Journal of Hydrology (2008) **356**, 106–118

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jhydrol

Comparison of conceptual model performance using different representations of spatial variability

Tapash Das [a,*,1], **András Bárdossy** [a,2], **Erwin Zehe** [b,3], **Yi He** [a,4]

[a] *Institute of Hydraulic Engineering, Universitaet Stuttgart, 70569 Stuttgart, Germany*
[b] *Institute for Geoecology, University of Potsdam, Germany*

Received 17 October 2006; received in revised form 1 March 2008; accepted 3 April 2008

KEYWORDS
Distributed modelling;
Spatial variability;
Rainfall–runoff model;
Predictive uncertainty

Summary The objective in this study is to explore a solution to the question whether model input data having higher spatial resolution and higher model resolution, as most people assume, lead to better model performance within a given modelling objective. An attempt was made to modify the conceptual rainfall–runoff model HBV to incorporate a spatially distributed structure. Additionally, three more model structures based on the HBV model concept were designed: lumped, semi-lumped and semi-distributed. An automatic calibration procedure based on simulated annealing optimization algorithm was followed for maximizing an objective function composed of Nash–Sutcliffe coefficients of several temporal aggregation steps. The predictive performance from each model was then assessed and compared with other model structures with respect to stream flow prediction at the catchment outlet. The spatial variation of the meteorological input was produced using external drift kriging method from available limited point measurements. The models were applied to a mesoscale catchment located in central Europe. The simulated hydrographs obtained using different model structures were analyzed through comparison of their Nash–Sutcliffe coefficients and other goodness-of-fit indices. For the present study, semi-distributed and semi-lumped model structures outperformed the distributed and fully-lumped model structures. A possible explanation why the distributed model did not perform better than the simpler model structures is the use of limited available spatial information. The models use interpolated precipitation and temperature as input, which probably cannot reflect the true spatial variability. Another possible explanation is

* Corresponding author. Tel.: +1 858 822 3572.
E-mail addresses: tadas@ucsd.edu (T. Das), Andras.Bardossy@iws.uni-stuttgart.de (A. Bárdossy), erwin.zehe@uni-potsdam.de (E. Zehe), yi.he@kcl.ac.uk (Y. He).
[1] Present address: Scripps Institution of Oceanography, University of California, San Diego, United States.
[2] Tel.: +49 711 6856 4663.
[3] Tel.: +49 331 977 2675.
[4] Present address: Department of Geography, King's College London, London WC2R 2LS, UK, Tel.: +44 20 7848 2916.

0022-1694/$ - see front matter
doi:10.1016/j.jhydrol.2008.04.008

that only discharge at the catchment outlet was predicted; which is the purpose for which lumped and semi-distributed models were actually designed.

Introduction

Spatially-distributed hydrological models have been increasingly applied to account for spatial variability of the main forcing variables (e.g. precipitation), physiographic characteristics of a catchment (e.g. topography, soil, land use) and detailed process calculation within a catchment (Liang et al., 1994; Beldring et al., 2003). Distributed models are utilized to undertake impact assessment studies (e.g. land use change and climate change studies) (Christensen and Lettenmaier, 2007), investigation of the influence of the spatial variability of catchment physiographic-climatic characteristics (Liang et al., 2004), estimation of internal fluxes and state variables at high spatial resolution (Beldring et al., 2003), and prediction in interior locations of a catchment (Brath et al., 2004; Reed et al., 2004). The effect of spatial variability of precipitation on the response of catchments has been reported in a large amount of studies. Krajewski et al. (1991) found higher sensitivity of catchment response in terms of peak magnitude, time-to-peak, and total runoff volume with respect to the temporal resolution than to the spatial resolution of the synthetically generated rainfall data. They also found considerable underestimation of the flood peaks by the lumped model. Obled et al. (1994) concluded that distributed rainfall was not ''an essential prerequisite for flow forecasting in temperate medium-sized rural catchments''. Haddeland et al. (2002) found that the Variable Infiltration Capacity (VIC) model (Liang et al., 1994) preserved the general form of the simulated hydrograph at the catchment outlet at a lower spatial model resolution using aggregated meteorological, soil and vegetation data. However, the simulated total runoff was lower at the lowest model resolution than at the highest-model resolution they applied. Liang et al. (2004) indicated that a critical spatial resolution might exist for the macro scale hydrological model (VIC). They indicated that better model performances in terms of runoff, evapotranspiration, and total zone soil moisture would not be necessarily expected for spatial resolutions finer than the critical resolution, provided that model parameters are being calibrated at each spatial resolution. Smith et al. (2004a) stated: ''A major reason for the use of distributed approaches is the hypothesis that by accounting for the spatial variability of rainfall and physical features within the basin, better simulations can be achieved at the basin outlet''. This is considered because of the fact that nature is highly heterogeneous and non-linear. A distributed model is not only capable of accounting for spatial variability of hydrological processes; it also enables computation of internal fluxes and state variables. A number of studies have compared results obtained with models of varying complexities. Smith et al. (2004a,b) analyzed a large number of studies and provided a comprehensive trend of the results. Refsgaard and Knudsen (1996) compared three different models having three different model structures and degrees of spatial distribution using a systematic calibration and validation procedure. They found that performance of all models under study was equivalent if calibration was done with at least 1 year of data while the distributed models performed slightly better without any calibration. Farmer et al. (2003) applied models of increasing complexity (ranging from the simplest to quasi-distributed framework) to identify the trade-off between model complexity and simulation accuracy in a number of temperate and semiarid catchments across Australia. They observed higher sensitivity due to small-scale perturbation in drier catchments than humid catchments. El-Nasr et al. (2005) assessed the performance of two different models, the fully distributed MIKE SHE and the semi-distributed model SWAT. Given the availability of hydrologic data, both models were equally able to simulate the hydrology in the catchment, despite that the MIKE SHE model predicts slightly better the overall variation of the stream flow. Atkinson et al. (2002) described a systematic approach to determine the minimum level of model complexity to predict stream flow in catchments located in New Zealand. They noticed that dry catchments are more sensitive to threshold storage parameters than the wetter catchments. Booij (2005) compared three spatial resolutions (considering 1, 15 and 118 subcatchments) using the conceptual HBV model to the Meuse river in Europe. The average and extreme discharge behavior at the catchment outlet was reproduced well by the all three versions of the HBV model, in the calibration and validation periods; however, increasing the resolution slightly improved the model results. Atkinson et al. (2003) tested eight different model types, originating from a simple conceptual model, with complexity ranging from lumped to fully distributed in a catchment located in New Zealand. They found the required model complexity to be a function of the season; however, the fully distributed representation was most appropriate under all seasonal climatic conditions to obtain accurate model predictions.

Additionally, Chaubey et al. (1999) found large uncertainty in estimated model parameters when detailed variations in the input rainfall were not taken into account. Zehe et al. (2005) showed that spatially highly resolved rainfall input leads to a reduced uncertainty of runoff predictions for a physically-based distributed model. They also indicated that the sensitivity of a rainfall–runoff model to input errors may depend partly on the model structure itself.

Moreover, a number of studies have been reported in the literature to examine the impact of catchment properties on model simulations, including the effect of grid size on catchment simulations. For example, Kuo et al. (1999) applied a variable-source-area hydrological model to grid sizes ranging from 10 to 600 m. They observed an increasing misrepresentation of the curvature of the landscape with increasing grid size; however, soil properties and land use distribution were not affected. Romanowicz et al. (2005) investigated the sensitivity of the distributed hydrological

SWAT model to the quality of the soil and land use data for modelling rainfall–runoff processes. They indicated that the SWAT model is highly sensitive to the soil and land use data quality. Bormann (2006) showed that an aggregation of digital elevation model, soil map and land use classification for the calculation of regional water balances, using TOPLATS model, did not lead to significant errors up to a grid size of 300 m. However, there were significant errors in the simulated water balance when applying a grid size of 1 km and larger.

