**Hydrology and
Earth System
Sciences**

Influence of rainfall observation network on model calibration and application

A. Bárdossy[1] and T. Das[1,*]

[1]Institute of Hydraulic Engineering, Universitaet Stuttgart, 70569 Stuttgart, Germany
[*]now at: Scripps Institution of Oceanography, University of California, San Diego

Abstract. The objective in this study is to investigate the influence of the spatial resolution of the rainfall input on the model calibration and application. The analysis is carried out by varying the distribution of the raingauge network. A meso-scale catchment located in southwest Germany has been selected for this study. First, the semi-distributed HBV model is calibrated with the precipitation interpolated from the available observed rainfall of the different raingauge networks. An automatic calibration method based on the combinatorial optimization algorithm simulated annealing is applied. The performance of the hydrological model is analyzed as a function of the raingauge density. Secondly, the calibrated model is validated using interpolated precipitation from the same raingauge density used for the calibration as well as interpolated precipitation based on networks of reduced and increased raingauge density. Lastly, the effect of missing rainfall data is investigated by using a multiple linear regression approach for filling in the missing measurements. The model, calibrated with the complete set of observed data, is then run in the validation period using the above described precipitation field. The simulated hydrographs obtained in the above described three sets of experiments are analyzed through the comparisons of the computed Nash-Sutcliffe coefficient and several goodness-of-fit indexes. The results show that the model using different raingauge networks might need re-calibration of the model parameters, specifically model calibrated on relatively sparse precipitation information might perform well on dense precipitation information while model calibrated on dense precipitation information fails on sparse precipitation information. Also, the model calibrated with the complete set of observed precipitation and run with incomplete observed data associated with the data estimated using multiple linear regressions, at the locations treated as missing measurements, performs well.

Correspondence to: T. Das
(tadas@ucsd.edu)

1 Introduction

Precipitation data is one of the most important inputs required in hydrological modeling and forecasting. In a rainfall-runoff model, accurate knowledge of precipitation is very essential for accurately estimating discharge. This is due to that fact that representation of precipitation is important in determining surface hydrological processes (Syed et al., 2003; Zehe et al., 2005). Beven (2001) noted that no model, however well founded in physical theory or empirically justified by past performance, will be able to produce accurate hydrograph predictions if the inputs to the model do not characterize the precipitation inputs. Precipitation is governed by complicated physical processes which are inherently nonlinear and extremely sensitive (Bárdossy and Plate, 1992). Precipitation is often significantly variable in space and time within a catchment (Krajewski et al., 2003). Singh (1997) provides detailed hydrological literature on the effect of spatial and temporal variability in hydrological factors on the stream flow hydrograph. Wilson et al. (1979) indicated that the spatial distribution and the accuracy of the rainfall input to a rainfall-runoff model influence considerably the volume of storm runoff, peak runoff and time-to-peak. Sun et al. (2002) demonstrated that errors in storm-runoff estimation are directly related to spatial data distribution and the representation of spatial conditions across a catchment. They found that the accuracy of storm-runoff prediction depends very much on the extent of spatial rainfall variability. However, Booij (2002) showed that the effect of the model resolution on extreme river discharge is much higher compared to the effect of the input resolution. Bormann (2006) indicated that high quality simulation results require high quality input data, but not necessarily always highly resolved data.

Raingauges are fundamental tools that provide an estimate of rainfall at a point. Generally, point measurements of raingauge accumulations are distributed in space over the catchment by interpolation techniques (e.g., kriging, Thiessen polygons, and inverse distance method). A large number

of earlier studies investigated the influence of the density of the raingauge network on the simulated discharge, with both real and synthetic precipitation and discharge data sets (Krajewski et al., 1991; Peters-Lidard and Wood, 1994; Seed and Austin, 1990; Duncan et al., 1993; St-Hilarie et al., 2003). Michaud and Sorooshian (1994) observed that inadequate raingauge densities in the case of the sparse network produced significant errors in the simulated peaks in a midsized semi-arid catchment. They also found considerable consistent reductions in the simulated peaks due to the spatial averaging of rainfall over certain spatial resolution. Many researchers have reported the effect of raingauge network degradation on the simulated hydrographs (e.g., Brath et al., 2004; Dong et al., 2005). More recently, Anctil et al. (2006) showed that model performance reduces rapidly when the mean areal rainfall is computed using a number of raingauges less than a certain number. They also observed that some raingauge network combinations provide better forecasts than when all available raingauges were used to compute areal rainfall. Nevertheless, inadequate representation of spatial variability of precipitation in modeling can be partially responsible for modeling errors. This may also lead to the problem in parameter estimation of a conceptual model. Chaubey et al. (1999) found large uncertainty in estimated model parameters when detailed variations in the input rainfall were not taken into account. Oudin et al. (2006) observed that random errors in rainfall input considerably affect model performance and parameter values, although, model results were nearly insensitive to random errors in potential evapotranspiration input. They also indicated that the sensitivity of a rainfall-runoff model to input errors might depend partially on the model structure itself. Chaplot et al. (2005) investigated the effect of the accuracy of spatial rainfall information on the modeling of water, sediment, and NO_3-N loads for two medium sized catchments under a range of climates, surface areas and environmental conditions. They observed that at both catchments, runoff and nitrogen fluxes are varied only slightly with decreasing gauge concentration. They argued that model performance is only slightly affected by data errors because they are able to adjust their parameters in order to compensate for input errors within a reasonable range.

Therefore, it may be of interest to investigate the results of the simulations obtained with the rainfall input when the model is parameterized according to a different type of input data. In fact, it is, frequently the case that a raingauge network changes due to an addition or subtraction of raingauges. The raingauge network can be strengthened by the addition of new instruments or by using weather radar, so that a more detailed representation of rainfall is allowed, but for calibration purposes, past observations are available only over the original, less numerous measuring points. Conversely, in the case of an operational flood forecasting system, the opposite situation may occur. In the flood forecasting system, the rainfall-runoff model is usually calibrated using all the available flood events and precipitation data. However, during the operational forecasting time, the precipitation data from all past observation stations may not be available due to a malfunctioning of a few of the observations in the network or the observation data may not be available online. In such cases, it is important to understand if the parameters calibrated using the rainfall coming from one type of network have the ability to represent the phenomena governing the rainfall-runoff process with the input provided by the different configuration of the raingauge network.

