Pergamon

*Phys. Chem .Earth (B)*, Vol. 25, No. 1, pp. 59–64, 2000
© 2000 Elsevier Science Ltd
All rights reserved
1464-1909/00/$ - see front matter

PII: S1464-1909(99)00121-5

# Multiscale Calibration and Validation of a Conceptual Rainfall-Runoff Model

**J. Seibert[1], S. Uhlenbrook[2], C. Leibundgut[2] and S. Halldin[1]**

[1]Uppsala University, Department of Earth Sciences, Hydrology, Villavägen 16, SE-75236 Uppsala, Sweden
[2]University Freiburg, Institute of Hydrology, Fahnenbergplatz, DE-79098 Freiburg, Germany

*Received 3 June 1998; revised 23 November 1998; accepted 27 January 1999*

**Abstract.** Model calibration and validation is usually limited to comparing streamflows at the basin outlet. In this study observed runoff series from nested basins were used for calibration and validation on different spatial scales. A conceptual rainfall-runoff model was applied to nested basins of different sizes (15.2, 40 and 257 km$^2$) located in the Black Forest in south-west Germany. The first step was to calibrate the model individually for each of the three basins and to use the runoff series from the other two basins for validation. Optimised parameter values were related to sizes and other properties of the basins. In the second step, the model was calibrated simultaneously to the runoff series from all three catchments using a genetic algorithm and a fuzzy combination of the individual objective function values. It was not possible to obtain as good fits as those achieved by separately calibrating the model to each sub-basin. However, the fit between measured and observed streamflow for the individual basins were acceptable (model efficiency values around 0.8) and significantly better than those obtained using a parameter set optimised in just one of the other basins.

## 1 Introduction

Calibration and validation of conceptual rainfall-runoff models is usually limited to comparing streamflow at the basin outlet. However, if runoff series from nested catchments are available multiscale validation and calibration become possible. A parameter set optimised for one catchment can be validated against runoff from another catchment or different runoff series can be utilised in the calibration procedure. By this means a parameter set can be determined which is not only valid for the runoff at the outlet, but also for flow at points within the catchment.

A difficulty in the application to catchments of different size might be that parameter values in a lumped or semi-distributed conceptual rainfall-runoff model are effective parameters at the catchment scale. They could be expected to depend on the spatial scale because runoff characteristics vary with basin size (*e.g.*, Mimikou, 1984) and the importance of different processes for the runoff formation changes with scale (Beven, 1991). If, on the other hand, a parameter set can be found that is valid for runoff simulations from different (sub)basins, it will be useful both for regional runoff prediction in ungauged catchments within the region and for comparison of parameter values between regions.

In this study a conceptual rainfall-runoff model, a modification of the HBV model (Bergstöm, 1976; 1992), was applied to the Dreisam catchment and two nested basins, located in the southern Black Forest in south-west Germany. The following questions were addressed: (1) Are parameters scale-dependent in this conceptual rainfall-runoff model? (2) Is a parameter set calibrated in one catchment valid in a similar catchment of different size (*i.e.*, multiscale validation)? (3) How can a model be calibrated simultaneously to runoff series from different (sub-)catchments?

## 2 Materials and methods

### 2.1 The HBV model

The HBV model is a conceptual model that simulates daily discharge using daily rainfall and temperature, and monthly estimates of potential evaporation as input. The model consists of different routines (Fig.1), where snowmelt is computed by a degree-day method, groundwater recharge and actual evaporation are functions of actual water storage in a soil box, runoff formation is represented by three linear reservoir equations and channel routing is simulated by a triangular weighting function.

---

*Correspondence to*: Jan Seibert (jan.seibert@hyd.uu.se)