Nevertheless, a higher spatial and temporal resolution of data and model application does not always lead to a better model performance (Reed et al., 2004). Results from the Distributed Model Intercomparison Project (DMIP) study (Reed et al., 2004; Smith et al., 2004a,b) indicate that distributed modelling approaches may not always provide improved outlet simulations compared to lumped conceptual model. They suggested that there may be a trade-off between the complexity of the model descriptions necessary to represent the catchment processes, the accuracy and the representativeness of the input data available and the accuracy required to achieve reliable simulation results. Butts et al. (2004) evaluated impacts of different model structures on hydrological modelling uncertainty for stream flow simulation. They showed that model performance is strongly dependent on model structures. Nevertheless, they observed distributed routing and to some extent distributed rainfall were the dominant processes controlling simulation accuracy in their study catchment. Perrin et al. (2001) carried out a comparative performance assessment of the structures of 19 daily lumped models on large number of catchments located in different regions. They observed the complex models outperform the simpler models in the calibration period, however, not in the verification period. They argued the main reason why the complex models lack stability is the complex structure is not appropriate to extracting information available in the hydrological time-series. They also indicate that inadequate complexity typically results in model over-parameterization and thus parameter uncertainty. Viney et al. (2005) evaluated 10 different models having different levels of complexity ranging from lumped to fully distributed models for a mesoscale catchment located in Germany. They found the simpler models tend to outperform the more complex models in both calibration and validation periods. They also showed all models tend to produce higher performance during the less-extreme validation periods, but this improvement was greatest for some of the more complex models applied in the study.

The main conclusions from these studies are that the model performance can vary depending on several factors, including the scale of the catchment, physiographic characteristics of the catchment, availability of the data required to set up the selected model, the dominating rainfall type (convective or advective), seasonality of precipitation, season of the year and dominating runoff producing mechanisms. For numerical models, it is expected that performance should in principle improve with increasing model resolutions because the discretization errors tend to zero when grid size goes to zero. But when applying models to heterogeneous landscapes this effect is often disturbed, because spatially resolved data are not at hand; for conceptual models there should be a lower limit of splitting the landscapes, simply because they represent spatial average behavior. Hence, we expect an optimum finite model resolution.

This paper attempts to answer the question whether input data having high spatial resolution and high model resolution lead to improved model results within a given modelling objective; in particular with respect to stream flow simulation in a mesoscale catchment. The conceptual rainfall–runoff model HBV was selected and modified to incorporate different degrees of spatial resolution, ranging from a lumped to a distributed structure. The predictive performance is then assessed and compared among different model structures. The models are applied to the Neckar catchment located in central Europe. The simulated hydrographs obtained using different model structures are analyzed through the comparison of different goodness-of-fit indices.

The remainder of this paper is organized as follows. The following section presents the description on the study catchment and data used in the study. A description of the model and methodology employed in the study is provided in model and methods section. The parameter estimation procedure is also outlined in the same section. The results are presented and discussed in the results and discussion section. Final section provide outlooks and conclusions.

Study catchment and data sets

This study was performed on the upper Neckar catchment, located in southwest Germany, situated downstream of Rottweil and upstream of Plochingen (Fig. 1). It covers an area of approximately 4000 km^2 which is approximately 28% of the entire Neckar catchment. The study catchment area was divided into 13 subcatchments depending on the available discharge gauges. Table 1 shows the sizes of different subcatchments; the subcatchment areas ranging between approximately 120 km^2 and 613 km^2. The elevation of the study area ranges between approximately 240 m above sea level (m.a.s.l.) to around 1010 m.a.s.l., and having a mean elevation of about 550 m.a.s.l. The long-term (for the period from 1961 to 1990) average daily temperature is 8.1 °C, ranging between −19.1 °C and 27.1 °C. The long-term average annual precipitation observed in the period from 1961 to 1990 is approximately 908 mm, ranging between 680 mm and 1230 mm with a standard deviation of 135 mm. There are snow fall events during the winter months, particularly in the upper parts of the catchment – northwest and southeast parts of the catchment (Fig. 1). The average daily discharge at the catchment outlet Plochingen (Neckar) is about 50.3 $m^3 s^{-1}$, ranging between 5.3 $m^3 s^{-1}$ and 1031.0 $m^3 s^{-1}$, while the standard deviation is 51.2 $m^3 s^{-1}$ for the period 1961–1990.

The required input data for the HBV model, applied in this study, are precipitation, air temperature and potential evapotranspiration. The daily precipitation total and daily maximum and minimum temperature data from 151 precipitation stations and 74 temperature stations distributed in and around the study catchment were collected from the

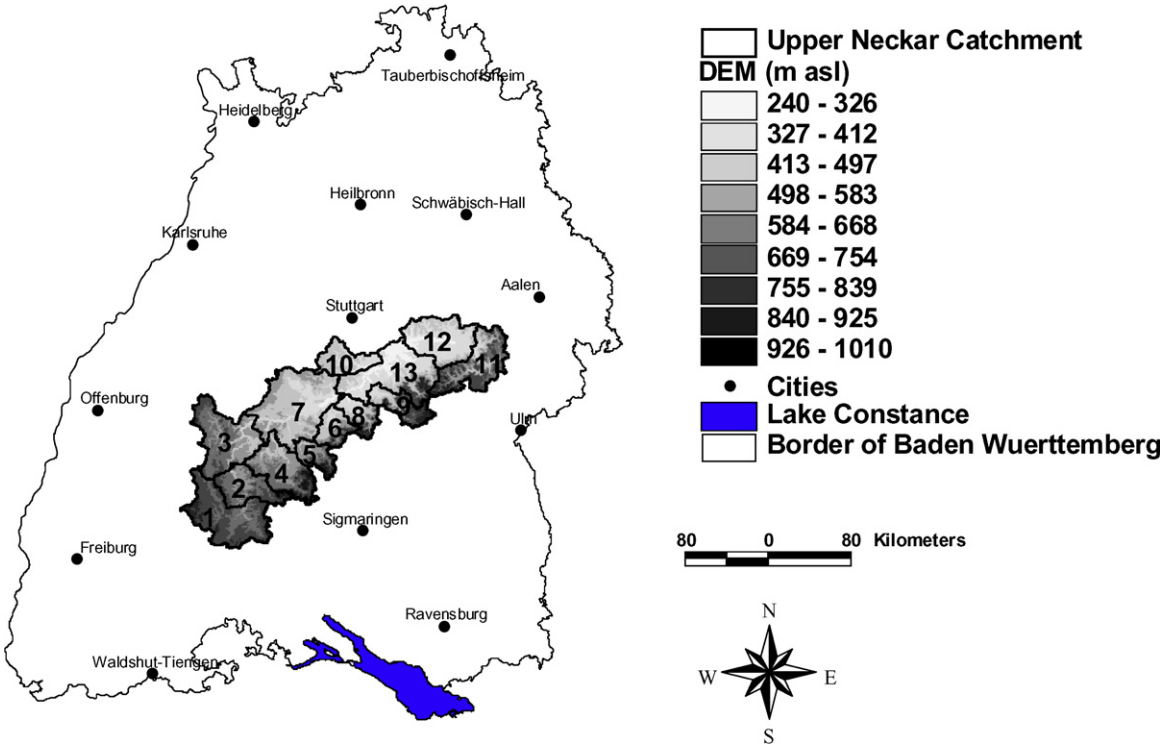

Figure 1 Upper Neckar catchment and subcatchment study areas, State of Baden-Wuerttemberg, southwest Germany.

