009).

CONCLUSION

In this paper a fully distributed, conceptual hydrological model was evaluated that can serve as a basis for water quality applications in the Fyrisån catchment in Sweden. To enable a better representation of small scale landscape features and related runoff generation processes, such as wetlands, a sub-grid parameterization scheme was incorporated into the model. Flow routing along stream networks and lakes was achieved by a new designed simple flow routing routine. The objective of the presented model was to find a spatially meaningful hydrological model that can provide driving variables for coupled distributed solute transport routines (Lindgren *et al.* 2007; Exbrayat *et al.* 2010).

Intensive model assessment and comparison throughout the study revealed limitations of the model capabilities, especially with respect to the transferability of model parameters and the simulation of spatial runoff patterns. The model performed only equally well compared to the much simpler lumped HBV model with regard to simulating runoff at the catchment outlets. However, it is important to note that the identification of these model shortcomings was only possible due to the rigorous tests that were carried out during the model evaluation process and once again highlight the importance of thorough model evaluation procedures (Andreassian *et al.* 2009). Besides the well-known, albeit less widely applied, test procedures suggested by Klemes (1986), it was demonstrated that the comparison with hydrologic benchmark models, such as the lumped HBV model, can provide further insights on model performance. Furthermore, inexpensive synoptic stream flow measurements were found to be valuable for model evaluation. Synoptic data allowed an assessment of the spatial representativity of the model simulation, whereas model evaluation is usually only focused on the temporal aspects of runoff dynamics at the catchment outlet. Despite these apparent model shortcomings of the distributed model concept, it provides a valuable basis for distributed water quality assessments and will be the subject of further model development.

ACKNOWLEDGEMENTS

Discharge and precipitation data were provided by SMHI (Swedish Meteorological and Hydrological Institute). The authors thank the Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, especially Mats Wallin and Jakob Nisell, for help with the land-use information. Funding provided by the German Academic Exchange Service (DAAD) supported an extended visit of the first author to Sweden.

REFERENCES

Andreassian, V., Perrin, C., Berthet, L., Le Moine, N., Lerat, J., Loumagne, C., Oudin, L., Mathevet, T., Ramos, M. H. & Valery, A. 2009 HESS Opinions 'Crash tests for a standardized evaluation of hydrological models'. *Hydrol. Earth Syst. Sci.* **13** (10), 1757–1764.

Arheimer, B. & Brandt, M. 1998 Modelling nitrogen transport and retention in the catchments of southern Sweden. *Ambio* **27** (6), 471–480.

Arheimer, B. & Wittgren, H. B. 1994 Modeling the effects of wetlands on regional nitrogen transport. *Ambio* **23** (6), 378–386.

Arheimer, B. & Wittgren, H. B. 2002 Modelling nitrogen removal in potential wetlands at the catchment scale. *Ecol. Eng.* **19** (1), 63–80.

Becker, A. & Braun, P. 1999 Disaggregation, aggregation and spatial scaling in hydrological modelling. *J. Hydrol.* **217** (3–4), 239–252.

Beldring, S., Engeland, K., Roald, L. A., Saelthun, N. R. & Vokso, A. 2003 Estimation of parameters in a distributed precipitation-runoff model for Norway. *Hydrol. Earth Syst. Sci.* **7** (3), 304–316.

Bergström, S. 1975 The development of a snow routine for the HBV-2 model. *Nord. Hydrol.* **6**, 73–92.

Bergström, S. 1990 *Parametervärden för HBV-modellen i Sverige: erfarenheter från modellkalibreringar under perioden 1975–1989*. SMHI, Norrköping.

Bergström, S. 1995 The HBV model. In: *Computer Models of Watershed Hydrology* (V. P. Singh ed.). Water Resources Publications, Highlands Ranch, Colorado, pp. 443–476.

Beven, K. J. 1996 A discussion of distributed hydrological modelling. In: *Distributed Hydrological Modelling* (M. B. Abbott & J. C. Refsgaard, eds). Kluwer Academic, Dordrecht, pp. 255–278.