Therefore, the aim of this paper is to investigate the influence of rainfall observation networks on model calibration and application. A method based on the combinatorial optimization algorithm simulated annealing (Aarts and Korst, 1989) is used to identify a uniform set of locations for a particular number of raingauges. First, the semi-distributed conceptual rainfall-runoff model HBV is used to investigate the effect of the number of raingauges and their locations on the sensitivity of the hydrological model results. The hydrological modeling performances of the networks are being analyzed through the comparison of Nash-Sutcliffe coefficient and other goodness-of-fit indices. Secondly, the influence of the rainfall observation network on model calibration and application is examined. This study seeks to determine whether the parameters calibrated using the rainfall coming from one type of network have the ability to represent the phenomena governing the rainfall-runoff process with the input provided by a different configuration of the raingauge network. The model is calibrated using precipitation interpolated from different raingauge networks. The calibrated model is then run for the validation period using the precipitation obtained from the raingauge network, which was not used for the calibration. Lastly, the simulation experiments are being carried out to analyze the reliability of supplementing missing precipitation measurements used for the calibration with data estimated using a multiple linear regression and running the model using that precipitation combined with available observed precipitation.

2 The study area and data

The upper Neckar catchment, located in southwest Germany, was selected as test catchment. The study area covers an area of approximately $4000\,km^2$. The study catchment area was divided into thirteen subcatchments depending on the available discharge gauges (Fig. 1).

The sizes of different subcatchments of the upper Neckar catchment are summarized in Table 1. The elevation for the catchment ranges from about 250 m a.s.l. to around 1000 m a.s.l., with a mean elevation of 546 m a.s.l. Slopes in general are mild; approximately 90% of the area has slopes varying from 0° to 15°. Though slopes in some areas in the Swabian Jura or in the Black Forest may have values up to 50°. The physiographical factors considered
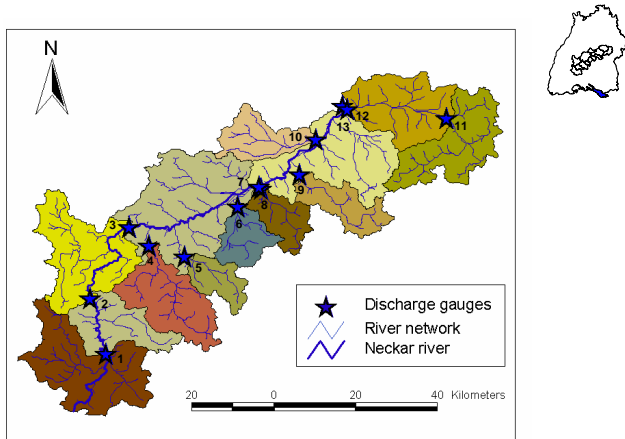
Fig. 1. Study area: Upper Neckar catchment showing different subcatchments and discharge gauges (Upper-right: location of the study domain in the state of Baden-Württemberg in southwest Germany).

Table 1. Summary of the sizes of the different subcatchments. The table also contains the drainage area of each discharge gauges.

	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 Blaesibg (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

in this study were derived from different sources: (1) Digital Elevation Model with a spatial resolution of 50 m×50 m; (2) a digitized soil map of the state of Baden-Württemberg at the scale 1:200 000 and (3) Land use map (LANDSAT satellite image for the year 1993) with a spatial resolution of 30 m×30 m. Daily discharge data from 13 gauging stations was used for model evaluation. All data was provided by the State Institute for Environmental Protection Baden-Württemberg (LUBW). The daily precipitation total, daily maximum and minimum temperature data distributed in and around the study catchment were acquired from the German Weather Service (DWD). The climate of the study area is characterized by warm-to-hot summers with generally mild winters. The coldest and hottest months in the study area are January and July respectively. The daily mean air temperature in January is about −0.8°C and in July is about 17°C according to the daily mean temperature records available for the period from 1961 to 1990. The annual variation of precipitation in the study area shows a multi-modal distribution. June and October are the wettest and driest months, with monthly means of 126 mm and 64 mm respectively, according to the daily amount of raingauge records available for the period from 1961 to 1990. The mean annual precipitation observed during this period is 908 mm. The study area has experienced some land use transitions from crop land or grass land to built-up area or industrial usages in the last several decades (Samaniego, 2003). The use of land cover information in the HBV model, applied in this study, is static. We have used land use information from the LANDSAT satellite image for the year 1993. The uncertainty might be introduced due to this reason and should be explored, but it was not considered in this study since it was not the primary aim.

2.1 Raingauge selection method and data preparation

The raingauges that have no missing measurements for the period from 1961 to 1990 and are located within or up to 30 km from the study catchment were used as a basis of complete raingauge network. The raingauge networks were selected from the complete network, consisting of 51 raingauges, using the combinatorial optimization algorithm simulated annealing (Aarts and Korst, 1989). The main idea behind the raingauge selection algorithm is to identify a uniform set of locations for a particular number of raingauges over the catchment. This means that for number of stations objective functions consisting of the mean distance of the station to the whole catchment and the minimum distance between the stations were considered. While the first objective was minimized to get internal stations and the second objective was maximized so not to take stations that were very close to each other. The selection algorithm was applied repeatedly to obtain optimal locations of different number of raingauges. Seven networks consisting of different number of raingauges ranging from 5 to 51 were obtained. Figure 2 shows the spatial distribution of the selected raingauge networks.

