**Hydrology and
Earth System
Sciences**

Robust estimation of hydrological model parameters

A. Bárdossy and S. K. Singh

Institute of Hydraulic Engineering, University of Stuttgart, Stuttgart 70569, Germany

Abstract. The estimation of hydrological model parameters
is a challenging task. With increasing capacity of computa-
tional power several complex optimization algorithms have
emerged, but none of the algorithms gives a unique and *very
best* parameter vector. The parameters of fitted hydrological
models depend upon the input data. The quality of input data
cannot be assured as there may be measurement errors for
both input and state variables. In this study a methodology
has been developed to find a set of robust parameter vectors
for a hydrological model. To see the effect of observational
error on parameters, stochastically generated synthetic mea-
surement errors were applied to observed discharge and tem-
perature data. With this modified data, the model was cal-
ibrated and the effect of measurement errors on parameters
was analysed. It was found that the measurement errors have
a significant effect on the best performing parameter vector.
The erroneous data led to very different optimal parameter
vectors. To overcome this problem and to find a set of robust
parameter vectors, a geometrical approach based on Tukey's
half space depth was used. The depth of the set of N ran-
domly generated parameters was calculated with respect to
the set with the best model performance (Nash-Sutclife ef-
ficiency was used for this study) for each parameter vector.
Based on the depth of parameter vectors, one can find a set of
robust parameter vectors. The results show that the parame-
ters chosen according to the above criteria have low sensitiv-
ity and perform well when transfered to a different time pe-
riod. The method is demonstrated on the upper Neckar catch-
ment in Germany. The conceptual HBV model was used for
this study.

1 Introduction

Hydrological models are used for different purposes such
as water management or flood forecasting. The estimation
of hydrological model parameters is a difficult task. Rea-
sons for this are the highly non-linear nature of hydrological
processes and the fact that different parameter vectors driv-
ing models describing the physical processes might have the
same effect on the discharge. This means that changes of
some parameters might be compensated by others. Unfor-
tunately traditional manual calibration of models with rea-
sonable parameter values often leads to weak results. Hence,
nowadays automatic procedures based on numerical methods
are used.

Many different optimization routines have been developed
to find optimum parameter vectors. A variety of objective
functions measuring model performance including multi-
objective approaches, have been tried to define optimality in
this context. Non-linearity of the hydrological models and
of the objective functions lead to very complex optimization
problems. Beven and Freer (2001) argue that there are no
optimum parameters, in fact there is a large set of parameter
vectors which all perform reasonably and one cannot eas-
ily distinguish between them. They call this an *equifinality*
problem which leads to high uncertainties in the model pre-
dictions. Frequently shown *dotty plots* give the impression
that the set of good parameter vectors can be found anywhere
in the space. But no clear convergence to a best single value
can be observed. However, in a previous paper (Bárdossy,
2007), the geometrical properties of a parameter vector with
good performance (from now on the *set of good parameters*)
were investigated for a two-dimensional case. It was shown
that the set of good parameters is well structured. Unfor-
tunately in higher dimensional spaces one can not *see* these
sets, thus it is not clear whether they are scattered or have
some clear structure. The high scatter observed in the good
individual parameters is very disturbing since, it does not

enable a classical identification of a single vector within corresponding confidence bounds.

The GLUE procedure (Beven and Binley, 1992) has widely been applied for uncertainty assessment and discussed in the scientific literature, although alternative procedures using parametric approaches to obtain *best* solutions have also been suggested. These approaches are optimal under certain assumptions, however, they are often selected purely for mathematical convenience and not necessarily based on experience with data.

In Kavetski et al. (2006a,b) it was noted that the performance metric of hydrological models is a *bumpy* function of the model parameters. They suggest different numerical procedures to smoothen parameter surfaces and to obtain optimal parameter vectors.

The purpose of this paper is to investigate the reasons leading to very different near optimum parameter vectors and to investigate the properties of the set of good parameters in high dimensional spaces. Our goal is not to find the parameter vectors which perform best for the calibration period but to find parameter vectors which:

1. lead to good model performance over the selected time period

2. lead to a hydrologically reasonable representation of the corresponding processes

3. are not sensitive: small changes of the parameters should not lead to very different results

4. are transferable: they perform well for other time periods and might also perform well on other catchments (i.e. they can be regionalized)

Concepts of computer geometry and multivariate statistics are used to identify the set of good parameters. Specifically, convex sets and the depth function defined in Tukey (1975) are used.

The concept of data depth has recently received much attention by Donoho and Gasko (1992), Rousseeuw and Struyf (1998), Rousseeuw and Ruts (1998), Liu et al. (1999), Zuo and Serfling (2000), Miller et al. (2003) and Lin and Chen (2006). It has been used for the investigation of large data sets. The application of depth function has been seen in several fields. Serfling (2002) used the depth function for nonparametric multivariate analysis. Cheng et al. (2000) had used data depth function for monitoring multivariate aviation safety data for control chart. They were also applied in quality control by Liu (1995), Hamurkaroğlu et al. (2004). The only hydrological application found so far is in Chebana and Ouarda (2008), where data depth was used to define weights for the regional estimation of hydrological extremes.

This paper is organized as follows: After the introduction, the case study area is introduced. In Sect. 3, the effect of observation errors on the identification of hydrological model parameters is discussed. In Sect. 4, the notion of statistical

depths is introduced. Geometrical properties of the set of *good* parameters is investigated with the help of the depth function and robust parameter vectors are identified. In the final section, results are discussed and conclusions are drawn.

2 Case study area and the hydrological model

The concept of this paper will be illustrated with examples from the Neckar catchment. The hydrological model chosen is a modified version of the HBV model. A short description of the catchment and the model is provided in this section.