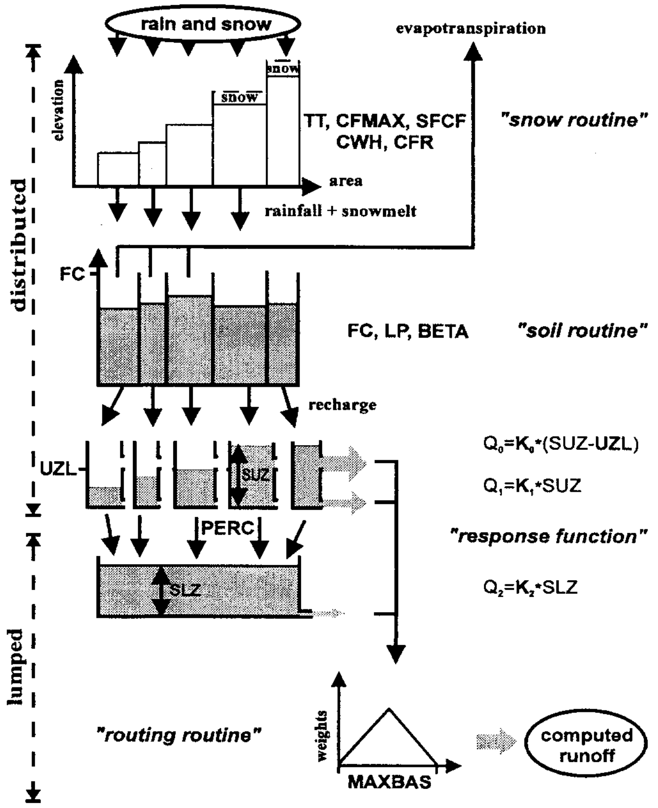

**Fig. 1.** Structure of the HBV model (for parameter descriptions see Tab. 2)

Descriptions of the model can be found in, *e.g.*, Bergström (1992), Harlin and Kung (1992) and Seibert (1997a). The version of the model used in this study, 'HBV light 1.2' (Seibert, 1997b) corresponds to the original version described by Bergström (1992) with three changes. Instead of prescribing initial states the new version uses a warming-up period. The restriction that only integer values are allowed for the routing parameter MAXBAS has been removed. In the original version of the HBV model (Bergström, 1992) computations in both the snow and the soil routine are performed individually for each elevation zone before the groundwater recharge is lumped together in the response routine. In the model version used in this study, the upper box in the response function is treated individually for each elevation zone additionally to the separate computations in the snow- and soil routines. This version was considered more logical than the original assumptions especially for mountainous catchments. Shallow groundwater responds and contributes to runoff according to local inputs, *e.g.*, snow melt which occurs only in the upper part of the basin will raise groundwater levels and generate runoff there without influencing groundwater levels in the lower parts.

## 2.2 Study site and model application

The study was performed in the Dreisam basin (257 km$^2$) and the two sub-basins Brugga (40 km$^2$) and Talbach (15.2 km$^2$), located in the southern Black Forest in south-west Germany (Fig. 2). The physiographic characteristics of the basins are similar (Table 1). The underlying bedrock consists mainly of gneiss, and is covered by soils of varying depths. Steep slopes, hilly uplands and relatively narrow valleys dominate the topography. An extended porous aquifer exists only in the main valley of the Dreisam basin (about 10 percent of the catchment area), and is used for the water supply of the city of Freiburg. All basins have a nival runoff regime, which is more pronounced for the smaller basins because of the higher elevations.

Each catchment was subdivided into elevation zones with a vertical extent of 100 m. The calibration period was November 1982 to October 1992 preceded by a warming-up period of one year. The model was run on a daily time step using daily temperature and precipitation as well as long-term mean monthly potential evapotranspiration as

**Table 1.** Physiographic and hydrological characteristics of the catchments

| Characteristic | Dreisam | Brugga | Talbach |
|---|---|---|---|
| Catchment size [km$^2$] | 257 | 40 | 15.2 |
| Mean elevation [m a.s.l.] | 780 | 986 | 1075 |
| *Land use* | | | |
| Forest [%] | 57.2 | 71 | 73.4 |
| Open [%] | 39.8 | 26.9 | 24 |
| Settlements [%] | 3 | 2.1 | 2.6 |
| Mean annual precipitation [mm/a] | 1500 | 1740 | 1850 |
| Mean annual runoff [mm/a] | 770 | 1200 | 1400 |