Beven, K. J. 2001 *Rainfall-Runoff Modelling: The Primer*. John Wiley & Sons, Ltd., Chichester.

Beven, K. 2006 A manifesto for the equifinality thesis. *J. Hydrol.* **320** (1–2), 18–36.

Blöschl, G. & Sivapalan, M. 1995 Scale issues in hydrological modelling: a review. *Hydrol. Proc.* **9**, 251–290.

Breuer, L., Huisman, J. A., Willems, P., Bormann, H., Bronstert, A., Croke, B. F. W., Frede, H. G., Graff, T., Hubrechts, L., Jakeman, A. J., Kite, G., Lanini, J., Leavesley, G., Lettenmaier, D. P., Lindstrom, G., Seibert, J., Sivapalan, M. & Viney, N. R. 2009 Assessing the impact of land use change on hydrology by ensemble modeling (LUCHEM). I: model intercomparison with current land use. *Adv. Water Res.* **32** (2), 129–146.

Carpenter, T. M. & Georgakakos, K. P. 2006 Intercomparison of lumped versus distributed hydrologic model ensemble simulations on operational forecast scales. *J. Hydrol.* **329** (1–2), 174–185.

Carpenter, S. R., Caraco, N. F., Correll, D. L., Howarth, R. W., Sharpley, A. N. & Smith, V. H. 1998 Nonpoint pollution of surface waters with phosphorus and nitrogen. *Ecol. Appl.* **8** (3), 559–568.

Darracq, A., Greffe, F., Hannerz, F., Destouni, G. & Cvetkovic, V. 2005 Nutrient transport scenarios in a changing Stockholm and Malaren valley region, Sweden. *Water Sci. Technol.* **51** (3–4), 31–38.

Doherty, J. 2005 *PEST: Model-independent Parameter Estimation, User Manual*. Watermark Numerical Computing, Brisbane.

Doherty, J. & Johnston, J. M. 2003 Methodologies for calibration and predictive analysis of a watershed model. *J. Am. Water Res. Assoc.* **39** (2), 251–265.

Donnelly-Makowecki, L. M. & Moore, R. D. 1999 Hierarchical testing of three rainfall-runoff models in small forested catchments. *J. Hydrol.* **219** (3–4), 136–152.

Eriksson, B. 1981 *Den 'potentiella' Evapotranspirationen i Sverige: The 'potential' Evapotranspiration in Sweden*. SMHI, Norrköping.

Exbrayat, J. F., Viney, N. R., Seibert, J., Wrede, S., Frede, H. G. & Breuer, L. 2010 Ensemble modelling of nitrogen fluxes: data fusion for a Swedish meso-scale catchment. *Hydrol. Earth Syst. Sci.* **14**, 2383–2397.

Fenicia, F., Savenije, H. H. G., Matgen, P. & Pfister, L. 2008 Understanding catchment behavior through stepwise model concept improvement. *Water Resour. Res.* **44** (1), W01402.

Grayson, R. & Blöschl, G. 2001 *Spatial Patterns in Catchment Hydrology Observations and Modelling*. Cambridge University Press, Cambridge, UK.

Grayson, R. B., Moore, I. D. & Mcmahon, T. A. 1992 Physically based hydrologic modeling, 2. Is the concept realistic. *Water Resour. Res.* **28** (10), 2659–2666.

Gren, I.-M. 1995 The value of investing in wetlands for nitrogen abatement. *Eur. Rev. Agric. Econ.* **22** (2), 157–172.

Gretener, B. 1994 *The River Fyris: A Study of Fluvial Transportation*. Institute of Earth Sciences, Uppsala University, Uppsala.

Gupta, H. V., Sorooshian, S. & Yapo, P. O. 1998 Toward improved calibration of hydrologic models: multiple and noncommensurable measures of information. *Water Resour. Res.* **34** (4), 751–763.

Gupta, H. V., Wagener, T. & Liu, Y. Q. 2008 Reconciling theory with observations: elements of a diagnostic approach to model evaluation. *Hydrol. Proc.* **22** (18), 3802–3813.