The basic inputs for the HBV model are precipitation, air temperature and potential evapotranspiration. The point measurements obtained from the selected raingauge networks were interpolated on a 1 km^2 grid using the external drift kriging method (Ahmed and de Marsily, 1987). The networks were kept constant for each subcatchment during the interpolation of meteorological variables. This means that the weights were adjusted. For subcatchments some of the weights might have become very small – in this sense the number of raingauges with significant influence might be smaller. Due to the fact that the selected rainfall networks are
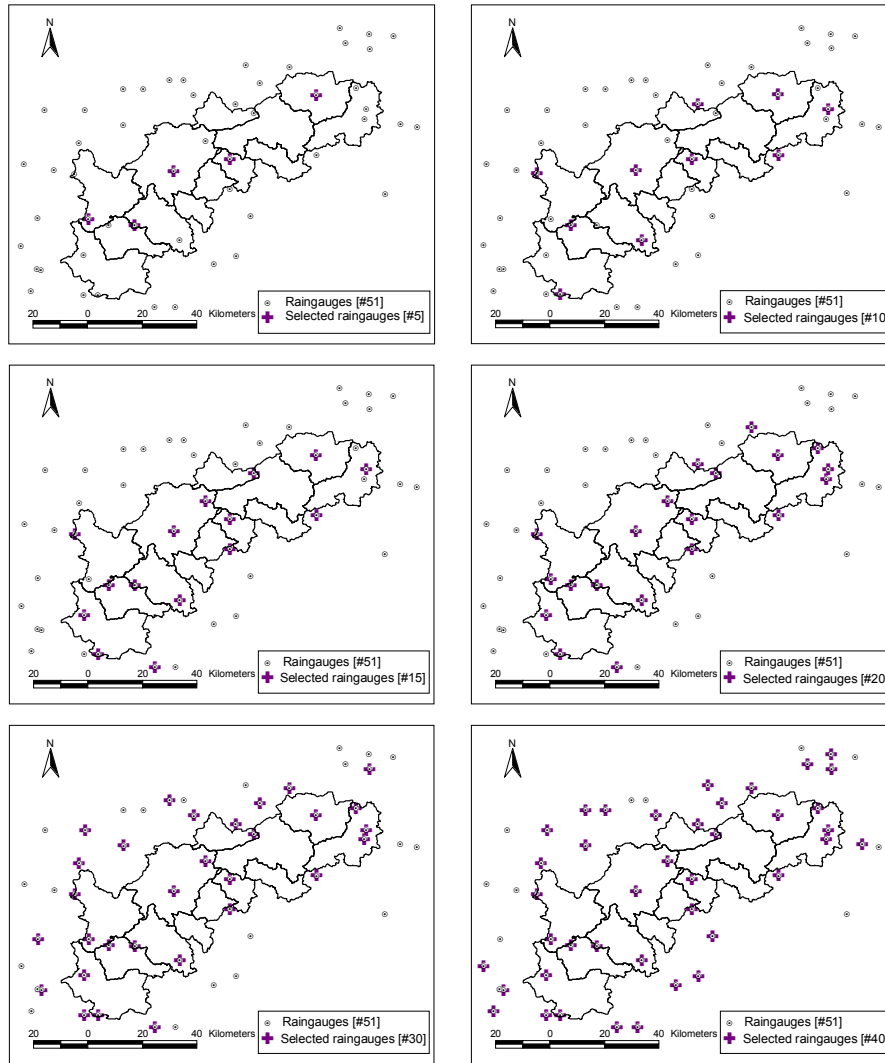
Fig. 2. Spatial distribution of the selected raingauge networks.

uniformly dense over the whole catchment they are more or less uniform for each subcatchment.

Note that the rate of increase of precipitation decreases with increasing elevation. The square root of the topographic elevation was assumed as a good approximation to account for this variation and it was used as the drift variable for precipitation (Hundecha and Bárdossy, 2004). Because the temperatures show a fairly constant lapse rate, topographic elevation was used as the drift variable for interpolating the temperature from the available point measurements. For precipitation the experimental variogram is calculated for each day when the daily precipitation amount exceeds some threshold values (maximum greater than 10 mm or mean greater than 5 mm). The experimental variogram is then fitted with theoretical variogram using automatic fitting procedure. The average variogram is used in the remaining days when the daily precipitation amount is low. The average variogram is also used through out the whole period for smaller network densities (e.g., 5 raingauge network) as experimental variograms should not be calculated from such small samples. For temperature interpolation the average variogram is used for every day. A combination of two theoretical variogram models, the spherical variogram and the pure nugget effect variogram (Kitanidis, 1997), is used in this study. Figure 3 shows the average daily variance over each grid of the catchment. The variance is calculated, using the interpolated precipitation computed with the 51 raingagues, over the calibration and validation periods. There is high variability in precipitation in the northwest part, close to Black forest, and southeast part, close to Swabian Alb, of the catchment. The temporal variability of precipitation is non homogeneous, which might influence the quality of interpolation and hydrological model results.
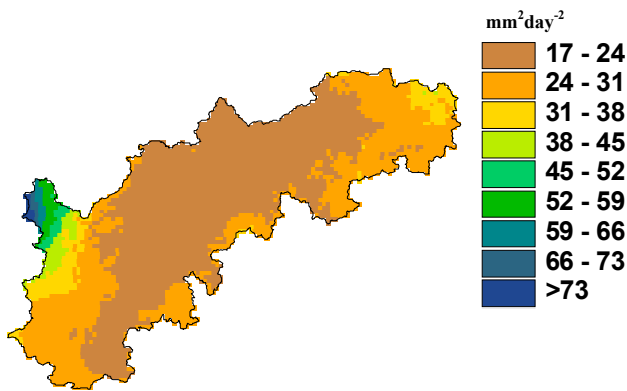
Fig. 3. Average daily variance ($mm^2 day^{-2}$) over each grid of the catchment. The variance is calculated, using the interpolated precipitation computed with the 51 raingagues for the period 1961-1990.

Fig. 4. Average daily variance ($mm^2 day^{-2}$) of areally averaged precipitation vs. number of raingauges. The variance is calculated, using the interpolated precipitation computed with different raingague networks for the period 1961–1990.

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.

Figure 4 depicts the average daily variance obtained using the averaged areal precipitation over the catchment, computed with the selected network densities, for the simulation period (1961–1990). As can be seen that the variability of the interpolated precipitation decreases with the increasing number of raingauges, but there is no change in the variability beyond a certain number of raingauges. The reason for this is that the contribution of individual stations to the areal average is the higher if less number of stations are used for interpolation. Thus, the interpolation using the smallest number of observations resembles the most variances calculated for each single location.

3 Model and methods

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 modeled is divided into a number of subcatchments and each subcatchment is further divided into a number of zones based on 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 accumulation and melt 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 sub-catchment scale are conceptually represented by two linear reservoirs. The upper reservoir simulates the near surface
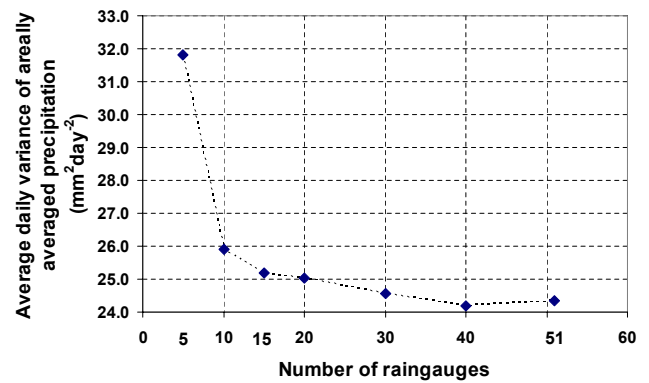
and interflow in the sub-surface layer, while the lower reservoir represents the base flow. Both reservoirs are connected in series by a constant percolation rate. Finally, there is a transformation function, consisting of a triangular weighting function with one free parameter, for smoothing the generated flow. The flow is routed from one node to the other of the river network by means of the Muskingum method. Additional description on the HBV model can be found in Lindström et al. (1997) and Hundecha and Bárdossy (2004).