Table 1 Summary of the sizes of different subcatchments

	Gauging station (River)	Subcatchment size [km^2]	Drainage area [km^2]
1	Rottweil (Neckar)	454.65	454.65
2	Oberndorf (Neckar)	240.13	694.78
3	Horb (Neckar)	420.18	1114.96
4	Bad Imnau (Eyach)	322.94	322.94
5	Rangendingen (Starzel)	119.89	119.89
6	Tuebingen (Steinlach)	140.21	140.21
7	Kirchentellinsfurt (Neckar)	613.33	2311.33
8	Wannweil (Echaz)	135.26	135.26
9	Riederich (Erms)	169.84	169.84
10	Oberensingen (Aich)	178.18	178.18
11	Suessen (Fils)	345.74	345.73
12	Plochingen (Fils)	349.09	694.83
13	Plochingen (Neckar)	472.05	3961.49

German Weather Service (DWD). The daily precipitation and the daily average, maximum and minimum temperatures were interpolated on a 1×1 km^2 grid using the external drift kriging method (Ahmed and de Marsily, 1987) from the available point measurements. Topographic elevation was used as external variable (drift) for the external drift kriging method (Hundecha and Bárdossy, 2004). Cross validation results showed the proposed kriging model was an appropriate selection.

In this study, the potential evapotranspiration was computed using the Hargreaves and Samani method (Hargreaves and Samani, 1985) on the same grid used for the interpolation of meteorological variables.

The period from 1961 to 1970 was used for calibration of the models. The subsequent period from 1971 to 1990 was used to validate the models. The meteorological conditions do not differ strongly between the calibration and validation periods. The mean daily temperature in the validation period is slightly higher than that of the calibration period in the majority of the study area. The difference in the mean daily temperature between the validation and calibration periods ranges between 0 and 0.4 °C. In contrast, the mean daily precipitation over the validation period is 0—11% lower than the corresponding value in the calibration period. The variability of both the daily mean temperature and daily amount of precipitation are more or less similar in the calibration and validation period.

Different GIS based digital data for the study catchment, including Digital Elevation Model (DEM) (30 m $\times$ 30 m horizontal spatial resolution), a land use/land cover map (LANDSAT satellite image for the year 1993) (50 m $\times$ 50 m horizontal spatial resolution), a soil map (1:200,000 scale) and the river network, were obtained from the State Agency for Environmental protection (LUBW, Baden-Wuerttemberg). The DEM, soil map, and land use were resampled on 1×1 km^2 spatial resolution. This is the model resolution adopted for the distributed model structure. The study area has experienced some land use changes in the last several decades (Samaniego, 2003; Samaniego and Bárdossy, 2006). A study performed in the upper Neckar catchment by Samaniego (2003) gives statistics on the land use changes in the catchment for the period 1981—1997. As reported in Samaniego (2003), agriculture showed a steady average growth rate of approximately 0.48% per year. Built-up, commerce, recreation, transportation, water bodies and wetlands showed growth at average rates of 0.90%, 1.67%,

3.73%, 0.40% and 1.01% per year, respectively. Forests had an annual growth rate of approximately 0.67% for the period from 1985 to 1993, and followed by a reduced annual growth rate of 0.15% from 1993 to 1997. In this study, we used land use information from the LANDSAT satellite image for the year 1993 and assumed it is more or less stationary throughout the calibration and validation periods. That is not only because the impact of land use change was not the primary goal of the study, but also the HBV model used in the study is not sensitive in response to land use change. Hundecha and Bárdossy (2004) only obtained a rather weak influence for floods by doubling of the forested area. Additionally, the ET routine implemented in the HBV model did not take plant morphology into consideration and, therefore, is not sensitive to changes in cropping patterns. Nevertheless, it is worth noting that the land use change can be a potential source of uncertainty.

Model and methods

The HBV model and different model structures

The HBV model is a semi-distributed conceptual model and was originally developed at the Swedish Meteorological and Hydrological Institute (SMHI) (Bergström and Forsman (1973)). The area to be modelled is divided into a number of subcatchments and each subcatchment is further divided into a number of zones according to elevation, land use or soil type or combinations of them. Snow accumulation and melt, actual soil moisture and runoff generation processes are calculated for each zone using conceptual routines. The snow routine uses the degree-day approach. Actual soil moisture is calculated by considering precipitation and evapotranspiration. Runoff generation is estimated by a non-linear function of actual soil moisture and precipitation. The dynamics of the different flow components at the subcatchment scale are conceptually represented by two linear reservoirs. The upper reservoir simulates the near surface and interflow in the sub-surface layer, while the lower reservoir represents the base flow. Both reservoirs are connected in series by a constant percolation rate. Lastly, one transformation function is used for smoothing the generated flow. The transformation consists of a triangular weighting function with one free parameter. Flood routing of the discharge between the river network nodes are carried out using the Muskingum method (Cunge, 1969). More detailed descriptions on the HBV model can be found elsewhere (Bergström and Forsman, 1973; Lindström et al., 1997; Hundecha and Bárdossy, 2004). It is worth noting that the distribution of each subcatchment into different zones is not spatially fixed in the original semi-distributed HBV model structure. This is because HBV modelling units (zones) are generally accomplished based on common elevation class or land use class or soil class or combination of the mentioned physical characteristics within each subcatchment without giving exact consideration where the considered elevation class or land use class or soil class are actually located in that subcatchment (Ehret, 2002).

Based on the HBV model concept, four different model structures, namely distributed, semi-distributed, semi-lumped and fully-lumped, were developed for this study purpose. A short description on different model structures is provided below.

Distributed structure

In the distributed model, the modifications were undertaken in the distributed model structure to account for detailed catchment characteristics (e.g. soil and land use), processes calculation and highly resolved meteorological variables (e.g. precipitation). In the modified structure, the catchment is divided into a number of subcatchments, as in the case of existing semi-distributed model structure. Each subcatchment is further divided into a number of grid cells in the distributed structure. Fig. 2 illustrates the primary difference effected in the distributed structure where the grid cells are used as computing units for snow accumulation and melt, and soil moisture accounting processes. The runoff response process, which is represented conceptually by two reservoirs, for direct discharge and base flow calculation for each subcatchment is kept unchanged, as in the original model structure, to restrict the number of model parameters to be optimized. Thus, in principle the model parameters describing the snow accumulation and melt, and soil moisture accounting processes can be adjusted differently for each individual grid in the modified model structure.

In the present study, the subcatchments were divided into a number of regular grids with an area of 1 km^2. For the modelling of the processes, snow accumulation and melt, and soil moisture accounting, in each individual grid, the interpolated daily precipitation amount and mean daily temperature on a regular grid of 1×1 km^2 was assigned to each computational grid. The potential evapotranspiration calculated on 1×1 km^2 grid was also assigned to each model grid.

Semi-distributed structure

In the semi-distributed model structure, sub-division of the subcatchments into a number of different homogenous zones can be accomplished based on various catchment characteristics (for example, topographic elevation, soil type, and land use) (Hundecha, 2005). The meteorological variables, precipitation and temperature, and potential evapotranspiration are assigned to each zone. In principle, the model parameters describing the snow accumulation and melt and the soil moisture accounting routine are different for each zone.

In this study topographic elevation was considered in defining the zones; elevation affects the distribution of the precipitation and temperatures, the rate of evaporation, and the snow accumulation and melt. Elevation zones were defined using a contour interval of 75 m and a maximum of 10 elevation zones were defined in each subcatchment. The precipitation, temperature, and potential evapotranspiration on regular grids of 1×1 km^2 were averaged according to topographic zone and assigned as input for modelling the hydrological processes. The parameters describing the runoff response processes were estimated for each subcatchment as in the case of the distributed model structure.