2.1 Study area

This study was carried out on the upper Neckar basin in South-West Germany in the state of Baden-Württemberg using data from the period 1961–1990. The region is flat, undulating in the east and north. The Black Forest and Swabian Alps are in the west and south. The $4000\,km^2$ large Upper Neckar basin was subdivided into 13 subcatchments (Fig. 1). Three of which were used for this study.

The study area elevations range from 238 m a.s.l. to 1010 m a.s.l. The dataset used in this study includes measurements of daily precipitation from 151 gauges and daily air temperature at 74 climatic stations. The meteorological input required for the hydrological model was interpolated from the observations with External Drift Kriging (Ahmed and de Marsily, 1987) using topographical elevation as external drift. The mean annual precipitation is 908 mm/year. Land use is mainly agricultural in the lowlands and forested in the medium elevation ranges. Hydrological characteristics of the three selected subcatchments are given in Table 1. For further details please refer to Samaniego (2003) and Bárdossy et al. (2005).

2.2 Hydrological model

The HBV model concept was developed by the Swedish Meteorological and Hydrological Institute (SMHI) in the early 1970's. It has been modified at the Institute of Hydraulic Engineering, University Stuttgart and used for this study. It includes conceptual routines for calculating snow accumulation and melts, soil moisture and runoff generation, runoff concentration within the subcatchment, and flood routing of the discharge in the river network. The snow routine uses the degree-day approach as set out in Eqs. (1) and (2). Soil moisture is calculated by balancing precipitation and evapotranspiration using field capacity and permanent wilting point as parameters, Eqs. (3) to (5).

$$\text{MELT} = DD \cdot (T - T_{\text{crit}}) \tag{1}$$

$$DD = DD_0 + k \cdot P \tag{2}$$

$$P_{\text{eff}} = (SM/FC)^{\beta} \cdot P + \text{MELT} \tag{3}$$
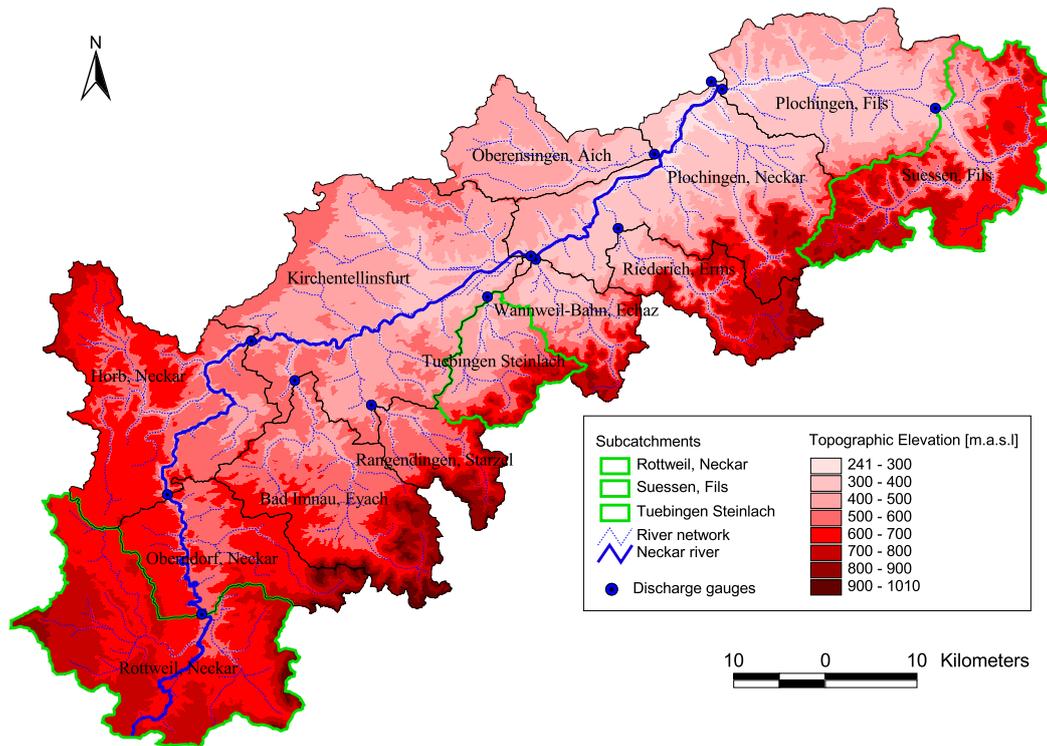
Fig. 1. Study area: upper Neckar catchment in south-west Germany.

Table 1. Summary of the size of the different subcatchment in the study area.

	Subcatchment	Subcatchment size (km^2)	Elevation (m)	Slope (degree)	Mean Discharge (m^3/s)	Annual Precipitation (mm)
1	Rottweil (Neckar)	454.65	555–1010	0–34.2	5.1	968.16
2	Tübingen (Steinlach)	140.21	340–880	0–38.8	1.7	849.84
3	Süssen (Fils)	345.74	360–860	0–49.3	5.9	1003.45

Where: P_{eff} is the effective precipitation, SM is the actual soil-moisture, FC is the maximum soil storage capacity, β is a model parameter (shape coefficient), P is the depth of daily precipitation, MELT is the amount of snow melt, DD is degree day factor, T is the mean daily air temperature, T_{crit} is threshold temperature, DD_0 is degree day factor when there is no rainfall and k is a positive number.

$$PE_a = (1 + C \cdot (T - T_m)) \cdot PE_m \tag{4}$$

Where: PE_a is the adjusted potential evapotranspiration, C is a model parameter, T is the mean daily air temperature, T_m is the long term mean monthly air temperature and PE_m is the long term mean monthly potential evapotranspiration.