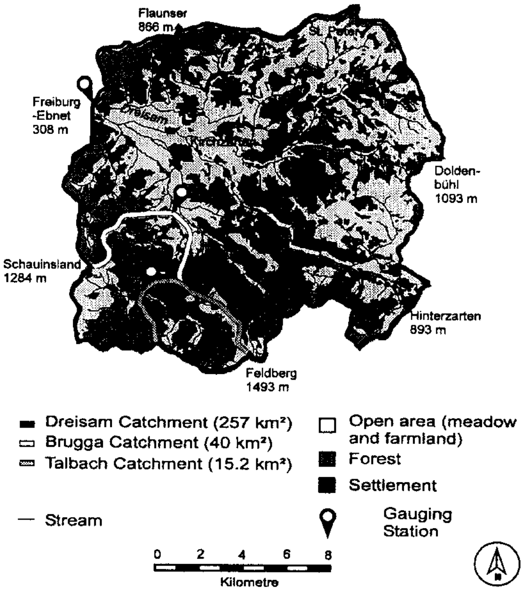

**Fig. 2.** Dreisam basin and the two nested sub-basins.

driving variables. Temperature data was available from only one station (Feldberg, 1493 m a.s.l.), whereas the areal mean precipitation was computed for each catchment from measurements at nine (or less) stations located in the Dreisam basin. Based on data from the two meteorological stations Freiburg (269 m a.s.l.) and Feldberg, the mean variations of temperature and precipitation with elevation were estimated as a decrease of temperature by 0.6 °C per 100 m and a relative increase of precipitation of 6 percent per 100 m. The potential evaporation was estimated for each catchment using the Turc-Wendling approach (Wendling, 1995).

### 2.3 Model calibration and validation

In the first step the model was calibrated for each catchment individually using a genetic algorithm, which mimics evolution in nature (e.g., Wang, 1991; Beasley *et al.*, 1993a, 1993b; Franchini, 1996). Subsequently, runoff from each catchment was simulated using the optimised parameter sets from the other two catchments. The second step was to calibrate the model simultaneously to all three catchments (*i.e.*, determining a single parameter set) using a genetic algorithm and a combined objective function (see below). As objective function the efficiency, $R_{\mathrm{eff}}$, (Nash and Sutcliffe, 1970) was used throughout this study. The efficiencies of runoff predictions for one catchment using the specific runoff from another catchment were computed for comparison.

In the genetic calibration algorithm optimised parameter sets were found by an 'evolution' of parameter sets. In each generation the chance to become a 'parent' of a new parameter set was related to the 'goodness', defined as efficiency, of the parameter set. A new parameter set was generated from the two 'parent' sets (sets A and B) by applying one of the following four rules for each parameter randomly with certain probabilities ($p$): value of set A ($p$=0.4), value of set B ($p$=0.4), random value between the values of set A and set B ($p$=0.19), random value within the limits of allowed parameter values (mutation) ($p$=0.01).

For the simultaneous calibration in the three catchments, four 'populations' of parameter sets were used. In three of them 'goodness' was defined as the efficiency, $R_{\mathrm{eff}}$, in one of the catchments, while in the fourth population 'goodness' was defined as a combined fuzzy measure (see below). After a certain number of iterations (10 in this study) parameter sets were exchanged between the populations. By this means characteristics of 'good' parameter sets in one catchment helped to optimise one parameter set valid for all catchments.

For each catchment the efficiency, $R_{\mathrm{eff}}$, of a simulation in one catchment was transformed into fuzzy measures, $f_i$, based on the efficiency obtained by the individual calibration, $R_{\mathrm{eff,\,max}}$ (Eq. 1). These fuzzy measures evaluate the degree of truth of the statement 'this parameter set is the best possible one'. The combined fuzzy measure, $F$, was computed as geometric mean of the three fuzzy measures according to each catchment (Eq. 2).

$$f_i = \max\left(0, \frac{R^i_{\mathrm{eff}} - 0.8\,R^i_{\mathrm{eff,max}}}{0.2\,R^i_{\mathrm{eff,max}}}\right) \qquad (1)$$

$R^i_{\mathrm{eff}}$ efficiency of simulation for catchment $i$

$R^i_{\mathrm{eff,max}}$ efficiency of individual calibration in catchment $i$

$$F = \sqrt[3]{f_1 \cdot f_2 \cdot f_3} \qquad (2)$$

## 3 Results

The individual calibrations resulted in good fits for the Dreisam catchment and the two sub-basins with $R_{\mathrm{eff}}$ values between 0.81 and 0.85 (Table 2). Multiscale validation, *i.e.*, runoff simulation using a parameter set optimised in one catchment in the other two catchments, gave acceptable but significantly poorer fits ($R_{\mathrm{eff}}$ on average 0.76 compared to 0.84) (Fig. 3).

The optimised parameter sets were rather similar and no obvious relations between calibrated model parameters and catchment size could be found (Table 2).