For this study, we considered topographic elevation in defining the zones. This is due to the reason that elevation affects the distribution of the basic meteorological variables such as precipitation and temperatures as well as the rate of evaporation and snow melt and accumulation. Elevation zones were defined using a contour interval of 75 m. The elevation of the study area varies from about 250 m to around 1000 m, so therefore a maximum of 10 elevation zones were defined in each subcatchment. The mean daily precipitation amount and the mean daily temperature were assigned as input to each zone. The meteorological variables for each zone were estimated as the mean of the interpolated values on the regular grids of $1 km^2$ located within a given zone. The potential evapotranspiration was also averaged over each zone from the potential evapotranspiration calculated on $1 km^2$ grids located within a given zone.

3.1 Model calibration and simulations

The HBV model was calibrated using the interpolated precipitation obtained from the different raingauge networks. Other input data, daily mean temperature and daily potential evapotranspiration, were kept constant for each calibration. The automatic calibration method based on the combinatorial optimization algorithm simulated annealing (Aarts and Korst, 1989) was used to optimize the model parameters. For this optimization, an objective function composed

of the Nash-Sutcliffe coefficient corresponding to daily and annual time steps was maximized, while a reasonable range was fixed to constrain model parameters. Note that the model calibration and validation were performed using the daily discharges at each subcatchment outlet (total number of sub-catchments is 13). During the calibration, the model parameters of the head water subcatchments were optimized before the mixed subcatchments.

The standard split sampling model calibration procedure was followed. The model calibration period runs from 1961 to 1970. The subsequent period up to 1990 was used to validate the calibrated model. The interpolated precipitation, based on daily recorded observations, from the different raingauge networks was used to simulate the model discharges for the calibration and validation periods. The meteorological conditions do not differ strongly between the calibration and validation periods for the same raingauge network.

3.2 Simulation comparison statistics

The simulation results obtained using different raingauge networks were compared using the following different statistical criteria: the Nash-Sutcliffe coefficient, the relative bias, the peak error and the root mean squared error.

The Nash-Sutcliffe coefficient (R_m^2) (Nash and Sutcliffe, 1970) is defined as

$$R_m^2 = 1 - \frac{\sum\limits_{i=i}^{N} (Q_s(t_i) - Q_o(t_i))^2}{\sum\limits_{i=1}^{N} (Q_o(t_i) - \overline{Q_o})^2} \qquad (1)$$

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 relative bias was computed to examine the model performance with regard to its ability to maintain the water balance. Additionally the peak error was calculated to check the model's estimation capacity of the peak flow.

The relative bias (rel. bias) is defined as:

$$\text{rel.bias} = \frac{\sum\limits_{i=1}^{N} (Q_s(t_i) - Q_o(t_i))}{\sum\limits_{i=1}^{N} Q_o(t_i)} \qquad (2)$$

Accordingly, 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 $\overline{Q_{o\,(\max)}}$:

$$\text{peak error} = \frac{\bar{Q}_{s(\max)} - \bar{Q}_{o(\max)}}{\bar{Q}_{o(\max)}} \qquad (3)$$

The root mean squared error (*RMSE*) was also calculated.

Further more, the mean model performance (R_{mm}^2) is calculated using the Nash-Sutcliffe coefficient values obtained at the discharge gauges during the calibration and validation periods.

$$R_{\text{mm}}^2 = \frac{1}{L} \sum_{i=1}^{L} \frac{[R_m^2(\text{calibration})_i + R_m^2(\text{validation})_i]}{2} \qquad (4)$$

where $R_m^2(\text{calibration})_i$ and $R_m^2(\text{validation})_i$ are Nash-Sutcliffe coefficients during calibration and validation periods for gauge i and L is the total number of gauges.

Note that higher values of R_{mm}^2 indicate better mean model performance.

Models are not developed for reproducing known past observations. The purpose is to apply them under different conditions (weather, climate or land use). Therefore model quality should also be measured from this viewpoint. The value of model parameters' transferability (T_m) is computed by subtracting the model performance for the validation period from the model performance obtained in the calibration period.

$$T_m = \max \left(R_m^2(\text{calibration})_i - R_m^2(\text{validation})_i, 0 \right)_{i=1,\ldots,L} \qquad (5)$$

A better performance on the validation period could be considered as purely random, thus the difference is limited by zero. Lower values of T_m indicate less loss of the model performance in the validation period and better model parameters' transferability.

Additionally, the mean absolute error and root mean squared error were calculated using the model simulated and observed discharges for each annual maximum flood event.

4 Results and discussion

4.1 Model results using different raingauge densities

A summary of the model performance for the calibration and validation periods for selected three gauges is shown in Table 2. The model performances are shown for the gauges at Horb (Neckar) and Suessen (Fils) because these two gauges are representative gauges located in the upstream and downstream of the catchment and Plochingen (Neckar) are being the outlet of the catchment. The network consisting of 5 raingauges yields the minimum model performance both the calibration and validation periods, whereas the highest model performance was observed using the 20 raingauge network for the gauge at Horb (Neckar) in the validation period. Moreover, increasing the raingauge numbers above 20 did not improve the model performance. The best model performance in the validation period for Suessen (Fils) was observed using the 15 raingauge network. On the other hand, the best model performance for the Plochingen (Neckar) was observed using the 30 raingauge network. This shows the influence of the spatial distribution of raingauges within each

Table 2. Model performance using precipitation interpolated from different numbers of raingauges in the calibration and validation periods.

| Number of raingauges | Horb (Neckar) | | | | Calibration period Suessen (Fils) | | | | Plochingen (Neckar) | | | |
	R_m^2	Rel. bias	Peak error	RMSE	R_m^2	Rel. bias	Peak error	RMSE	R_m^2	Rel. bias	Peak error	RMSE
5	0.82	−0.05	−0.17	7.70	0.72	0.00	−0.14	3.33	0.84	−0.03	−0.06	20.54
10	0.83	0.04	−0.10	7.37	0.77	−0.02	−0.12	3.08	0.86	0.00	−0.10	18.80
15	0.86	0.01	−0.13	6.76	0.75	−0.02	−0.12	3.18	0.87	0.01	−0.10	18.53
20	0.86	0.02	−0.11	6.80	0.77	0.01	−0.10	3.03	0.87	0.00	−0.08	18.42
30	0.85	0.02	−0.08	6.86	0.77	−0.01	−0.11	3.02	0.88	−0.01	−0.12	18.59
40	0.85	0.02	−0.08	7.04	0.77	−0.03	−0.10	3.05	0.86	0.00	−0.07	18.97
51	0.84	0.04	−0.05	7.24	0.76	0.00	−0.12	3.11	0.86	−0.02	−0.08	19.13