$$E_a = (SM/PWP) \cdot PE_a \tag{5}$$

Where: E_a is the actual evapotranspiration, SM is the actual soil-moisture and PWP is limiting soil-moisture at which potential evapotranspiration take place. Runoff generation is simulated by a nonlinear function of the actual soil-moisture and precipitation. The runoff concentration is modeled by two parallel nonlinear reservoirs representing the direct discharge and the groundwater response. Flood routing between the river network nodes uses the Muskingum method. Additional information about the HBV model in general can be found in Hundecha and Bárdossy (2004), and Bergström (1995). Direct runoff and percolation from each subcatchment are calculated using Eqs. (6) to (10).

$$Q_0 = k_0 \cdot (S_1 - L) \tag{6}$$

$$Q_1 = k_1 \cdot S_1 \tag{7}$$
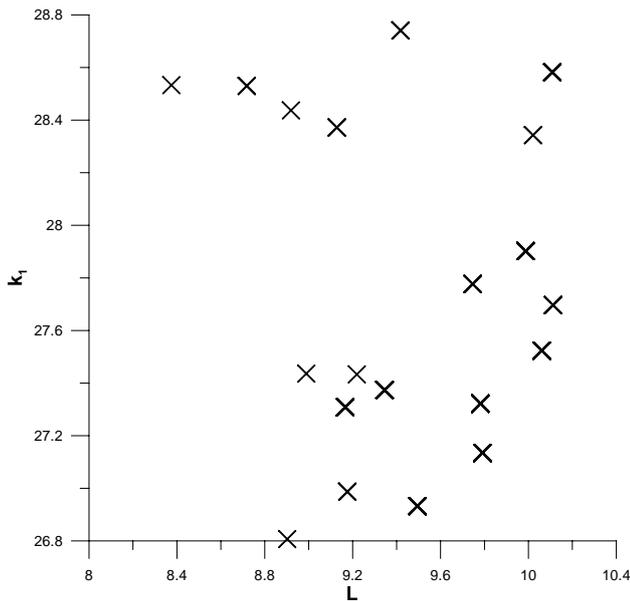
Fig. 2. Scatter plot of model parameters obtained by optimization using random discharge errors.

$$Q_{\text{perc}} = k_{\text{perc}} \cdot S_1 \tag{8}$$

$$Q_2 = k_2 \cdot S_2 \tag{9}$$

Where: Q_0 is near surface flow, Q_1 is interflow, Q_{perc} is percolation, Q_2 is baseflow, k_0 is the near surface flow storage constant, k_1 is the interflow storage constant, k_{perc} is the percolation storage constant, k_2 is the baseflow storage constant, S_1 is upper reservoir water level, S_2 is lower reservoir water level, L is threshold water level for near surface flow. The total runoff is computed as the sum of the outflows from the upper and lower reservoirs. The total flow is then smoothed using a transformation function, consisting of a triangular weighing function with one free parameter, MAXBAS.

$$Q = g(t, \text{MAXBAS}) \cdot (Q_0 + Q_1 + Q_2) \tag{10}$$

Where: Q is current overall discharge and MAXBAS is the duration of the triangular weighting function (Unit Hydrograph). There are 15 parameters to describe the model, out of which 9 parameters are used for this study (Table 2).

3 The effect of observation errors

In rainfall runoff modeling, input errors play a crucial role but the problem of input errors is generally neglected by hydrologists (Paturel et al., 1995). Hydrological models use observation data for the identification of model parameters. Unfortunately many of the hydrological observations contain partly systematic and partly random errors. Precipitation is measured at a few selected locations and typically interpolated for the catchment area. Thus, precipitation values used

in the model can be wrong due to measurement (for example caused by evaporation or wind) and to interpolation errors. The impact due to error in precipitation has been investigated by Ibbitt (1972), Troutman (1985), Paturel et al. (1995), Andréassian et al. (2001) and Oudin et al. (2006), who found that error in precipitation has significant influence on model performance.

Observed discharge is used as the main calibration quantity, thus their errors may have significant influence on model performance. In most cases water levels are observed and rating curves are used to transform them to discharges. This is an important source of partly systematic error and, when combined with other errors, compromise the identification of the model parameters.

The parameter vectors obtained by model parameter optimization algorithms are optimal with respect to an erratic objective function. Measurement errors and errors due to model structure are mixed (Todini, 2007) and cannot be separated directly. The following examples illustrate the effect of observation uncertainty on parameter estimation.

Firstly, consider the observed meteorological variables and discharge. We assume that due to measurement errors the accuracy of the measured discharge $Q_M(t)$ is $q\%$. Thus, the real but unknown discharge $Q_E(t)$ can be written as:

$$Q_E(t) = Q_M(t)(1 + \varepsilon_Q(t)) \tag{11}$$

with $\varepsilon_Q(t)$ being a random error. This random error is due to uncertainties of the rating curve, non-uniqueness of the stage discharge relationship, changes of the cross section etc. Here we assume that the error follows a normal distribution $N(0, \frac{q}{100})$. This means we assume a constant relative random error and further, that the errors are independent (error dependence would increase the effect of observation uncertainty).

To quantify the effect of the flow error on model performance a set of $M=100$ realizations of $Q_E(t)$ was generated with $q=5$. Note that the rating curve related errors are usually higher than this, especially in the case of extreme flows. Consequently the parameters of the hydrological model were estimated using simulated annealing by maximizing the Nash-Sutcliffe coefficient as if each parameter vector Q_E was the observed series. The model parameters obtained show a considerable scatter. For example in two dimensions, Fig. 2 shows the scatter plot for the selected model parameters L and k_1, where $M=20$.

The uncertainty of the estimated model parameters with respect to input error structure can also be investigated. With respect to temperature observations one can assume that the real but unknown temperature $T_E(t)$ can be written as:

$$T_E(t) = T_I(t) + \varepsilon_T(t) \tag{12}$$

In this case an additive error of the catchment mean temperature was assumed. As in the previous case, M realizations were generated and model parameters were optimized for each of the series separately.