Hjulström, F. 1993 Studies of the morphological activity of rivers as illustrated by the River Fyris. *Bulletin of the Geological Institution of the University of Uppsala* **25**, 221–527.

Hooper, R. P. 2001 Applying the scientific method to small catchment studies: a review of the Panola Mountain experience. *Hydrol. Proc.* **15** (10), 2039–2050.

Hundecha, Y. & Bardossy, A. 2004 Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model. *J. Hydrol.* **292** (1–4), 281–295.

Hutchinson, M. F. 1989 A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. *J. Hydrol.* **106** (3–4), 211–232.

Jakeman, A. J. & Hornberger, G. M. 1993 How much complexity is warranted in a rainfall-runoff model. *Water Resour. Res.* **29** (8), 2637–2649.

Johst, M., Uhlenbrook, S., Tilch, N., Zillgens, B., Didszun, J. & Kirnbauer, R. 2008 An attempt of process-oriented rainfall-runoff modeling using multiple-response data in an alpine catchment, Loehnersbach, Austria. *Hydrol. Res.* **39** (1), 1–16.

Karssenberg, D., Burrough, P. A., Sluiter, R. & de Jong, K. 2001 The PC raster software and course materials for teaching numerical modelling in the environmental siences. *Trans. GIS* **5** (2), 99–110.

Klemes, V. 1986 Operational testing of hydrological simulation models. *Hydrol. Sci. J.–J. Sci. Hydrol.* **31** (1), 13–24.

Kotlarski, S. & Jacob, D. 2005 Development of a subgrid scale parameterization of mountain glaciers for use in regional climate modelling. WGNE Blue Book, WMO, Geneva, pp. 4–13.

Krasovskaia, I. & Gottschalk, L. 1992 Stability of river flow regimes. *Nord. Hydrol.* **23** (3), 137–154.

Lantmäteriet. 2003 *Terrängkartan/Gröna Kartan 1:50.000*. Lantmäteriet, Gävle.

Lindgren, G. A., Wrede, S., Seibert, J. & Wallin, M. 2007 Nitrogen source apportionment modeling and the effect of land-use class related runoff contributions. *Nord. Hydrol.* **38** (4–5), 317–331.

Lindström, G., Johansson, B., Persson, M., Gardelin, M. & Bergström, S. 1997 Development and test of the distributed HBV-96 hydrological model. *J. Hydrol.* **201** (1–4), 272–288.

Lischeid, G. & Uhlenbrook, S. 2003 Checking a process-based catchment model by artificial neural networks. *Hydrol. Proc.* **17** (2), 265–277.

Michaud, J. & Sorooshian, S. 1994 Comparison of simple versus complex distributed runoff models on a midsized semiarid watershed. *Water Resour. Res.* **30** (3), 593–605.

Motovilov, Y. G., Gottschalk, L., Engeland, K. & Rodhe, A. 1999 Validation of a distributed hydrological model against spatial observations. *Agric. Forest Meteorol.* **98** (9), 257–277.

Mroczkowski, M., Raper, G. P. & Kuczera, G. 1997 The quest for more powerful validation of conceptual catchment models. *Water Resour. Res.* **33**, 2325–2335.

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

O'Callaghan, J. F. & Mark, D. M. 1984 The extraction of drainage networks from digital elevation data. *Comput. Vis. Graph. Image Proc.* **28**, 328–344.

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** (1), 62–78.

Perrin, C., Oudin, L., Andreassian, V., Rojas-Serna, C., Michel, C. & Mathevet, T. 2007 Impact of limited streamflow data on the efficiency and the parameters of rainfall-runoff models. *Hydrol. Sci. J.–J. Sci. Hydrol.* **52** (1), 131–151.

Rabus, B., Eineder, M., Roth, A. & Bamler, R. 2003 The shuttle radar topography mission – a new class of digital elevation models acquired by spaceborne radar. *ISPRS J. Photogram. Remote Sens.* **57** (4), 241–262.