| | Horb (Neckar) | | | | Validation period Suessen (Fils) | | | | Plochingen (Neckar) | | | |
	R_m^2	Rel. bias	Peak error	RMSE	R_m^2	Rel. bias	Peak error	RMSE	R_m^2	Rel. bias	Peak error	RMSE
5	0.81	0.06	−0.12	8.47	0.76	0.08	−0.19	3.03	0.84	0.05	−0.01	20.85
10	0.81	0.05	−0.11	8.61	0.79	0.09	−0.14	2.82	0.87	0.04	−0.09	19.20
15	0.83	0.09	−0.12	8.05	0.80	0.09	−0.19	2.76	0.87	0.07	−0.06	18.96
20	0.85	0.09	−0.12	7.99	0.79	0.13	−0.15	2.86	0.87	0.06	−0.06	18.99
30	0.84	0.09	−0.09	7.80	0.80	0.10	−0.17	2.78	0.89	0.05	−0.10	18.65
40	0.83	0.10	−0.09	7.99	0.79	0.09	−0.15	2.81	0.86	0.06	−0.06	19.35
51	0.82	0.11	−0.07	8.16	0.77	0.12	−0.16	2.93	0.87	0.04	−0.06	19.10
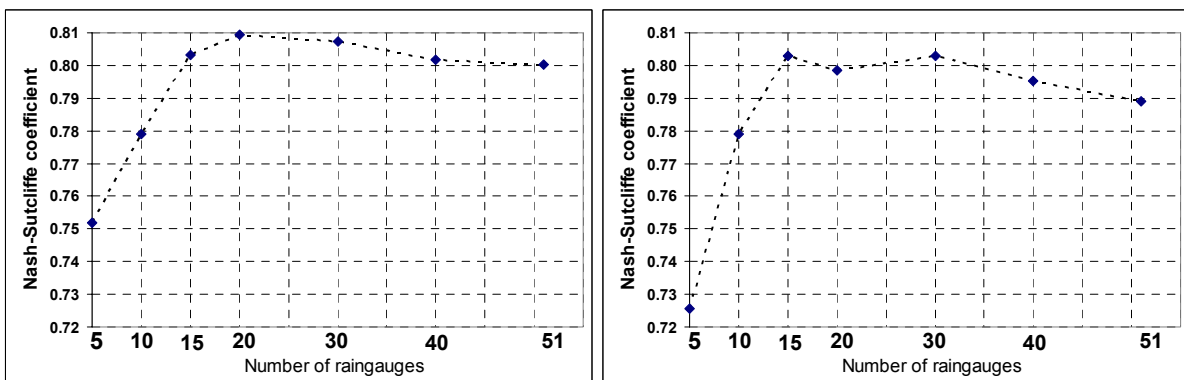
Fig. 5. Mean Nash-Sutcliffe coefficient calculated on the daily time step resulting from different raingauge networks for the calibration period (left panel) and validation period (right panel).

subcatchment. It seems that the hydrological model used for this study can be well adjusted to the different precipitation observation densities. A higher spatial model resolution might have lead to different results. The number of raingauges within a subcatchment is not necessarily increasing with the increase of the total number of stations. For example, for Horb the number of stations within the catchment is 5 for the 15 raingauge network and 6 for all denser networks. It can be seen that all stations within the whole catchments are included in all networks consisting of at least 20 raingauges (Fig. 2). The improvement in precipitation interpolation seems to be very small, as expected, if raingauges outside the investigation area are also considered. Thus, the subsequent hydrological modeling cannot be improved.

Figure 5 shows the mean Nash-Sutcliffe coefficient calculated on the daily time step for the calibration and validation periods. The mean values were calculated from the Nash-Sutcliffe coefficients obtained at all gauges over the catchment. The mean values were calculated to assess the model performance on the calibration and validation periods separately, and to check the model transferability.

A considerable deterioration in model performance was observed when using the network consisting of 5 raingauges, for both the calibration and validation periods (Fig. 5). The mean performances also show that consideration of stations located at far outside the catchment cannot improve the precipitation interpolation considerably so that it could be reflected through improved model performance.
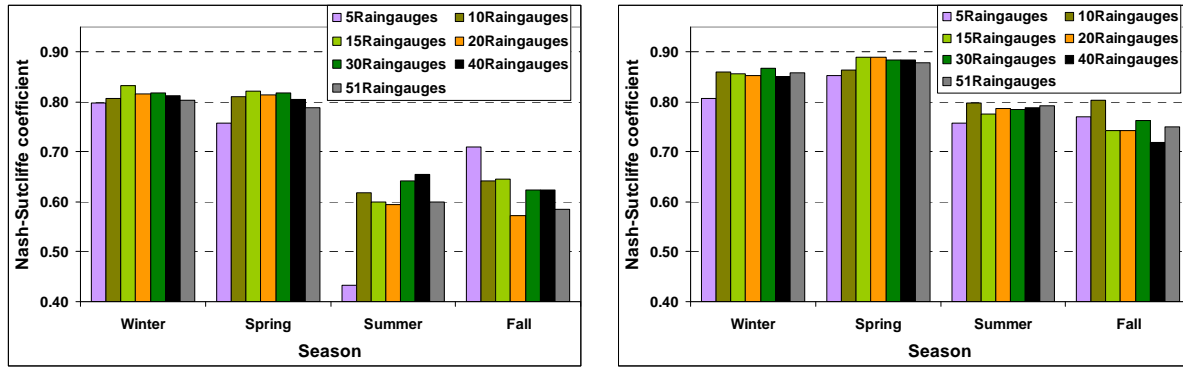
Fig. 6. Seasonal Nash-Sutcliffe coefficients of the daily step using the precipitation interpolated from different number of raingauges during the validation period for the gauges at Suessen (Fils) (left panel) and Plochingen (Neckar) (right panel).

Table 3. Mean model performance and parameters' transferability obtained using the precipitation interpolated from different raingauge networks. The values are calculated from the Nash Sutcliffe coefficients on daily time step at all the discharge gauges over the catchment.

Number of raingauges	Mean model performance	Model parameters' transferability
5	0.74	0.12
10	0.78	0.03
15	0.80	0.04
20	0.80	0.04
30	0.82	0.04
40	0.80	0.05
51	0.80	0.05

Table 3 represents the mean model performance and model parameters' transferability, calculated using the Eqs. (4) and (5) respectively, corresponding to the different raingauge networks.

The mean transferability of the model is nearly constant for all precipitation networks except the 5 raingauge case. The reason might be the model resolution lead to an "over fit" for this case. The low transferability values indicate that the model parameters could be reasonably assessed if at least 10 precipitation stations were considered. A least value of mean model performance was observed using the network consisting of 5 raingauges.

There is a considerable difference in the seasonal variability of precipitation; winter precipitation covers the area more or less uniformly, whereas, convective rainfall in summer leads to high spatial variability. Therefore a seasonal investigation of the model performances is reasonable. Figure 6 shows the seasonal model performance for the Suesen (Fils) and Plochingen (Neckar) gauges.