Hydrol. Earth Syst. Sci., 12, 1273–1283, 2008

www.hydrol-earth-syst-sci.net/12/1273/2008/

Table 2. Model parameters range for Rottweil (Neckar) Catchment.

Parameter	Description	Iteration 1		Iteration 4	
		Min	Max	Min	Max
T_{crit}	Threshold temperature for snow melt initiation	−1.50	2.50	−0.16	0.42
DD	Degree-day factor	0.12	2.12	1.30	2.11
Dew	Precipitation/Degree-day relation	0.01	1.09	0.287	1.06
β	Model parameter (shape coefficient)	0.01	2.01	0.81	1.08
L	Threshold water level for near surface flow	8.30	10.29	8.419	10.28
K_0	Near surface flow storage constant	0.77	2.77	1.75	2.68
K_1	Interflow storage constant	26.80	28.81	26.82	28.62
K_{per}	Percolation storage constant	19.98	21.98	20.05	21.94
K_2	Baseflow storage constant	36.80	38.79	36.83	38.72

Table 3. Model performance for the observed series using optimal parameters obtained using 100 randomly perturbed discharge data sequences.

Catchments	Mean NS	Median NS	Max NS	Min NS	Standard deviation
Rottweil	0.699	0.698	0.733	0.675	0.0123
Tübingen	0.716	0.716	0.733	0.703	0.0054
Süssen	0.751	0.751	0.775	0.733	0.0083

Tables 3 and 4 show the effect of observation error on the Nash-Sutcliffe coefficient for both discharge and temperature measurement errors, respectively. The uncertainty of model parameters with respect to precipitation uncertainty can be considerable, depending on the density of the observation network. This problem was investigated by Das (2006).

These examples show that model parameters and model performance are highly influenced by measurement errors. Two parameter vectors, with model performances differing in the range of the measurement error caused fluctuations of the Nash-Sutcliffe value, cannot be distinguished from each other. Either of them might lead to a better description of the hydrological system. The parameters obtained by sophisticated optimization procedures might thus be suboptimal in reality. Thus, it is reasonable to investigate the set of parameters which gives similar performance as the numerical optimum. These parameters will be called *good parameters* in the subsequent sections.

4 Geometrical structure of the parameter set

One of the major problems is that there is a large number of parameter vectors which perform nearly equally well. It is difficult, then to decide which of these should be taken for prediction. Scatter plots showing model performances as a function of individual parameters indicate that a wide range of parameter values can lead to good model performance. At present it seems impossible to know a priori if a fitted given parameter vector leads to good or bad performance when ap-

plied to a model. In Bárdossy (2007), the geometrical structure of the best performing parameters of the unit hydrograph (impulse response function) of the Nash cascade were investigated. It was shown that the set has a very clear geometrical structure. In this paper models with many more than two parameters are considered. It is difficult to visualise the subset of best parameters in higher dimensions; instead methods of computational geometry are used herein.

In order to investigate the properties of the set of good parameter vectors, the concept of data depth was used. Depth functions were first introduced by Tukey (1975) to identify the center (a kind of generalized median) of a multivariate dataset. Several generalizations of this concept have been defined in Rousseeuw and Struyf (1998), Liu et al. (1999) and Zuo and Serfling (2000).

Definition: The halfspace depth of a point p with respect to the finite set X in the d dimensional space $\Re^d$ is defined as the minimum number of points of the set X lying on one side of a hyperplane through the point p. The minimum is calculated over all possible hyperplanes.

Formally the halfspace depth of the point p with respect to set X is:

$$D_X(p) = \min_{n_h} \left(\min \left(|\{x \in X \langle n_h, x - p \rangle > 0\}| \right), \right.$$
$$\left. \left(|\{x \in X \langle n_h, x - p \rangle < 0\}| \right) \right) \tag{13}$$

Here $\langle x, y \rangle$ is the scalar product of the d dimensional vectors, and n_h is an arbitrary unit vector in the d dimensional space representing the normal vector of a selected hyperplane.

Table 4. Model performance for the observed series using optimal parameters obtained using 100 randomly perturbed temperature data sequences.

Catchments	Mean NS	Median NS	Max NS	Min NS	Standard deviation
Rottweil	0.690	0.691	0.723	0.650	0.014
Tübingen	0.722	0.723	0.740	0.706	0.0062
Süssen	0.750	0.750	0.777	0.719	0.0121
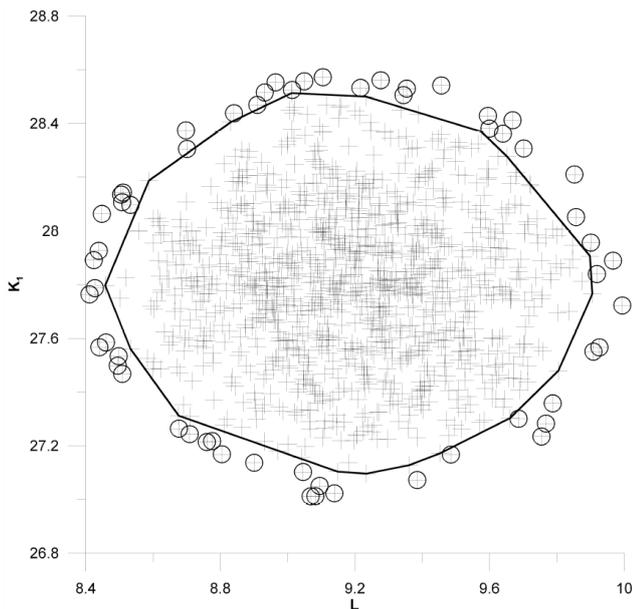
Fig. 3. Points with low depth ≤ 4 (circles) of a two dimensional set of model parameters (crosses).