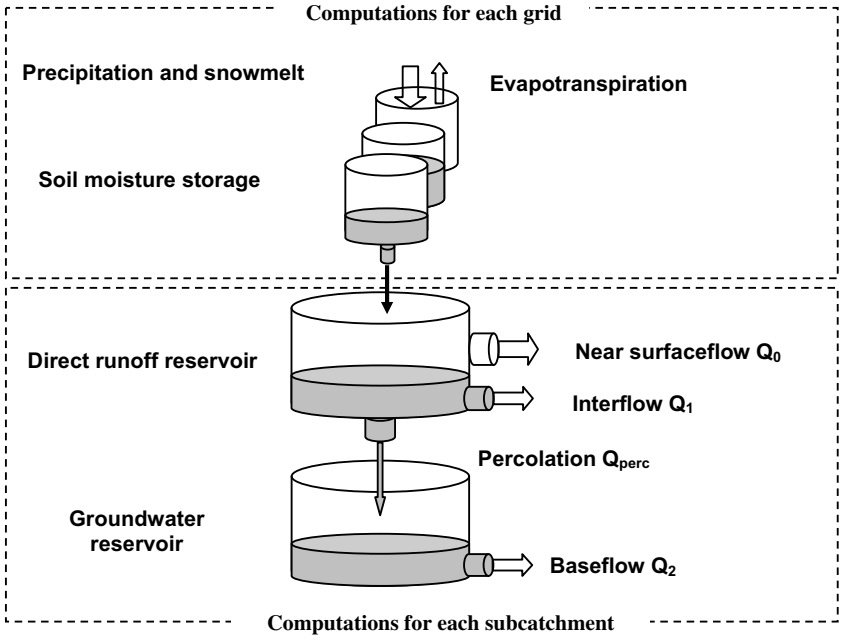

Figure 2 Modified distributed structure of the HBV model and representation of primary hydrological processes.

Semi-lumped structure

In this study, the only difference between the semi-distributed and semi-lumped model structures lies in the fact that the model parameters describing the snow accumulation and melt and soil moisture accounting processes were assigned the same values to each zone within a subcatchment. However, these parameter values are different for different subcatchments. The meteorological variables, precipitation and temperature, and potential evapotranspiration were assigned to each zone similar to the semi-distributed model structure.

Fully-lumped structure

In the fully-lumped model, the meteorological variables, precipitation and temperature and potential evapotranspiration were assigned to each subcatchment. All the processes, snow accumulation and melt, soil moisture accounting and runoff response were calculated for each individual subcatchment. Thus, the model parameters describing the different processes were estimated separately for each subcatchment.

For each subcatchment, the meteorological variables were estimated as the mean of the interpolated values on the regular grids of 1×1 km^2 located within a given subcatchment. Similarly, the potential evapotranspiration was averaged over each subcatchment from the calculated potential evapotranspiration on 1×1 km^2 grids located within a given subcatchment.

Table 2 illustrates the features of the different model structures of the HBV model.

Modelling time step

Daily meteorological data were used in this study; however precipitation data with a high temporal resolution are required, in general, in order to properly model the magnitude and the temporal distribution of high flows. A simple disaggregation procedure was used. The daily amount of precipitation was uniformly disaggregated through the middle 12 h of the day, as implemented in Bárdossy and Das (2008). The model was then run at a time step of 6 h. The disaggregation was important because the flow times between the different gauges were less than a day. This disaggregation method improved the performance of the models (not shown). Still, a uniform distribution of rainfall over the day can be a source of uncertainty; more sophisticated disaggregation schemes could have been used in a stochastic framework, but were not applied in this study.

Model parameter estimation

Conceptual rainfall–runoff models are designed to approximate the dominant physical processes of the hydrological cycle with the help of a series of connected mathematical equations (Beven, 2001). The parameters are generally not measurable and thus are determined through the calibration procedure (Yapo et al., 1996; Götzinger and Bárdossy, 2007). In this study the different model structures were calibrated based on the combinatorial optimization algorithm, simulated annealing (Aarts and Korst, 1989), to maximize the objective function composed of Nash–Sutcliffe coefficients (Nash and Sutcliffe, 1970) of several temporal aggregation steps based on the procedure described by Hartmann and Bárdossy (2005). This was achieved by a simultaneous calibration of the model for the daily and longer time steps.

The calibration of the different model structures were carried out using daily and mean annual discharges. Thus, the overall objective function to be optimized was the arithmetic sum of the weighted Nash–Sutcliffe coefficients corresponding to the discharges at the daily and annual time steps.

Table 2 General features summary of the various HBV model structures

Model structure	Meteorological variables scale		Processes calculation scale			Parameterization scale		Discharge comparison
	Precip. & Temp.	Potential ET	Effective Precip.	Actual ET	Runoff concentration	Soil parameters	Snow parameters	
Fully-lumped	SCat	SCat	SCat	SCat	SCat	SCat	SCat	Scat
Semi-lumped	Zone	Zone	Zone	Zone	SCat	SCat	SCat	Scat
Semi-distributed	Zone	Zone	Zone	Zone	SCat	Zone	Zone	SCat
Distributed	Grid	Grid	Grid	Grid	Scat	HRU_a (SC & LC)	HRU_b (LC)	Scat

Total number of SCats: 13.
Total number of grids in the catchment: 3962.
Total number of zones in the catchment: 82.
Total number of HRU (based on SC and LC): 28.
ET — evapotranspiration; Precip. — precipitation; Temp. — temperature; SCat — subcatchment; HRU — hydrological response unit; SC — soil class; LC — landuse class.

In the distributed model structure, the model parameters (responsible for snow accumulation and melt, and soil moisture accounting processes) were calibrated based on the concept of hydrological response units (HRU) mainly to reduce the problem of over-parameterization. A HRU is defined on the basis of soil type and land use information. The soil type and land use type were re-classified into a smaller number of classes; 7 classes for soil types and 4 classes for land use types. Specific catchment characteristics were then categorized into 28 HRUs (HRU_a) and the parameters describing the soil moisture accounting process were optimized based on the HRU_a. Therefore, for each parameter of each HRUa, a range of maximum and minimum was calculated based on the soil type. Specifically, the maximum value of field capacity multiplied by soil thickness (FC) and permanent wilting point multiplied by soil thickness (PWP), among the 7 soil types, were used as maximum value of FC and PWP during the optimization (FC and PWP are two conceptual model parameters). Finally, any optimized parameters corresponding to the 28 HRUs were assigned to the model grids in such a way that grids with similar HRUs received the same value irrespective of the fact where the grid was located in space. In the same way, the parameters describing the snow accumulation and melt process were optimized for 4 HRUs (grouped based on the land use class) (HRU_b).

In the semi-distributed model structure, the model parameters describing the snow accumulation and melt, and soil moisture accounting processes were optimized for each zone within each subcatchment.

In the semi-lumped model structure, the parameters describing the snow accumulation and melt and soil moisture accounting processes were optimized for only one zone for each subcatchment. Then the optimized value of that parameter was simultaneously assigned to all zones within that subcatchment. As a result, the value of any parameter is different for different subcatchments, but that value is same for all the zones within a subcatchment.

In the fully-lumped model structure, the parameters were optimized individually for each subcatchment.

The same calibration period (1961—1970), calibration objective, and number of iterations (100 iterations and each iteration step had a number of repetition steps) employed in the automatic calibration algorithm were used for each model structure. Appropriate constraints were assigned to the parameter ranges based on previous modelling experiences (Hundecha, 2005). Other physically significant constraints were also applied to some of the model parameters, as implemented in Hartmann and Bárdossy (2005). In order to avoid effects from the selection of the initial conditions, the first year was used as a warm-up for the calibration, which has been suggested by Hundecha (2005).

Goodness-of-fit criteria in simulations

A number of statistical criteria were used to compare the simulation results obtained using the four different model structures, including:

(1) Nash—Sutcliffe coefficient (Nash and Sutcliffe, 1970),
(2) Relative bias,
(3) Peak error, and
(4) Root mean squared error (RMSE)

The Nash—Sutcliffe coefficient (R^2_m) (Nash and Sutcliffe, 1970) is defined as

$$R^2_m = 1 - \frac{\sum_{i=i}^{N}(Q_s(t_i) - Q_o(t_i))^2}{\sum_{i=1}^{N}(Q_o(t_i) - \overline{Q}_o)^2} \quad (1)$$

where $Q_o(t_i)$ and $Q_s(t_i)$ are observed and simulated daily discharge at time step t_i, respectively, and $\overline{Q}_o$ is mean observed daily discharge and N is the total number of time steps.