Figure 6 indicates that the poorest model performance for

all raingauge densities was observed in summer. The results corresponding to the 5 raingauge network are the least conclusive. All other networks lead to similar performances. This is not surprising as external stations cannot capture local convective events. A better performance could be expected only from a denser internal network.

The purpose of modeling is often related to floods; therefore the model performance for large events is of special interest. Subsequently event statistics were calculated for each annual maximum flood event. Figure 7 shows the mean absolute error and root mean squared error for the gauge at Horb (Neckar).

The results show a considerable scatter. Differences in the performance are mainly event dependent. On average, the absolute error with respect to the annual maximum discharges for the gauge at Horb (Neckar) ranges between 6.9% and 8.4% using the precipitation interpolated from varying raingauge networks. For the 5 and 10 station networks for some events unusually bad performances could be observed. In general the higher densities lead to slightly better model results. The results support the findings corresponding to the mean performance, specifically that improvements will mainly be achieved by using all internal observations.

The simulations were not insensitive to the spatial variability of the precipitation fields obtained using different number of raingauges; the model results using interpolated precipitation from different raingauge networks did not produce the same hydrograph (not shown). In general, this study showed that using too coarse a raingauge network for estimating the rainfall fields can give rise to remarkable poor hydrological simulation results, even though, the simulation results can not be improved further with addition of more raingauges beyond a certain number. The results are in agreement with other similar hydrological studies, for example Brath et al. (2004) and Dong et al. (2005).
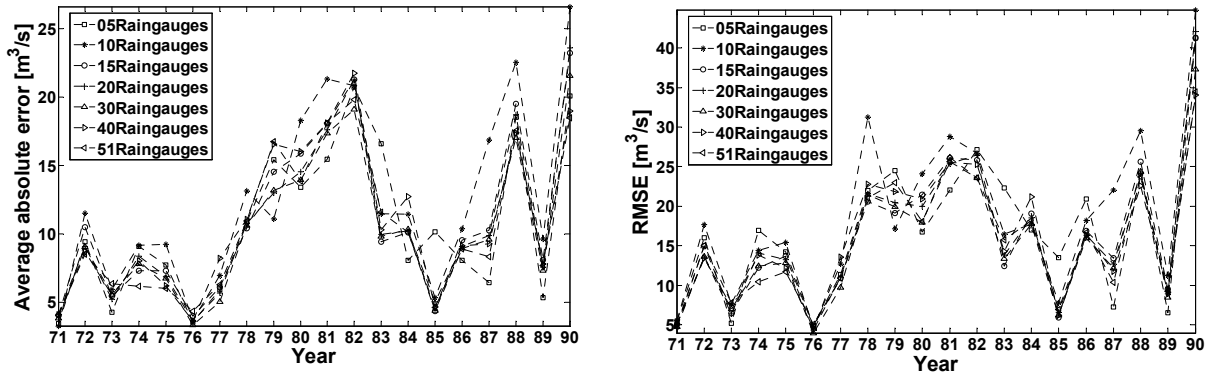
Fig. 7. Event statistics for each annual maximum flood event using different raingauge networks during the validation period for the gauge at Horb (Neckar): mean absolute error (left panel) and root mean squared error (right panel).

4.2 Influence of the rainfall observation network on model calibration and application

In the following section, the aim of the simulation experiment was to investigate the influence of the spatial resolution of the rainfall input on the calibration of a conceptual model. First, the semi-distributed HBV model was calibrated with the precipitation interpolated from the available observed rainfall of varying raingauge networks. The calibrated model was then run using the same precipitation used for the calibration as well as interpolated precipitation based on networks of reduced and increased raingauge density.

For example, the model was first calibrated using precipitation interpolated from 10 and 20 raingauges. The calibrated model using 10 raingauges was then run using precipitation obtained from 20 raingauges for the validation period and vice versa. This experiment is indicated in tables and figures, later on, as follows: **10/10** is calibrated with 10 raingauges and simulated with 10 raingauges, **20/20** is calibrated with 20 raingauges and simulated with 20 raingauges, **10/20** is calibrated with 10 raingauges and simulated with 20 raingauges and **20/10** is calibrated with 20 raingauges and simulated with 10 raingauges.

The model calibrated using less detailed precipitation (precipitation from 10 raingauges) slightly improves more often when it was run using relatively more detailed precipitation (precipitation from 20 raingauges) (Table 4). On the other hand, the model performance obtained using precipitation from 20 raingauges deteriorated when the same model was run using precipitation obtained from 10 raingauges. Three of the five subcatchments, shown in Table 4, demonstrate the same pattern.

The parameter values, as noted by Brath et al. (2004), may compensate for an incomplete representation of the precipitation field within a reasonable range. Yet, only provided the model parameters were updated by performing recalibration, for which the new input precipitation was estimated from the reduced raingauges network. But there was
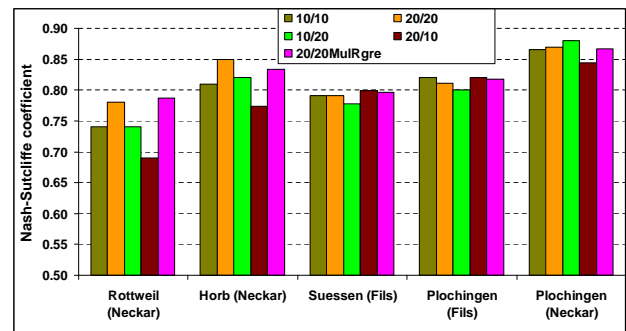

Fig. 8. Nash-Sutcliffe coefficient on the daily time step obtained using different level of precipitation input information for the validation period for selected five gauges.

no such type of compensation for the second case when the calibrated model using 20 raingauges was run using precipitation obtained from the 10 raingauge network. This demonstrates the inability of the 10 raingauges to adequately represent the precipitation field for the catchment.

The reason for bad performance with missing data is due to possible systematic differences in precipitation. These can occur due to different elevations or geographical exposition of the whole dataset compared to the remaining stations. A systematic under or overestimation should be avoided.

However, these can be reduced by a data filling algorithm. The following simulation experiment was carried out in order to investigate whether the estimated precipitation at raingauges with missing values (for example offline stations only used for model calibration), together with the precipitation data from the remaining stations that were used during the model calibration, has any benefit over the model operated by precipitation from the reduced raingauges. As the first step missing precipitation values are estimated using a multiple linear regression (Montgomery and Peck, 1982) based on the long available common time series. The interpolation was then carried out using external drift kriging treating the

Table 4. Model performances using the input precipitation information obtained from different number of raingauges.