Fig. 4. The performance of the model using different depth.

If the point p is outside the convex hull of X then its depth is 0. Points on and near the boundary have low depth while points *deeply* inside have high depth.

One advantage of this depth function is that it is invariant to affine transformations of the space. This means that the different ranges of the parameters have no influence on their depth.

The calculation of the halfspace depth is computationally very expensive if the number of points in X is large or the dimension is high. Efficient algorithms are available for $d=2$ from Miller et al. (2003). In this study the approximate calculation suggested in Rousseeuw and Struyf (1998) was used.

To illustrate the concept of data depth let us consider a two dimensional data set. The parameter vectors with good performance are shown on Fig. 3. We have chosen L and k_1 in this example. The first parameter corresponds to the x axis and the second to the y axis. Points with low depth ≤ 4 are marked with circles. The convex polygon separates the points with low depth (≤ 4) from those with high depth (>5). Note that points near the boundary of the set have a

low depth while the interior points in the middle of the set have the greatest depth.

4.1 Data depth of the good parameter set

In order to explore the set of reasonably performing parameter vectors using the above introduced concept, nine parameters of the HBV model were considered. The parameter ranges used for the initial Monte Carlo simulation for the subcathment Rottweil (Neckar) is given in Table 2. N random parameter vectors where generated in a rectangle bounded by reasonable limits in the $d=9$ dimensional space. For each of these parameter vectors the hydrological model was applied and the performance was calculated. This set of parameter vectors is denoted as X_N. A subset $X_N^* \subset X_N$ of the best performing parameter vectors (in our case we chose the upper 10%) were identified. The depth of each point in X_N with respect to X_N^* was calculated. Figure 4 shows the histogram of the performance of the hydrological model for the points $\theta \in X_N$ with depth $D(\theta) > L$. One can see that all points with high depth (being in the geometrical interior of the set X_N) lead to good model performance. The reason for this is that one assumes that the low depth points can be regarded as an iso-hypersurface corresponding to the selected level. If one assumes continuity of the objective function then higher values of the function are expected in the interior of the set.

Table 5. Model performance for the $N=10000$ random parameter sets with respect to the data depth calculated on the basis of the points selected corresponding to the upper 10% performance.

Depth	Number of points	Mean NS	Standard deviation
–	10000	0.3132	0.6766
≥ 1	1743	0.6720	0.0198
≥ 10	893	0.6839	0.0135
≥ 50	182	0.6931	0.0090
> 100	33	0.6971	0.0069

In order to check this statement an independent second set Y_N of N random parameter vectors were generated. The depth of the points of Y_N with respect to X_N^* was calculated. For all parameters $\theta \in Y_N$, the hydrological model was run and the performances calculated. The results are evaluated for parameters such that $D(\theta) \geq L$, exemplified in Table 5 with the statistics of the performances. One can see that the randomly generated parameter vectors which posses high depth have good model performance. The standard deviation of the performance decreases with increasing depth, showing that in the *deep* interior of the set all parameter vectors perform similarly. These results show that for this case one can geometrically identify parameter vectors which are good. Note that even if the best performance is related to the deepest subset, this is not necessarily always the case, since the global optimum might itself correspond to a low depth.

4.2 Transferability

In order to investigate the transferability of the parameters with respect to their depth, two experiments were carried out.

As a first test the total observation period of 30 years was divided into three 10 year periods. The hydrologic characteristics of the three time periods are listed in Table 6. The model performance was calculated for each time period. The set of good parameter vectors was identified for each time period separately and the depth of each parameter with respect to this set was calculated. In this way, three depth values were assigned to each parameter vector.

The sets with 50 and 150 deepest parameter vectors were identified for each time period. The intersection of the convex sets corresponding to the 50 deepest points consisted of 36 for the 150 point set 84 points indicating that depth is stable over all time periods. Note that a parameter vector was considered to be in the intersection if it had positive depth with respect the sets considered. As a set of 10 000 points were considered; an independence selection of two sets with 150 points would have led with high probability no points in the intersection. This means that parameters with large depth are robust with respect to the selected time period.
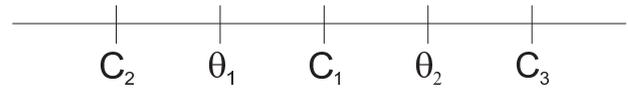
Fig. 5. Construction of the points C_1, C_2 and C_3 in one dimension for sensitivity analysis of parameters.

As a second test the parameters with greater depth for one time period were used for another time period and their performance was calculated. In Table 7 the results of the transferred model quality with respect to depth corresponding to the time period 1961–1970 are shown. Note that the subset of the boundary points was selected by choosing only points for which the performance exceeds a given threshold. This way we obtained two sets with the same mean performance. Note that for the interior points, the performance in the other time periods is significantly better than those of the boundary points. The standard deviations of the performance for the validation time periods are smaller for the interior points which indicates that the transfer of these parameters is more reasonable for the parameter vectors from the interior.

4.3 Sensitivity

The sensitivity is not investigated in the usual way to see how the model reacts to changes of individual parameters. Instead the parameter vectors are considered as sensitive if a small change of the whole vector might lead to a big change (usually drop) in the performance of the model. The sensitivity of boundary points and inside points was compared. For this purpose parameter vectors were altered from the boundary ($D(\theta_1=1)$) and from the inside ($D(\theta_2)>1$) of the set. A specific vector η was added and subtracted from the selected parameter vectors. This way the vectors θ_1 and θ_2 were altered to the same extent. Four new parameter vectors

$$C_1 = \theta_1 - \eta$$
$$C_2 = \theta_1 + \eta$$
$$C_3 = \theta_2 - \eta$$
$$C_4 = \theta_2 - \eta$$

are created. We select $\eta = \frac{\theta_1 - \theta_2}{2}$ thus $C_2 = C_4$. Due to the definition of the depth $D(C_1) \geq 1$ while $D(C_2) \leq 1$. For C_3 one cannot make any statements on the depth.