The simultaneous calibration provided one 'regional' parameter set. The fits using this parameter set were not as good as those obtained by individual calibration, but better than those using a parameter set optimised in another catchment ($R_{\mathrm{eff}}$ in average 0.81 compared to 0.84 and 0.76, respectively) (Fig.3). Individual values of this regional parameter set were approximately an average value of the individually optimised sets for most parameters.

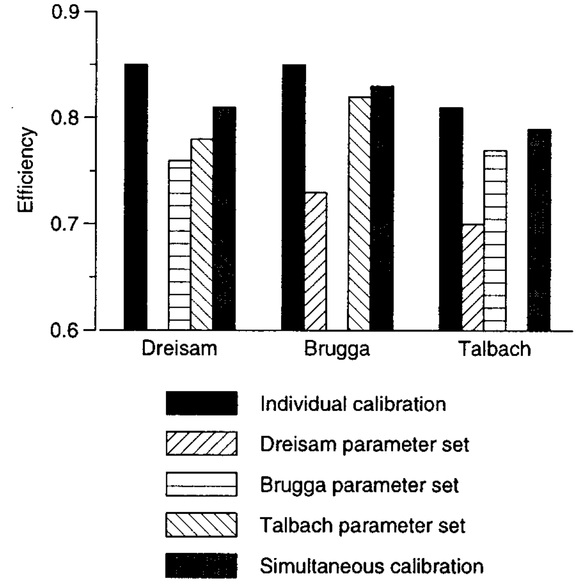

**Fig. 3.** Goodness of simulations using different parameter sets.

**Table 2.** Optimised parameter sets and efficiency of simulations

| Parameter | Explanation | Unit | Dreisam 257 km$^2$ | Brugga 40 km$^2$ | Talbach 15.2 km$^2$ | All catchments |
|---|---|---|---|---|---|---|
| **Snow routine** | | | | | | |
| TT | Threshold temperature | °C | 0.2 | 0.0 | -0.7 | 0.0 |
| CFMAX | Degree-day factor | mm °C$^{-1}$d$^{-1}$ | 1.9 | 2.1 | 1.9 | 2.1 |
| SFCF | Snowfall correction factor | - | 0.47 | 0.83 | 1.00 | 0.80 |
| CWH | Water holding capacity | - | 0.19 | 0.01 | 0.11 | 0.01 |
| CFR | Refreezing coefficient | - | 0.04 | 0.06 | 0.07 | 0.07 |
| **Soil routine** | | | | | | |
| FC | Maximum of SM (storage in soil box) | mm | 300 | 387 | 406 | 315 |
| LP | Threshold for reduction of evaporation (SM/FC) | - | 0.57 | 0.78 | 0.49 | 0.63 |
| BETA | Shape coefficient | - | 3.2 | 4.1 | 4.2 | 3.3 |
| CET | Correction factor for potential evaporation | C$^{-1}$ | 0.11 | 0.01 | 0.01 | 0.07 |
| **Response routine** | | | | | | |
| $K_0$ | Recession coefficient (upper box) | d$^{-1}$ | 0.46 | 0.22 | 0.27 | 0.30 |
| $K_1$ | Recession coefficient (upper box) | d$^{-1}$ | 0.19 | 0.14 | 0.19 | 0.18 |
| $K_2$ | Recession coefficient (lower box) | d$^{-1}$ | 0.07 | 0.05 | 0.07 | 0.06 |
| UZL | Threshold for $Q_0$ outflow | mm | 33 | 41 | 40 | 42 |
| PERC | Maximal flow from upper to lower box | mm d$^{-1}$ | 4.1 | 4.5 | 4.7 | 4.5 |
| MAXBAS | Routing, length of weighting function | d | 1.0 | 1.2 | 1.1 | 1.0 |
| $R_{eff}$ | Efficiency of simulation | - | 0.85 | 0.85 | 0.81 | 0.81/0.83/0.79 |

## 4 Discussion

In general, the variations of parameter values between the three catchments were small. This was not surprising given the similarity of the catchments. However, at least for some of the parameters variations caused by scale effects could have been expected from the differences of the catchment area.

Only three catchments were used in this study and for some parameters the optimised values must be considered to be subject to large parameter uncertainty, *i.e.*, values over wide ranges may provide good simulations (Seibert, 1997a). Variations of parameter values with changes in catchment characteristics can, therefore, only be interpreted as indications. Simulations are not very sensitive to changes in parameter values for some parameters in the HBV model. For these parameters, variations between the catchments may be caused mainly by chance. Examples of such parameters are the refreezing coefficient (CFR) or the correction coefficient for evaporation (CET). Parameter sensitivity and its relationship to catchment characteristics should be investigated in more detail.