The range of R^2_m lies between $-\infty$ and 1.0 (perfect fit). Higher values of R^2_m indicate better model performance.

The relative bias was computed to study the total volume error for various model structures. The relative bias (rel. bias) is defined as

$$\text{Rel.bias} = \frac{\sum_{i=i}^{N}(Q_s(t_i) - Q_o(t_i))}{\sum_{i=i}^{N}Q_o(t_i)} \quad (2)$$

The peak error was calculated to examine the peak flow estimation capacity by the models. The peak error is defined based on the relative difference between the mean

annual simulated peak discharge $\overline{Q_{s(max)}}$ and the mean annual observed peak discharge $\overline{Q_{o(max)}}$:

$$\text{Peak error} = \frac{\overline{Q}_{s(max)} - \overline{Q}_{o(max)}}{\overline{Q}_{o(max)}} \quad (3)$$

Additionally, the root mean squared error (defined in Eq. (4)) between the observed and simulated discharges was computed,

$$\text{RMSE} = \left(\frac{1}{N}\left(\sum_{i=1}^{N}(Q_s(t_i) - Q_o(t_i))^2\right)\right)^{0.5} \quad (4)$$

Furthermore, the model mean performance (R^2_{mm}) was estimated as the mean of the Nash–Sutcliffe coefficient values obtained during the calibration and validation periods for the gauges,

$$R^2_{mm} = \frac{1}{L}\sum_{i=1}^{L}\frac{[R^2_m(calibration)_i + R^2_m(validation)_i]}{2} \quad (5)$$

where $R^2_m(calibration)_i$ and $R^2_m(validation)_i$ are Nash–Sutcliffe coefficients during calibration and validation periods for gauge i and L is the total number of gauges.

Higher values of R^2_{mm} indicate better model mean performance.

The models also should be tested for their transferability. This is because the purpose of hydrological modelling is the application of the model for different conditions, for example different time periods or locations. A possibility to measure transferability is to compare model performances over the calibration and the validation periods. A coefficient of transferability T_m can be defined as

$$T_M = \max(R^2_m(calibration)_i - R^2_m(validation)_i, 0)_{i=1,\ldots,L} \quad (6)$$

A better model performance on the validation period can be considered as purely random; thus the difference (in Eq. (6)) was limited by zero.

Lower values of T_m indicate less loss of the model performance in the validation period and hence better model parameters' transferability.

The same statistical measures were also used in Bárdossy and Das (2008).

Results and discussion

The evaluated performance statistics for the calibration period (1961–1970) and the validation period (1971–1990) are presented in Table 3. With respect to overall performance measures, distributed model has the least conclusive performance in the calibration period. In particular, the Nash–Sutcliffe coefficient for the simulation of the daily discharge in the subcatchments during the calibration period has values ranging between 0.61 and 0.84, with a median value of 0.74 for the distributed structure. The minimum, maximum and median values of Nash–Sutcliffe coefficients obtained using the semi-distributed model structure are 0.68, 0.86 and 0.79, respectively. The values of the Nash–Sutcliffe coefficients obtained using the semi-lumped model structure ranges between 0.65 and 0.87, and have a median value of 0.78. The Nash–Sutcliffe coefficient obtained using the fully-lumped model structure ranges between 0.64 and 0.86, and has a median value of 0.77. The highest Nash–Sutcliffe coefficient value was observed at the catchment outlet Plochingen (Neckar) using all four model structures. There are higher errors in water balance (relative bias and root mean squared error) and peak discharge estimation using the distributed and fully-lumped model structure.

The model performance in the validation period also followed the same pattern observed in the calibration period – the distributed model does not outperform the other model structures with respect to any of the overall performance measures. The Nash–Sutcliffe coefficient for the simulation of the daily discharge in the subcatchments during the validation period has values ranging between 0.59 and 0.82, with a median value of 0.74 for distributed model, as shown in Table 3. The minimum, maximum and median values of the Nash–Sutcliffe coefficient obtained using the semi-distributed model are 0.55, 0.87 and 0.79, respectively. In the semi-lumped model structure, the values of the Nash–Sutcliffe coefficient range between 0.58 and 0.86, and have a median value of 0.77. The Nash–Sutcliffe coefficient obtained using the fully-lumped model structure ranges between 0.55 and 0.87, and has a median value of 0.77. As observed in the calibration period, the highest Nash–Sutc-

Table 3 Model performance statistics obtained using various HBV model structures during the 1961–1970 calibration period and 1971–1990 validation period

Model structure	R^2_m			Rel. bias			Peak error			RMSE		
	Min	Max	Median	Min	Max	Median	Min	Max	Median	Min	Max	Median
Calibration period												
Distributed	0.61	0.84	0.74	−0.11	0.16	0.02	−0.56	−0.07	−0.30	1.09	20.34	3.55
Semi-distributed	0.68	0.86	0.79	−0.03	0.02	0.01	−0.20	−0.06	−0.11	0.95	19.23	3.17
Semi-lumped	0.65	0.87	0.78	−0.01	0.08	0.02	−0.22	−0.06	−0.16	0.96	18.68	3.21
Fully-lumped	0.64	0.86	0.77	0.00	0.06	0.02	−0.25	−0.07	−0.13	0.99	18.87	3.32
Validation period												
Distributed	0.59	0.82	0.74	−0.03	0.29	0.10	−0.59	−0.01	−0.34	1.08	22.03	3.40
Semi-distributed	0.55	0.87	0.79	0.02	0.24	0.07	−0.20	0.07	−0.08	1.13	18.78	3.11
Semi-lumped	0.58	0.86	0.77	0.03	0.23	0.08	−0.23	0.03	−0.11	1.09	19.27	3.30
Fully-lumped	0.55	0.87	0.77	0.03	0.24	0.08	−0.24	0.04	−0.14	1.14	19.08	3.29

liffe coefficient value was observed at the catchment outlet Plochingen (Neckar) using all model structures. The lowest value of the Nash–Sutcliffe coefficient was observed at the gauge Rangendingen (Starzel), which is partly located within a karst geological formation. Among the four different model structures, the semi-distributed and semi-lumped structures show relatively higher mean values of the Nash–Sutcliffe coefficient, as also observed in the calibration period. There are also higher errors in water balance and peak discharge estimation using the distributed and fully-lumped model structure, as in the calibration period. Among the distributed and semi-distributed model structures, relatively higher differences are observed in smaller subcatchments.

This result is surprising for the calibration period as one would expect a better fit with a larger number of model parameters. The reasons for this can be problems with the parameter optimization as the larger number of parameters requires a larger number of iterations. Further, as the response surface becomes more complicated, the likeliness of ending in a local optimum is more likely for the distributed case. Even if the ''real'' optimum might not have been achieved, the results show, with a reasonable calibration effort, it was not possible to obtain better model performance for the distributed case.

A further source of error is the fact that precipitation input was obtained using interpolation. The individual high resolution (1×1 km^2) grids have higher estimation errors than the integrals corresponding to zones, subcatchments and catchments.

On average, it was observed that low flows and medium peak flows were well estimated, but higher peak flows were underestimated. The peak flows generally occurred during the winter period in the study catchment. Peak flow underestimation may be due to an improper representation of the snow accumulation and snow melting phenomena or the simplification adopted in the model to simulate water redistribution in the soil.