Figure 5 explains the construction of the three points in one dimension.

The above construction of parameter vectors C_1, C_2 and C_3 was carried out for a large number of randomly selected pairs θ_1 and θ_2. The θ_1 and θ_2 were selected in such a manner that their mean performance was the same. Table 8 shows the statistics of the Nash-Sutcliffe coefficients for the sets corresponding to C_1, C_2 and C_3. One can see that the inside points all have good performance and the standard deviation

Table 6. Runoff characteristics for different time periods.

Subcatchment	Rottweil (Neckar)		Tübingen (Steinlach)		Süssen (Fils)	
Time period	Annual Precipitation (mm)	Annual Discharge (mm)	Annual Precipitation (mm)	Annual Discharge (mm)	Annual Precipitation (mm)	Annual Discharge (mm)
1961–1970	997.53	375.26	851.84	400.36	1007.94	575.55
1971–1980	908.48	309.36	808.14	366.62	960.02	512.62
1981–1990	997.21	385.66	888.84	404.86	1041.72	541.81

Table 7. Model performance for parameter vectors according to their depth corresponding to the time period 1961–1970.

Time period	Boundary points				Points with depth>5			
	Mean	Std	Min	Max	Mean	Std	Min	Max
1961–1970	0.682	0.010	0.667	0.711	0.682	0.015	0.647	0.705
1971–1980	0.630	0.043	0.488	0.714	0.673	0.019	0.634	0.726
1981–1990	0.751	0.029	0.641	0.798	0.776	0.017	0.715	0.804

Table 8. Model performance for the inner and the shifted boundary and deep points.

Variable	Mean NS	Standard deviation	Skewness	Max NS	Min NS
C_1	0.692	0.005	0.30	0.710	0.677
C_2	0.576	0.101	–6.95	0.658	–0.491
C_3	0.686	0.024	–5.58	0.713	0.363

is small. Points at C_2 (outside points) have the worst performance while C_3 is better than C_2 but worse than C_1. The skewness of the performance is nearly zero for the inside set C_3, while in other cases the strong negative skew indicates that in some cases the performance loss due to the shift outside the set is extremely high. The same alteration of the parameters leads to less performance loss for deep points than for shallow points. Further, there is no loss if the parameter vector remains in the convex set of deep parameters. This again highlights the advantage of deep parameter vectors.

5 Robust parameter estimation (ROPE)

In Sect. 4 it was shown that parameters in the interior (expressed through data depth) of the set of good points are themselves good and transferable and not very sensitive. A possible explanation for this is that these parameters can be regarded as a kind of compromise solution – where none of the processes represented by the parameters is overemphasized.

For modelling purposes one might be interested in finding the set of good parameters and also the identification of the deep parameter vectors for robust modelling. For this purpose the following procedure is suggested:

1. the limits for the d selected parameters are identified

2. N random parameter vectors forming the set X_N are generated in the d dimensional rectangle bounded by the limits defined in 1.

3. the hydrological model is run for each parameter vector in X_N and the corresponding model performances are calculated

4. the subset X_N^* of the best performing parameters is identified. This might be for example the best 10% of X_N.

5. M random parameter sets forming the set Y_M are generated, such that for each parameter vector $\theta \in_M$, $D(\theta) \geq L$ (with $L \geq 1$) where the depth is calculated with respect to the set X_N^*.

6. the set Y_M is relabeled as X_N and steps 3–6 are repeated until the performance corresponding to X_N and Y_M does not differ more than what one would expect from the observation errors.

Note that the ROPE algorithm can be easily modified to a general multivariate optimization procedures.

The algorithm was used for three selected subcatchments (Rottweil, Tübingen, Süssen). Four iterations were enough to find a *good* set of parameters for all the catchments. The performance of the model using different calibration and validation periods is summarized in Fig. 6 for catchment Süssen.
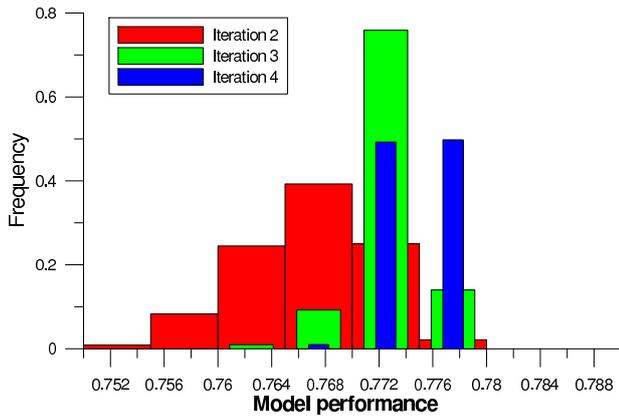
Fig. 6. Histograms of the model performances for the different iterations of the algorithm for catchment Süssen.

Fig. 7. Parameter value vs. model performance for the sets obtained in iteration 2 (crosses) and iteration 4 (circles) for catchment Süssen.

As one can see the subsequent iterations of the algorithm deliver better sets. The mean NS for the iterations increased form 0.767 (iteration 2) through 0.773 (iteration 3) to 0.775 (iteration 4). The improvement in the last iteration is very small and less than what one would expect to be caused by measurement errors. Therefore the algorithm stopped after this iteration.