The parameter SFCF (correction factor for snowfall) can be interpreted as a representation of losses from the snow pack caused by evaporation of intercepted snow (Seibert, 1998). A decrease of optimised values with decreasing elevations is therefore reasonable. The large decrease for the Dreisam catchment can not be explained by increasing snow losses only. The low value for SFCF in this catchment was partly caused by too low values of observed runoff. There are two main sources for systematic errors: from the aquifer in the main valley, water is withdrawn for water supply for which the runoff data was approximately corrected and water bypasses the gauging station within the porous aquifer (not corrected for).

The maximum storage in the soil box (FC) was expected to increase from the Talbach to the Dreisam catchment both from scale and physiographic considerations (Uhlenbrook *et al.*, 1998). The optimised values, however, varied in the opposite way. For a given soil moisture status the BETA parameter in the soil routine determines the portion of rainfall and snowmelt which contributes to groundwater storage, *i.e.*, the recharge decreases with increasing BETA values. The BETA value can be interpreted as an indicator of the relationship between the catchment soil moisture and the extension of contributing areas. BETA can also be interpreted as a representation of the permeability of the soils. For both reasons, it could have been expected to increase with catchment size. The relative contributing area tends to be larger in small catchments than in larger ones and the soils in the Dreisam catchment are more permeable in areas with higher altitudes (Stahr, 1979). On the other hand, steeper slopes and the absence of extensive aquifers can explain the larger values of BETA in the smaller catchments.

The recession coefficients ($K_0$, $K_1$ and $K_2$) were expected to decrease with increasing catchment size because of a more damped and even hydrograph in a larger catchment. The values were almost identical for the three catchments except for $K_0$ where the value was much larger in the Dreisam catchment. The outflow controlled by $K_0$ is active only during conditions with very high flows and its values were influenced to a large degree by an extreme single event in December 1991 in the Dreisam catchment, which was the highest discharge ever measured. This event was mainly caused by contributions from the northern sub-basins whereas the event was less extreme in the southern sub-basins (Talbach and Brugga).

It might be possible that variation in different characteristics cancelled one another. Whilst higher values for the recession coefficients, for instance, had been expected from scale considerations in the two smaller catchments, the opposite could be expected from the geological differences with an extended porous aquifer in the main valley of the Dreisam catchment compared to mainly fissured hard rock aquifers in the smaller basins.

The routing parameter MAXBAS was expected to increase with increasing catchment size because of the increasing channel length. However, the differences in catchment size were small compared to the daily time step and the values close to one were reasonable, *i.e.*, water entering the streams contributes to runoff at the outlet within one time step.

The fact that a parameter set optimised in one catchment was, more or less, valid in the other catchments was related to the small differences in the optimised parameter values. This result indicated that parameter values did not depend strongly on catchment size. Parameter value variations were probably caused primarily by changes in other catchment characteristics, even if these were small.

Runoff prediction for ungauged catchments is an important task in hydrology. One possibility is to use a hydrological model with parameter values optimised in a nearby gauged catchment. A simple alternative is the direct use of runoff series from a nearby catchment scaled by the catchment size. Compared to this alternative, the use of the model approach to predict runoff in one catchment (Fig. 3) reduced the errors significantly in two of the three cases, whereas the two methods were of a similar quality for predictions of runoff from the Brugga catchment using the Talbach catchment and *vice versa* (Table 3). The main reason for these results are assumed to be the differences in temperature and precipitation, which are considered only in the modelling approach.

**Table 3.** Efficiency of runoff predictions using specific runoff series from one of the other catchments

| Catchment used to compute runoff | Efficiency of computed series for... | | |
|---|---|---|---|
| | Dreisam | Brugga | Talbach |
| Dreisam | - | 0.64 | 0.52 |
| Brugga | 0.42 | - | 0.87 |
| Talbach | -0.17 | 0.80 | - |

## 5 Concluding remarks

The results have to be interpreted with care since only three catchments were used in this study. Although the catchments were nested and thus not independent, the direct dependence was minor as the portion of the nested catchment area was not large in any cases. The results indicate that parameter values may not depend strongly on variations of the spatial scale within the range of 15 $km^2$ to 250 $km^2$. Results were different in a comparable study on the same catchments (Uhlenbrook *et al.*, 1998), where another calibration method, a slightly different model structure, and different data series were used.