Refsgaard, J. C. & Henriksen, H. J. 2004 Modelling guidelines – terminology and guiding principles. *Adv. Water Res.* **27** (1), 71–82.

Refsgaard, J. C. & Knudsen, J. 1996 Operational validation and intercomparison of different types of hydrological models. *Water Resour. Res.* **32** (7), 2189–2202.

Refsgaard, J. C., Storm, B. & Abbott, M. B. 1996 Comment on 'a discussion of distributed hydrological modelling'. In: *Distributed Hydrological Modelling* (M. B. Abbott & J. C. Refsgaard, eds). Kluwer Academic, Dordrecht, pp. 279–288.

Saelthun, N. R. 1996 The 'Nordic' HBV model. Description and documentation of the model version developed for the project climate change and energy production. *Norwegian Water Resources and Energy Administration Publication* **7**, Oslo, pp. 26.

Seibert, P. 1994 Hydrological characteristics of the NOPEX research area. Masters thesis, Department of Earth Sciences (Hydrology), Uppsala University, Uppsala, 51 pp.

Seibert, J. 1997 Estimation of parameter uncertainty in the HBV model. *Nord. Hydrol.* **28** (4–5), 247–262.

Seibert, J. 1999 Regionalisation of parameters for a conceptual rainfall-runoff model. *Agric. Forest Meteorol.* **98** (9), 279–293.

Seibert, J. & Beven, K. J. 2009 Gauging the ungauged basin: how many discharge measurements are needed? *Hydrol. Earth Syst. Sci.* **13** (6), 883–892.

Sieber, A. & Uhlenbrook, S. 2005 Sensitivity analyses of a distributed catchment model to verify the model structure. *J. Hydrol.* **310** (1–4), 216–235.

Seibert, J., Uhlenbrook, S., Leibundgut, C. & Halldin, S. 2000 Multiscale calibration and validation of a conceptual rainfall-runoff model. *Phys. Chem. Earth B – Hydrol. Oceans Atmos.* **25** (1), 59–64.

Singh, V. P. 1995 *Computer Models of Watershed Hydrology*. Water Resources Publications, Highlands Ranch, Colorado.

Sooroshian, S. & Gupta, V. K. 1995 Model calibration. In: *Computer Models of Watershed Hydrology* (V. P. Singh, ed.). Water Resources Publications, Highlands Ranch, Colorado, pp. 23–68.

Strasser, U. & Etchevers, P. 2005 Simulation of daily discharges for the upper Durance catchment (French Alps) using subgrid parameterization for topography and a forest canopy climate model. *Hydrol. Proc.* **19** (12), 2361–2373.

Uhlenbrook, S. & Sieber, A. 2005 On the value of experimental data to reduce the prediction uncertainty of a process-oriented catchment model. *Environ. Model. Softw.* **20** (1), 19–32.

Uhlenbrook, S., Roser, S. & Tilch, N. 2004 Hydrological process representation at the meso-scale: the potential of a distributed, conceptual catchment model. *J. Hydrol.* **291** (3–4), 278–296.

Uhlenbrook, S., Seibert, J., Leibundgut, C. & Rodhe, A. 1999 Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model paramters and structure. *Hydrol. Sci. J.–J. Sci. Hydrol.* **44** (5), 779–797.

Wagener, T., Boyle, D. P., Lees, M. J., Wheater, H. S., Gupta, H. V. & Sorooshian, S. 2001 A framework for development and application of hydrological models. *Hydrol. Earth Sys. Sci.* **5** (1), 13–26.

Wissmeier, L. & Uhlenbrook, S. 2007 Distributed, high-resolution modelling of 18O signals in a meso-scale catchment. *J. Hydrol.* **332** (3–4), 497–510.

Xu, C. Y. 1999 Operational testing of a water balance model for predicting climate change impacts. *Agric. Forest Meteorol.* **98** (9), 295–304.

First received 28 June 2010; accepted in revised form 8 September 2011. Available online 21 July 2012