Figure 7 shows the dotty plots of the selected model parameters for catchment Süssen. The performance corresponding to the parameters of iteration 4 are better than those corresponding to iteration 2 but the parameter range remains the same. Figure 8 shows the two dimensional scatter plot for the two model parameters for iterations 2 and 4 obtained for catchment Tübingen. One has the impression that these parameters can take a wide range of values, and that there is no difference between the the two sets. The ranges of the parameters for the catchment Rottweil for iteration 1 and 4 are listed in Table 2. Even if the ranges are very similar for many parameters one has to bear in mind that these are two dimensional projections of 9 dimensional sets. The sets themselves are very different, the ratio of their 9 dimensional volume is approximately 0.01 (calculated as Monte Carlo integral).

Figure 9 shows the sensitivity of the calculated discharge with respect to two sets of parameter vectors with different depth. 1000 Parameter vectors from the boundary (depth=1) and from the interior (depth>1) were taken and the corresponding hydrographs were calculated. The 95% and the 5% lines show that the interior parameter vectors lead to smaller differences in calculated discharge. The differences between the 95% and the 5% values are plotted separately on Fig. 10 showing that taking interior parameter vectors leads to an approximately 20% reduction.

In the case study presented here all parameters in the final set Y_M performed well. In the case of other models or other performance measures this may not necessarily be the case.
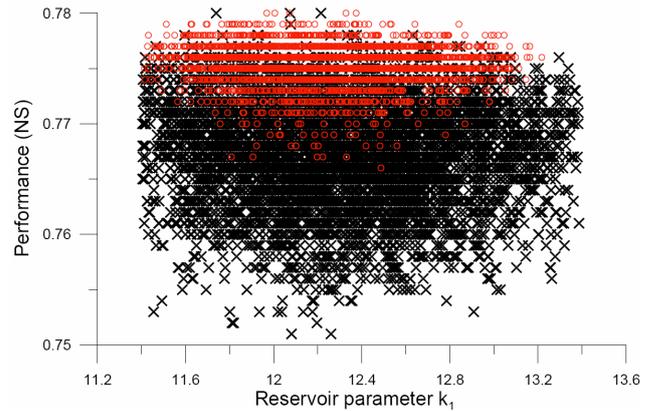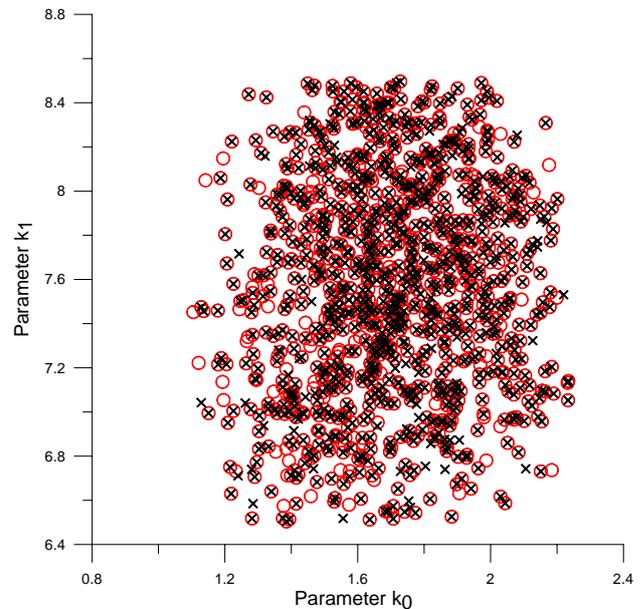
Fig. 8. Parameter value for the sets obtained in iteration 2 (crosses) and iteration 4 (circles) for catchment Tübingen.

However the set Y_M always contains a large portion of good parameters and possible transformations (for example taking the logarithm of some parameters) might fix this problem.

6 Conclusions

Discussion and conclusions

– In this paper the effect of observation uncertainty on the parameter estimation was investigated. It could be shown that observation errors can lead to very different optimal model parameters if the uniqueness of the
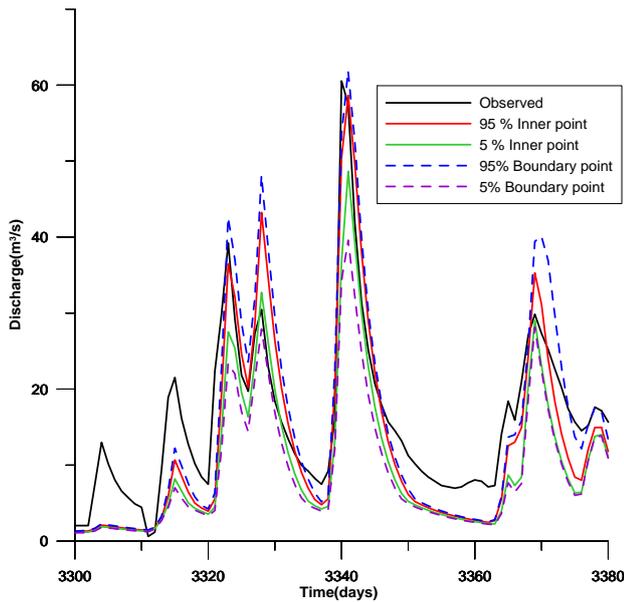
Fig. 9. Hydrograph with confidence interval for boundary points and inner points.

Fig. 10. Confidence band width of high depth vs confidence band width of low depth.

parameters is assumed and the parameters corresponding to the optimum of the performance function are identified.

– Observational uncertainty of the input and the discharge leads to variability of the model performance. This variability has to be considered in model parameter estimation. All model parameters which do not differ more in their performance than what can be caused by measurement errors could themselves be the best parameters.

– Data depth is a useful tool to identify robust parameter vectors. Parameters with low data depth are near the boundary and are sensitive to small changes and do transfer to other time periods less well as high depth ones.