In this study, the catchments were not divided into sub-basins in which the model was run in parallel for the simulations. The catchment is often subdivided for different reasons, *e.g.*, to divide the catchment into areas above and below a lake (Bergström, 1992), in other applications of the HBV model or other lumped models. This study indicates that such a subdivision, within the investigated scale, probably does not introduce scale effects in parameter values.

The simultaneous calibration to runoff series from different catchments within a region provides a regional parameter set. The method based on the genetic algorithm proposed in this study was found to be a suitable tool for this simultaneous calibration. Future studies will show if this parameter set is the most efficient for simulation of runoff from an ungauged catchment within the region. For comparison and regionalisation of parameter values between different regions such ‘regional’ parameter sets may be more suitable and robust than parameter sets optimised for individual catchments as they represent a regional average.

**Acknowledgements.** The runoff data used in this study were collected by the Gewässerdirektion Südlicher Oberrhein / Hochrhein, Waldshut, Germany. The precipitation and temperature data were provided by the German weather service (Deutscher Wetterdienst, DWD), Offenbach, Germany. The help from J. Holocher during data composition and in creating Fig. 2 is also gratefully acknowledged.

## References

Beasley, D., Bull, D.R., and Martin, R.R. (1993a) An overview of genetic algorithms: Part 1, Fundamentals, *University Computing*, **15**: 58-69

Beasley, D., Bull, D.R., and Martin, R.R. (1993b) An overview of genetic algorithms: Part 2, Research Topics, *University Computing*, **15** 170-181

Bergström, S. (1976) Development and application of a conceptual runoff model for Scandinavian catchments, *SMHI, Report* No. RHO 7, Norrköping, 134 pp.

Bergström, S. (1992) The HBV model - its structure and applications, *SMHI Hydrology, RH* No.4, Norrköping, 35 pp.

Beven, K. J. (1991) Scale considerations, In: D. S. Bowles and P. E. O'Connell (eds.) *Recent Advances in the Modeling of Hydrologic Systems*, Kluwer Academic Publishers, Netherlands, pp. 357-371

Franchini, M. (1996) Use of a genetic algorithm combined with a local search method for the automatic calibration of conceptual rainfall-runoff models, *Hydrological Sciences Journal*, **41**: 21-40

Harlin, J. and Kung, C.-S. (1992) Parameter uncertainty and simulation of design floods in Sweden, *Journal of Hydrology*, **137**, 209-230

Mimikou, M. (1984) Regional relationships between basin size and runoff characteristics, *Hydrological Sciences Journal* **29**: 63-73

Nash, J. E. and Sutcliffe, J. V. (1970) River flow forecasting through conceptual models; Part I - A discussion of principles, *Journal of Hydrology*. **10**, 282-290.

Seibert, J. (1997a) Estimation of parameter uncertainty in the HBV model, *Nordic Hydrology*, **28**, 247-262

Seibert, J. (1997b) HBV light, version 1.2, User's manual, *Uppsala University, Department of Earth Sciences, Hydrology*, 31 pp.

Seibert, J. (1999) Regionalisation of parameters for a conceptual rainfall-runoff model, accepted by *Agricultural and Forest Meteorology*

Stahr, K. (1979) Die Bedeutung periglazialer Deckschichten für Bodenbildung und Standorteigenschaften im Südschwarzwald, *Freiburger Bodenkundliche Abhandlungen*, No. 9, 273 pp.

Uhlenbrook, S.; Holocher, J., Leibundgut, Ch., and Seibert, J. (1998) Using a conceptual rainfall-runoff model on different scales by comparing a headwater with larger basins; In: Kovar, K.; Tappeiner, U.; Peters, N.E. and R.G. Craig: *Hydrology, Water Resources and Ecology in Headwaters*; *IAHS-Pub.* No. 24, 297-306

Wang, Q.J. (1991) The genetic algorithm and its application to calibrating conceptual rainfall-runoff models, *Water Resources Research*, **27**: 2467-2471

Wendling, U. (1995) Berechnung der Gras-Referenzverdunstung mit der FAO Penman-Monteith-Beziehung, *Wasserwirtschaft*, **85**: 602-604