In Fig. 3, the Nash–Sutcliffe coefficients of different model structures at various gauges in the validation period are presented. The model performance obtained using the semi-distributed and semi-lumped model structures are relatively better, in general, as compared to the other model

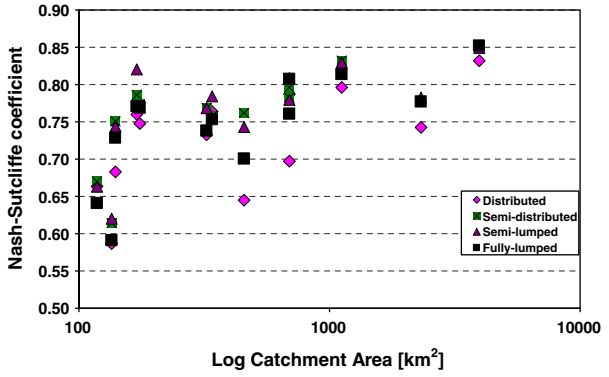

Figure 3 Nash–Sutcliffe coefficients obtained using various structures of the HBV model during the 1971–1990 validation period.

structures. This can be due to the ''overparametrization'' of the high and low resolution models. Another possible reason for this might be that the calibration of the higher and lower resolution models was overemphasizing differences of interpolated precipitation. The model performance is better for larger subcatchments using all the model structures (Fig. 3). This is due to a kind of error compensation; random errors corresponding to different subcatchments might cancel out on larger scales.

The performance of different model structures was also assessed at higher aggregation time steps to assess the correctness of water balance of longer time periods which is important when models are applied for purposes of water resource management. Regarding modelling of runoff at higher time steps (at a 7 day and 30 day time step), the performance of the models in terms of the Nash–Sutcliffe coefficient shows a similar trend (Table 4) as observed at the daily time step shown in Table 3. The increase of Nash–Sutcliffe coefficient with aggregation over increasing time periods occurs because the under- or overestimation of daily discharges does not matter when the model performance is estimated at higher aggregation time steps. A summary of the Nash–Sutcliffe coefficient at a 7 day and 30 day time step obtained using the different model structures in the validation period is shown in Table 4. As can be seen, the model performance improves with aggregation over increasing time periods using the all model structures. The distributed model does not outperform at 7 days aggregation time steps, even though at 30 days the distributed model performs similarly (median value) with the semi-distributed model structure. This indicates that the model error is diminishing at higher aggregation time periods.

The model mean performance and model parameters' transferability, calculated using Eqs. (5) and (6), respectively, are shown in Table 5. The model performances obtained at the Kirchentellinsfurt (Neckar), Wannweil (Echaz) and Riederich (Erms) gauges were not considered during these calculation. Kirchentellinsfurt (Neckar) was not considered due to the large number of missing measurements (61% days of the simulation period 1961–1990). The performances at the Wannweil (Echaz) and Riederich (Erms) gauges were not considered because the drainage area of those gauges is located within a karst formation. The highest value of model mean performance was observed using the semi-distributed and semi-lumped model structures, and least model mean performance was observed using the distributed model structure (Table 5). The least conclu-

Table 4 Statistics of Nash–Sutcliffe coefficients at 7 days and 30 days time step using various HBV model structures during the 1971–1990 validation period

Model structure	Nash–Sutcliffe coefficient					
	7 days time step			30 days time step		
	Min	Max	Median	Min	Max	Median
Distributed	0.69	0.88	0.82	0.67	0.90	0.86
Semi-distributed	0.78	0.89	0.84	0.80	0.91	0.86
Semi-lumped	0.78	0.89	0.82	0.80	0.91	0.85
Fully-lumped	0.78	0.89	0.82	0.80	0.91	0.85

Table 5 Model mean performance and parameter transferability using various HBV model structures

Model Structure	Model mean performance	Model parameters' transferability
Distributed	0.75	0.06
Semi-distributed	0.80	0.05
Semi-lumped	0.80	0.04
Fully-lumped	0.79	0.04

sive model parameters' transferability was observed for the distributed model structure. This indicates the optimized model parameters for the distributed model structure are less transferable for the validation period. As the same amount of information was used for each model resolution one might conclude that the distributed models were overparametrized, and thus could not be identified well from only the calibration period.

The Nash—Sutcliffe coefficient was also calculated for different years during the calibration and validation periods for different gauges to investigate the variation of the model performance on a year-to-year basis. On average, the standard deviation of the yearly Nash—Sutcliffe coefficients was relatively higher for the distributed model structure compared to the other model structures. This again suggests the problem of overparametrization and implies that the model performance using the distributed structure is not consistent. Additionally, this indicates averaging in space compensates for stochastic errors as stated above for aggregating over time.

There are high variations in spatial variability in winter (more advective precipitation) and summer precipitation (more convective precipitation). For that reason, it is also interesting to calculate the seasonal model performance. Fig. 4 illustrates the seasonal model performance in the validation period for the gauges at Horb (Neckar) and Plocingen (Neckar). The gauges at Horb (Neckar) and Plochigen (Neckar) were chosen because Horb is a representative gauge at the upper stream catchment and Plochingen (Neckar) is the outlet of the catchment. As shown in Fig. 4, the seasonal model performances are observed better, in general, for the semi-distributed model structure than the other models. But, the performance is comparatively poor in the summer season using all the model structures. Also, peak error (defined by Eq. (3)) was calculated in different seasons using the four models and there was found to be higher peak error in summer seasons than in winter seasons — indicating peak flows in summer time were not simulated as well as in the winter time (not shown). These results are from the convective precipitation events, which are more localized and are not well captured by the coarse raingauge networks used in the present study. Consequently, none of the model structures showed better model performance. Seasonal model performance, in general, is lower in the smaller drainage area than the higher drainage area, considered in Fig. 4.

It is worth noting that the model performances comparisons in this study were based on stream flow prediction at different subcatchment outlets. The results demonstrated that the highest-model resolution considered in this study did not provide the best model performance with respect to stream flow simulation. The additional information gained by interpolation for a high spatial resolution of the precipitation and temperature together with other model data (elevation, soil and land use information) did not provide any additional improvement with respect to stream flow simulation. It is important to note that the main forcing inputs, precipitation and temperature, for the different model structures were interpolated from the coarse observation network. The estimation of precipitation input from the point measurements to the grid cells used in the distributed model is more difficult.

Additionally, due to the high degree of model parameter interaction and dependency, the calibration task for a distributed model is even more difficult.

There were number of means to optimize the model parameters for the distributed model, including increasing the number of iterations employed in the automatic calibration scheme. Table 6 compares the model performance using two different attempts for the calibration of the model parameters of the distributed model structure. For the second attempt, the number of iterations was doubled (200 iterations) compared to the first attempt in the automatic optimization algorithm. There was an improvement in the calibration period as a result of doubling the number

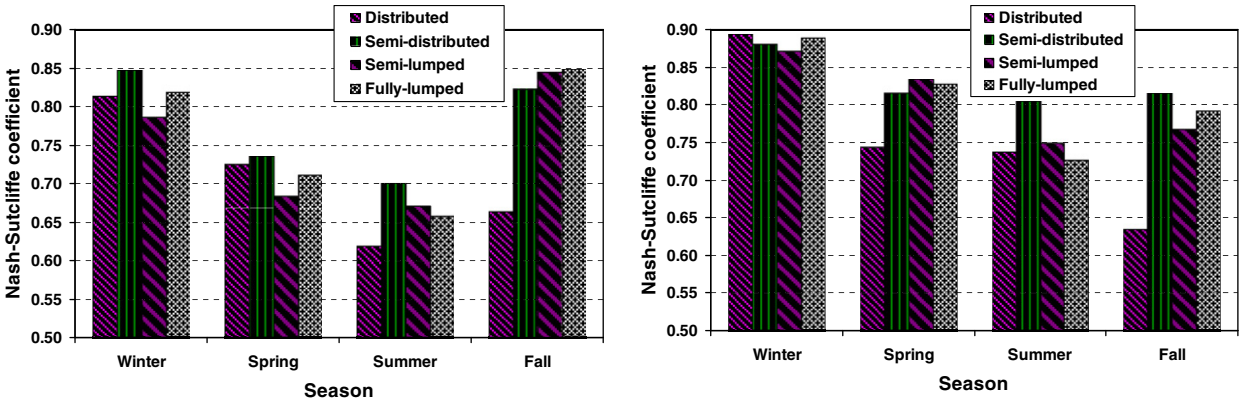

Figure 4 Seasonal model performance using various model structures for the 1971—1990 validation period for Necker River gauges at Horb (left) and Plochingen (right).