		Rottweil (Neckar)	Horb (Neckar)	Suessen (Fils)	Plochingen (Fils)	Plochingen (Neckar)
10/10	R_m^2	0.74	0.81	0.79	0.82	0.87
	Rel. bias	−0.05	0.05	0.09	0.12	0.04
	Peak error	−0.23	−0.11	−0.14	−0.16	−0.09
	RMSE	3.74	8.61	2.82	5.03	19.20
20/20	R_m^2	0.78	0.85	0.79	0.81	0.87
	Rel. bias	0.00	0.09	0.13	0.17	0.06
	Peak error	−0.07	−0.12	−0.15	−0.16	−0.06
	RMSE	3.49	7.99	2.86	5.07	18.99
10/20	R_m^2	0.74	0.82	0.78	0.80	0.88
	Rel. bias	−0.18	−0.03	0.13	0.19	0.01
	Peak error	−0.36	−0.25	−0.13	−0.12	−0.13
	RMSE	3.75	8.10	2.91	5.22	18.06
20/10	R_m^2	0.69	0.77	0.79	0.82	0.84
	Rel. bias	0.13	0.19	0.09	0.10	0.08
	Peak error	0.08	0.00	−0.16	−0.20	−0.01
	RMSE	4.28	9.32	2.79	4.95	20.84

10/10: calibrated with 10 raingauges and simulated with 10 raingauges; **20/20:** calibrated with 20 raingauges and simulated with 20 raingauges; **10/20:** calibrated with 10 raingauges and simulated with 20 raingauges and **20/10:** calibrated with 20 raingauges and simulated with 10 raingauges.

values estimated for the missing stations as uncertain data. The model, calibrated with the precipitation data obtained from 20 raingauges, was then run in the validation period using the precipitation field above described.

Figure 8 shows the model performance for selected five gauges during the validation period using the different level of input precipitation information. The data shown in Table 4 is partly used to prepare Fig. 8 for easier comparison. In the following tables and figures **20/20MulRgre** indicates model calibrated with 20 raingauges and simulated with 20 raingauges (rainfall estimated at 10 locations considered as missing measurements). The results from five gauges are shown as they are representative and also these gauges are wide spread in upstream and downstream over the catchment.

The model performed well when it was calibrated using precipitation from 20 raingauges and was run with an incomplete observed data set combined with data generated using the multiple linear regression technique at the locations of the remaining 10 raingauges (Fig. 8). The reason for this is that systematic differences of the rainfall fields are removed by the multiple linear regressions. The similar results are also observed in other subcatchments (not shown).

The performance for flood events was also investigated. Figure 9 depicts the mean absolute error and root mean squared error for the gauge at Rottweil (Neckar).

On average, the absolute error with respect to the annual maximum discharges for the gauge at Rottweil (Neckar)

ranges between 6.8% and 8.2%. The highest error was observed when the calibrated model using 20 raingauges was run using 10 raingauges. The error reduced to 6.9% when the calibrated model using 20 raingauges was run using 20 raingauges, however, with 10 raingauges of precipitation data estimated using the multiple linear regression technique and the remaining 10 from the observed data. This analysis also supports that the missing measurements can and should be supplemented using a data filling algorithm (in our case multiple linear regression) if additional precipitation information was used for model calibration.

Models are also used for water management purposes, thus their correct balances of different time periods are also of importance. A summary of the Nash-Sutcliffe coefficients at a 7 day and 30 day time step in the validation period is shown in Table 5. The Nash-Sutcliffe coefficients of the daily time step for the same gauges are shown in Fig. 8. Regarding modeling of runoff at higher time steps, the model performance in terms of the Nash-Sutcliffe coefficient shows a similar pattern as that shown in Fig. 8 at the daily time step. As shown that the model performance improves at the higher aggregation time steps for three of the five subcatchments and slightly deteriorates for the catchments Suessen (Fils) and Plochingen (Fils). The reason for this might be that the Fils originates from a karstic area where subsurface catchments might differ from the surface ones.
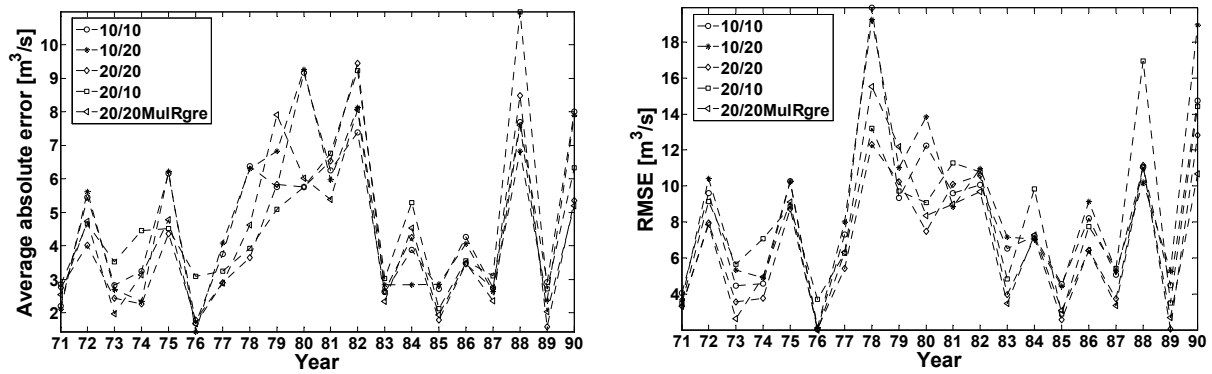
Fig. 9. Event statistics for each annual maximum flood event during the validation period using precipitation obtained from different raingauge networks and estimated precipitation for the gauge at Rottweil (Neckar): mean absolute error (left panel) and root mean squared error (right panel).

Table 5. Nash-Sutcliffe coefficients at 7 days and 30 days time step obtained using different level of precipitation input information for selected five gauges for the validation period.

Gauge	Number of raingauges	Nash-Sutcliffe coefficient	
		7 days time step	30 days time step
Rottweil (Neckar)	20/10	0.76	0.82
	20/20MulRgre	0.86	0.90
Horb (Neckar)	20/10	0.80	0.82
	20/20MulRgre	0.89	0.91
Suessen (Fils)	20/10	0.82	0.80
	20/20MulRgre	0.82	0.80
Plochingen (Fils)	20/10	0.85	0.83
	20/20MulRgre	0.85	0.82
Plochingen (Neckar)	20/10	0.88	0.90
	20/20MulRgre	0.90	0.92

20/10: calibrated with 20 raingauges and simulated with 10 raingauges and **20/20MulRgre** indicates model calibrated with 20 raingauges and simulated with 20 raingauges (rainfall estimated at 10 locations considered as missing measurements) (see text).

5 Conclusions

In this paper attempts have been made to investigate the influence of the spatial representation of the precipitation input, interpolated from different raingauge density, on the calibration and application of the semi-distributed HBV model. The precipitation input was interpolated using the external drift kriging method from the point measurements of the selected raingauge networks. The performance of the HBV model was assessed using different model performance evaluation criteria for the calibration and validation periods.