– From the examples discussed in this paper, one could see that equally performing parameters are not necessarily equally transferable or equally sensitive. Data depth can help to find domains with robust and transferable parameters.

– A iterative algorithm to find a convex set containing good model parameters was developed.

– In this paper, model performance was measured by the traditional Nash-Sutcliffe coefficient. Other measures can be treated similarly - but might lead to different parameter sets.

Further research is needed to use the concepts developed in this paper for other purposes and models. Robust estimation of the model parameters might be very useful for regionalization and could contribute to a better prediction in ungaged basins. The suggested methodology can be extended for uncertainty analysis by relating the likelihood of the parameters to their depth, however further research is required to complete this task.

Edited by: F. Laio

References

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

Andréassian, V., Perrin, C., Michel, C., Usart-Sanchez, I., and Lavabre, J.: Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models, J. Hydrol., 250, 206–223, 2001.

Bárdossy, A.: Calibration of hydrological model parameters for un-gauged catchments, Hydrol. Earth Syst. Sci., 11, 703–710, 2007, http://www.hydrol-earth-syst-sci.net/11/703/2007/.

Bárdossy, A., Pegram, G. S., and Samaniego, L.: Modeling data relationships with a local variance reducing technique: Applications in hydrology, Water Resour. Res., 41, W08404, doi:10.1029/2004WR003851, 2005.

Bergström, S.: The HBV model, in: Computer Models of Watershed Hydrology, Water Resources Publications, Littleton, Colorado, USA, 1995.

Beven, K. and Binley, A. M.: The future of distributed models: model calibration and uncertainty prediction, Hydrol. Process., 6, 279–298, 1992.

Beven, K. and Freer, J.: Equifinality, data assimilation, and data uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, J. Hydrol., 249, 11–29, 2001.

Chebana, F. and Ouarda, T. B. M. J.: Depth and homogeneity in regional flood frequency analysis, Water Resour. Res., 44, W11422, doi:10.1029/2007WR006771, 2008.

Cheng, A. Y., Liu, R. Y., and Luxhéj, J. T.: Monitoring multivariate aviation safety data by data depth: control charts and threshold systems, IIE Transactions, 32, 861–872, 2000.

Das, T.: The Impact of Spatial Variability of Precipitation on the Predictive Uncertainty of Hydrological Models, Ph.D. dissertation No. 154, University of Stuttgart, 2006.

Donoho, D. L. and Gasko, M.: Breakdown properties of location estimates based on halfspace depth and projected outlyingness, The Annals of Statistics, 20(4), 1808–1827, 1992.

Hamurkaroğlu, C., Mert, M., and Saykan, Y.: Nonparametric control charts based on mahalanobis depth, Journal of Mathematics and Statistics, 33, 57–67, 2004.

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.

Ibbitt, R. P.: Effects of random data errors on the parameter values for a conceptual model, Water Resour. Res., 8(1), 70–78, 1972.

Kavetski, D., Kuczera, G., and Franks, S.: Calibration of conceptual hydrological models revisited: 2. Improving optimisation and analysis, J. Hydrol., 320, 187–201, 2006a.

Kavetski, D., Kuczera, G., and Franks, S. W.: Calibration of conceptual hydrological models revisited: 1. Overcoming numerical artefacts, J. Hydrol., 320, 173–186, 2006b.

Lin, L. and Chen, M.: Robust estimating equation based on statistical depth, Stat. Pap., 47, 263–278, 2006.

Liu, R. Y.: Control charts for multivariate processes, American Statistical Association, 90, 1380–1387, 1995.

Liu, R. Y., Parelius, J. M., and Singh, K.: Multivariate analysis by data depth: descriptive statistics, graphics and inference, The Annals of Statistics, 27(3), 783–858, 1999.

Miller, K., Ramaswami, S., Rousseeuw, P., Sellares, T., Souvaine, D., Streinu, I., and Struyf, A.: Efficient computation of depth contours by methods of computational geometry, Statistics and Computing, 13, 153–162, 2003.

Oudin, L., Perrin, C., Mathevet, T., Andréassian, 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.

Paturel, J., Servat, E., and Vassiliadis, A.: Sensitivity of conceptual rainfall-runoff algorithms to errors in input data case of the GR2M model, J. Hydrol., 168, 111–125, 1995.

Rousseeuw, P. J. and Ruts, I.: Constructing the bivariate tukey median, Statistica Sinica, 8, 827–839, 1998.

Rousseeuw, P. J. and Struyf, A.: Computing location depth and regression depth in higher dimensions, Statistics and Computing, 8, 193–203, 1998.

Samaniego, L.: Hydrological Consequences of Land Use/ Land Cover Change in Mesoscale Catchments, Transactions of the institute of hydraulic engineering, University of Stuttgart, Faculty of Civil Engineering, Stuttgart, Ph.D. dissertation No. 118, 2003.

Serfling, R.: Generalized Quantile Processes Based on Multivariate Depth Functions, with Applications in Nonparametric Multivariate Analysis, J. Multivariate Anal., 83, 232–247, 2002.

Todini, E.: Hydrological catchment modelling: past,present and future, Hydrol. Earth Syst. Sci., 11, 468–482, 2007, http://www.hydrol-earth-syst-sci.net/11/468/2007/.

Troutman, B. M.: Errors and parameter estimation in precipitation-runoff modeling, 2. Case study., Water Resour. Res., 21, 1214–1222, 1985.

Tukey, J.: Mathematics and picturing data, in: In Proceedings of the 1975 International 17 Congress of Mathematics, 2, 523–531, 1975.

Zuo, Y. and Serfling, R.: General notions of statistical depth function., The Annals of Statistics, 28(2), 461–482, 2000.

www.hydrol-earth-syst-sci.net/12/1273/2008/

Hydrol. Earth Syst. Sci., 12, 1273–1283, 2008