Table 6 Nash–Sutcliffe coefficients resulting from unique optimization schemes applied to the distributed model structure during the 1961–1970 calibration period and 1971–1990 validation period

	Rottweil (Neckar)	Oberndorf (Neckar)	Horb (Neckar)	Bad Imnau (Eyach)	Suessen (Fils)	Plochingen (Fils)	Plochingen (Neckar)
Calibration (I)	0.69	0.72	0.82	0.75	0.76	0.79	0.84
Calibration (II)	0.72	0.77	0.83	0.76	0.77	0.78	0.84
Validation (I)	0.70	0.72	0.76	0.76	0.81	0.77	0.82
Validation (II)	0.70	0.74	0.77	0.76	0.80	0.77	0.82

The number of repetitions was doubled to 200 iterations using the automatic optimization algorithm during the second model calibration (Calibration II) as compared to the initial calibration (Calibration I).

of iterations in the optimization scheme for five out of seven gauges as shown in Table 6, even though there was not any change in the model performance for the catchment outlet. Interestingly, there were slight improvements in model performance in the validation period for only two gauges, and there were eventually no changes for the remaining gauges. Also, the model performance did not improve considerably compared to the semi-distributed and semi-lumped model structures (not shown). Perhaps there was a higher compensation for the bias in the model input data during the calibration for the simpler model structures.

A great deal of effort was expended to optimize the model parameters using an automatic optimization algorithm. We have provided the results using the optimized model parameters obtained within a reasonable effort and as far as possible using the same information for each model. We also verified the results with re-running the optimization algorithm with changing random number (used in the optimization algorithm) and initial values of the parameters to be optimized. We argued that with a reasonable calibration effort it was not possible to obtain better model performance for the distributed model structure, even if the ''real'' optimum parameter values might not have achieved. Still, the potential effect of automatic calibration, including the initial configuration of the iteration, the number of iterations and acceptance criteria should not be neglected. A systematic investigation would be valuable to study the extent that they might influence the results.

Conclusions and outlook

The question was explored in this study as to whether input data having high spatial resolution and high model resolution leads to improved model results with respect to stream flow simulation. Four different model structures, with varying degrees of spatial resolution, based on the HBV model concept, were developed for simulating different hydrological processes in a mesoscale catchment. The models were applied to the Upper Neckar catchment, Germany, and verified using observed discharge data. The performance was then assessed using different model performance evaluation criteria.

For the present study catchment, semi-distributed and semi-lumped model structures outperformed the distributed and fully-lumped model structures. Among the model performances obtained using the distributed and semi-distributed structures, relatively higher differences were observed for smaller subcatchments and the differences in the model performances were less for the larger subcatchments.

The results suggested that more finely resolved input data and higher model resolution do not necessarily always improve the model performance, unless the input data corresponds to an increase of information.

The results of this study supported the general findings of Refsgaard and Knudsen (1996), El-Nasr et al. (2005), and Booij (2005), where eventually no significant improvement in the model simulations was observed due to high model resolution and higher model complexity. Also, the results of the DMIP project study showed that the lumped model outperformed distributed models in majority cases for the study catchments considered (Reed et al., 2004).

The findings of the presented study may be due to the following reasons:

(i) As noted by Beldring et al. (2003), highly resolved information about the physiographic characteristics of the catchment (for example, topography, soil and vegetation) might help in the determination of model parameters. Consequently, this may improve the stream flow estimation. However, the distributed model did not outperform the lower model resolution in the present study catchment. This can partially be attributed to the optimization procedure used for calibration. An advanced calibration of the distributed model would require assumptions of the spatial dependence (smoothness) of the model parameters.

Another possible explanation can be poor estimation of areal precipitation for the distributed model grid cells. The precipitation used in the present study was interpolated from the available point raingauges, which probably does not sufficiently account for the spatial variability of the precipitation. For example, the average variance of 20 raingauge stations (for the simulation period 1961–1990) within the catchment is 31.11 mm^2 day^2. Conversely, the average variance from the interpolated precipitation computed with the 20 raingauges located within the catchment in the simulation period is 25.02 mm^2 day^2.

(ii) In fact, the parameter values, as noted by Brath et al. (2004), may compensate for an incomplete representation of the precipitation field. Perhaps there was a higher compensation for the bias in the precipitation observations and other model input data by the cali-

bration for the relatively simpler model structures. Thus, in the semi-distributed and semi-lumped model structures, the error representing the spatial variability of precipitation and other model input data was possibly compensated by the model parameterization. However, the error in representing spatial variability of precipitation and other model input data in the finer model resolution dominates the bias compensation in the precipitation observations and other model input data by the calibration procedure for the distributed model structure. Additionally, the model parameterization for the distributed model structure is more challenging due to the parameters' large number of degrees of freedom and their underlying interactions.

(iii) The HBV model is originally a conceptually lumped model. Though a great deal of effort has been invested in accounting for spatial variability of precipitation and temperature in the ''distributed'' HBV model, it still falls short in its ability to account for full spatial variability due to its conceptual and lumped nature (for example no representation of lateral flow component in the designed ''distributed'' HBV model).

(iv) Additional model uncertainties can be incurred as a consequence of the interaction between the resolution of input data and the model structure itself.

The results of this study also indicate that physically-based distributed models may have the same problems. A systematic input error can be introduced by the insufficient representation of spatial variability of model input data; as such models are basically driven by data from point measurements with much coarser spatial resolution than the distributed model grid. Though, it is true that this may depend on scale, data availability, rainfall type (convective or advective) and seasonality of precipitation, and physiographic characteristics of the catchment (soil, vegetation, topography).

The main limitation of this study was that the comparison of different model structures was not carried out using the real highly distributed precipitation data. The use of fully, spatially resolved radar rainfall data or combination of radar-rain gauge rainfall data together with other highly resolved model input data can be used to carry out the comparison of different model structures for similar or different hydro-climatic regions to obtain the general findings of this study which is beyond the scope of this paper.

Acknowledgements

The work described in this paper is supported by a scholarship program initiated by the German Federal Ministry of Education and Research (BMBF) and coordinated by its International Bureau (IB) under the program of the International Postgraduate Studies in Water Technologies (IPSWaT). Its content is the sole responsibility of the authors and does not necessarily represent the official position or policy of the German Federal Ministry of Education and Research. We thank Jens Götzinger, Pawan K. Thapa, and Bill Schaal for their comments and suggestions on the initial version of the manuscript. We also wish to thank two anonymous reviewers for their constructive comments and suggestions, which substantially improved the manuscript.

References

Aarts, E., Korst, J., 1989. Simulated Annealing and Boltzmann Machines A Stochastic Approach to Combinatorial Optimization and Neural Computing. John Wiley & Sons Inc., Chichester.

Ahmed, S., de Marsily, G., 1987. Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity. Water Resources Research 23 (9), 1717–1737.

Atkinson, S.E., Woods, R.A., Sivapalan, M., 2002. Climate and landscape controls on water balance model complexity over changing time scales. Water Resources Research 38 (12), 1314, 50.1–50.17.

Atkinson, S.E., Sivapalan, M., Woods, R.A., Viney, N.R., 2003. Dominant physical controls of hourly streamflow predictions and an examination of the role of spatial variability: Mahurangi catchment, New Zealand. Advances in Water Resources 26 (2), 219–235.