A number of simulation experiments were carried out in accordance to the study objective. A first set of experi-

ments considered the spatial representation of precipitation from varying raingauge networks. It showed that the number and spatial distribution of raingauges affect the simulation results. It was found that the overall model performances worsen radically with an excessive reduction of raingauges. However, the overall performances were not significantly improved by increasing the number of raingauges more than a certain threshold number specially if stations around but outside the catchments are considered. A significant inability to represent the spatial precipitation fields using network consisting of less number of raingauges are observed in the summer season particularly for the smaller subcatchments. These results are model and resolution specific. Higher spatial

model resolution can lead to a higher sensitivity on precipitation observation density. The influence of raingauge density in other regions might be very different depending on the rainfall type (convective or advective), seasonality of precipitation, importance of snow accumulation and melt, topography and land use. The more the hydrological processes are complicated the more precipitation observations might be necessary. Temporal variability is influencing the hydrograph considerably. However, this was not the main interest of this paper - we tried to concentrate on the spatial aspect. A combined space-time investigation would of course be of great interest, which, is beyond the scope of this paper.

A second set of analysis considered the model calibration using precipitation interpolated from one type of raingauge network and was run using precipitation interpolated from another type of raingauge network. The analysis indicated that models using different raingauge networks might need their parameters recalibrated. Specifically, the HBV calibrated with dense precipitation information fails when run with relatively sparse precipitation information. However, the HBV model calibrated with sparse precipitation information can perform well when run with dense precipitation information.

A third set of experiments analyzed the reliability of supplementing missing precipitation measurements used for the calibration with data estimated using a multiple linear regression technique, and running the model using that precipitation combined with observed precipitation. The results showed that the model performs well when it was calibrated with a complete set of observed precipitation and run with an incomplete observed data set combined with estimated data instead of running the calibrated model using incomplete observed data only. This result offers an encouraging perspective for the implementation of such a procedure for an operational flood forecasting system. Further research is needed in this direction to prove the practical applicability.

References

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

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

Anctil, F., Lauzon, N., Andreassian, V., Oudin, L., and Perrin, C.: Improvement of rainfall-runoff forecasts through mean areal rainfall optimization, J. Hydrol., 328, 717–725, 2006.

Bárdossy, A. and Plate, E.: Space-Time Model for Daily Rainfall using Atmospheric Circulation Patterns, Water Resour. Res., 28(5), 1247–1259, 1992.

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

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

Booij, M. J.: Modelling the effects of spatial and temporal resolution of rainfall and basin model on extreme river discharge, Hydrol. Sci. J., 47(2), 307–320, 2002.

Bormann, H.: Impact of spatial data resolution on simulated catchment water balances and model performance of the multi-scale TOPLATS model, Hydrol. Earth Syst. Sci., 10, 165–179, 2006, http://www.hydrol-earth-syst-sci.net/10/165/2006/.

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

Chaplot, V., Saleh, A., and Jaynes, D. B.: Effect of the accuracy of spatial rainfall information on the modeling of water, sediment, and NO_3-N loads at the watershed level, J. Hydrol., 312, 223–234, 2005.

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

Dong, X., Dohmen-Janssen, C. M., and Booij, M. J.: Appropriate Spatial Sampling of Rainfall for Flow Simulation, Hydrol. Sci. J., 50(2), 279–297, 2005.

Duncan, M. R., Austin, B., Fabry, F., and Austin, G. L.: The effect of gauge sampling density on the accuracy of streamflow prediction for rural catchments, J. Hydrol., 142, 445–476, 1993.

Hargreaves, G. H. and Samani Z. A.: Reference crop evapotranspiration from temperature, Appl. Engr. Agric., 1, 96–99, 1985.

Hundecha, Y. and Bárdossy, A.: 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, 281–295, 2004.

Kitanidis, P. K.: Introduction to Geostatistics: Applications in Hydrogeology, Cambridge University Press, 249 pp., 1997.

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

Krajewski, W. F., Ciach, G. J., and Habib, E.: An analysis of small-scale rainfall variability in different climatic regimes, Hydrol. Sci. J., 48, 151-162, 2003.

Lindström, G., Johansson, B., Persson, M., Gardelin, M., and Bergström, S.: Development and test of the distributed HBV-96

hydrological model, J. Hydrol., 201, 272–288, 1997.

Michaud, J. D. and Sorooshian, S.: Effect of rainfall-sampling errors on simulations of desert flash floods, Water Resour. Res., 30(10), 2765–2776, 1994.

Montgomery, D. C. and Peck, E. A.: Introduction to linear regression analysis, John Wiley and Sons, New York, New York, USA, 1982.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models. Part I. A discussion of principles, J. Hydrol., 10, 282-290, 1970.

Oudin, L., Perrin, C., Mathevet, T., Andreassian, V., and Michel, C.: Impact of biased and randomly corrupted inputs on the efficiency and the parameters of watershed models, J. Hydrol., 320, 62–83, 2006.

Peters-Lidard, C. D. and Wood, E. F.: Estimating storm areal average rainfall intensity in field experiments, Water Resour. Res., 30(7), 2119–2132, 1994.

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

Seed, A. W. and Austin, G. L.: Sampling Errors for Raingage Derived Mean Areal Daily and Monthly Rainfall, J. Hydrol., 118, 163–173, 1990.

Singh, V.: Effect of Spatial and Temporal Variability in Rainfall and Watershed Characteristics on Stream flow Hydrograph, Hydrol. Processes, 11, 1649–1669, 1997.

St-Hilarie, A., Ouarda, T. B. M. J., Lachance, M., Bobée, B., Gaudet, J., and Gibnac, C.: Assessment of the impact of meteorological network density on the estimation of basin precipitation and runoff: a case study, Hydrol. Processes, 17(18), 3561–3580, 2003.

Sun, H., Cornish, P. S., and Daniell, T. M.: Spatial Variability in Hydrologic Modeling using Rainfall-Runoff Model and Digital Elevation Model, J. Hydrol. Eng., 7(6), 404–412, 2002.

Syed, K. H., Goodrich, D. C., Myers, D. E., and Sorooshian, S.: Spatial characteristics of thunderstorm rainfall fields and their relation to runoff, J. Hydrol., 271, 1–21, 2003.

Wilson, C. B., Valdes, J. B., and Rodriguez-Iturbe, I.: On the influence of the spatial distribution of rainfall on storm runoff, Water Resour. Res., 15(2), 321–328, 1979.

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

www.hydrol-earth-syst-sci.net/12/77/2008/

Hydrol. Earth Syst. Sci., 12, 77–89, 2008