Bárdossy, A., Das, T., 2008. Assessing the impacts of raingauge density on the simulation results of a hydrological model. Hydrological Earth System Sciences 12, 77–89.

Beldring, S., Engeland, K., Roald, L.A., Sælthun, N.R., Voksø, A., 2003. Estimation of parameters in a distributed precipitation-runoff model for Norway. Hydrological Earth System Sciences 7 (3), 304–316.

Bergström, S., Forsman, A., 1973. Development of a conceptual deterministic rainfall–runoff model. Nordic Hydrology 4, 170–174.

Beven, K.J., 2001. Rainfall–Runoff Modelling: The Primer. John Wiley and Sons, Chichester.

Booij, M.J., 2005. Impact of climate change on river flooding assessed with different spatial model resolutions. Journal of Hydrology 303, 176–198.

Bormann, H., 2006. Impact of spatial data resolution on simulated catchment water balances and model performance of the multi-scale TOPLATS model. Hydrological Earth System Sciences 10, 165–179.

Brath, A., Montanari, A., Toth, E., 2004. Analysis of the effects of different scenarios of historical data availability on the calibration of a spatially-distributed hydrological model. Journal of Hydrology 291, 232–253.

Butts, M.B., Payne, J.T., Kristensen, M., Madsen, H., 2004. An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation. Journal of Hydrology 298, 242–266.

Chaubey, I., Hann, C.T., Grunwald, S., Salisbury, J.M., 1999. Uncertainty in the models parameters due to spatial variability of rainfall. Journal of Hydrology 220, 46–61.

Christensen, N., Lettenmaier, D.P., 2007. A multimodel ensemble approach to assessment of climate change impacts on the hydrology and water resources of the Colorado River basin. Hydrological Earth System Sciences 11, 1417–1434.

Cunge, J.A., 1969. On the subject of a flood propagation computation method (Muskingum method). Journal of Hydraulic Research 7, 205–230.

El-Nasr, A.A., Arnold, J.G., Feyen, J., Berlamont, J., 2005. Modelling the hydrology of a catchment using a distributed and a semi-distributed model. Hydrological Processes 19, 573–587.

Ehret, U., 2002. Rainfall and flood nowcasting in small catchments using weather radar. Doctoral Thesis, University of Stuttgart, Germany.

Farmer, D., Sivapalan, M., Jothityangkoon, C., 2003. Climate, soil, and vegetation controls upon the variability of water balance in temperate and semiarid landscapes: downward approach to water balance analysis. Water Resources Research 39 (2), 1035.

Götzinger, J., Bárdossy, A., 2007. Comparison of four regionalisation methods for a distributed hydrological model. Journal of Hydrology 333, 374–384.

Haddeland, I., Matheussen, B.V., Lettenmaier, D.P., 2002. Influence of spatial resolution in a macroscale hydrologic model. Water Resources Research 38 (7), 1124–1133.

Hartmann, G., Bárdossy, A., 2005. In: Wagener, T., Stewart, F., Gupta, H.V., Bøgh, E., Bastidas, L., Nobre, C., de Oliveira Galão, C. (Eds.), Investigation of the Uncertainty of Hydrological Modeling for Climate Change Impact Assessment. Regional Hydrological Impacts of Climatic Change – Impact Assessment and Decision Making, vol. 295. IAHS Redbook, Publ., pp. 283–293.

Hargreaves, G.H., Samani, Z.A., 1985. Reference crop evapotranspiration from temperature. Applied Engineering in Agriculture 1, 96–99.

Hundecha, Y., Bárdossy, 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. Journal of Hydrology 292, 281–295.

Hundecha, Y., 2005. Regionalization of parameters of a conceptual rainfall runoff model. Doctoral Thesis, University of Stuttgart, Germany.

Krajewski, W.F., Lakshimi, V., Georgakakos, K.P., Jain, S.C., 1991. A monte carlo study of rainfall sampling effect on a distributed catchment model. Water Resources Research 27 (1), 119–128.

Kuo, W.-L., Steenhuis, T.S., McCulloch, C.E., Mohler, C.L., Weinstein, D.A., DeGloria, S.D., Swaney, D.P., 1999. Effect of grid size on runoff and soil moisture for a variable-source-area hydrology model. Water Resources Research 35 (11), 3419–3428.

Liang, X., Lettenmaier, D.P., Wood, E.F., Burges, S.J., 1994. A simple hydrologically based model of land surface water and energy fluxes for GSMs. Journal of Geophysics Research 99 (D7), 14415–14428.

Liang, X., Guo, J., Leung, L.R., 2004. Assessment of the effects of spatial resolutions on daily water flux simulations. Journal of Hydrology 298, 287–310.

Lindström, G., Johansson, B., Persson, M., Gardelin, M., Bergström, S., 1997. Development and test of the distributed HBV-96 hydrological model. Journal of Hydrology 201, 272–288.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models. Part I. A discussion of principles. Journal of Hydrology 10, 282–290.

Obled, C.H., Wendling, J., Beven, K., 1994. The sensitivity of hydrological models to spatial rainfall patterns: an evaluation using observed data. Journal of Hydrology 159 (1), 305–333.

Perrin, C., Michel, C., Andréassian, V., 2001. Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. Journal of Hydrology 242 (3–4), 275–301.

Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo, D.-J., DMIP Participants, 2004. Overall distributed model intercomparison project results. Journal of Hydrology 298(1–4) 27–60.

Refsgaard, J.C., Knudsen, J., 1996. Operational validation and intercomparisons of different types of hydrological models. Water Resources Research 32 (7), 2189–2202.

Romanowicz, A.A., Vanclooster, M., Rounsevell, M., La Junesse, I., 2005. Sensitivity of the SWAT model to the soil and land use data parametrisation: a case study in the Thyle catchment, Belgium. Ecological Model 187, 27–39.

Samaniego, L., 2003. Hydrological consequences of landuse/land cover and climatic changes in mesoscale catchments. Doctoral Thesis, University of Stuttgart, Germany.

Samaniego, L., Bárdossy, A., 2006. Simulation of the impacts of land use/cover and climatic changes on the runoff characteristics at the mesoscale. Ecological Modelling 196 (1–2), 45–61. doi:10.1016/j.ecolmodel.2006.01.00.

Smith, M.B., Koren, V.I., Zhang, Z., Reed, S.M., Pan, J.-J., Moreda, F., 2004a. Runoff response to spatial variability in precipitation: an analysis of observed data. Journal of Hydrology 298, 267–286.

Smith, M.B., Seo, D.-J., Koren, V.I., Reed, S.M., Zhang, Z., Duan, Q., Moreda, F., Cong, S., 2004b. The distributed model intercomparison project (DMIP): motivation and experiment design. Journal of Hydrology 298, 4–26.

Viney, N.R., Croke, B. F.W., Breuer, L., Bormann, H., Bronstert, A., Frede, T.H., Hubrechts, G.T., Huisman, J.A., Jakeman, A.J., Kite, G.W., Lanini, J., Leavesley, G., Lettenmaier, D.P., Lindström, G., Seibert, J., Sivapalan, M., Willems, P., 2005. Ensemble modelling of the hydrological impacts of land use change. In: Zerger, A., Argent, R.M. (Eds.), MODSIM 2005 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, pp. 2967–2973. ISBN: 0-9758400-2-9.

Yapo, P.O., Gupta, H.V., Sorooshian, S., 1996. Automatic calibration of conceptual rainfall–runoff models: sensitivity to calibration data. Journal of Hydrology 181, 23–48.

Zehe, E., Becker, R., Bárdossy, A., Plate, E., 2005. Uncertainty of simulated catchment scale runoff response in the presence of threshold processes: role of initial soil moisture and precipitation. Journal of Hydrology 315 (1–4), 183–202.