PREDICTION OF MEAN MONTHLY DISCHARGES IN UNGAUGED BASINS

Kamila Hlavčová, Ján Szolgay, Silvia Kohnová, Robert Danihlík, Marcel Zvolenský

Department of Land and Water Resources Management, Slovak University of Technology, Bratislava, Slovak Republic, e-mail: kamila.hlavcova@stuba.sk, jan.szolgay@stuba.sk, silvia.kohnova@stuba.sk, robert.danihlik@stuba.sk, marcel.zvolensky@stuba.sk

Abstract: The necessity to generate time series of runoff for planning and design purposes and environmental protection at ungauged sites is often the case in water resources studies. As in the case of the absence of measured runoff optimisation techniques cannot be used to estimate the parameters of rainfall-runoff models, regional estimation methods are used instead. In several studies regional regression was used for relating the parameters of rainfall-runoff models to catchment characteristics. In Hlavčová et al. (2000) a different method for the regional calibration of a monthly water balance model was proposed. Instead of using the regional regression, the new method involved the regional calibration of a monthly rainfall runoff model to several gauged catchments in a given region simultaneously using a regional fitness function. In the first study a small group of similar catchments was subjectively selected using hydrological reasoning; in the second three homogeneous groups of catchments were pooled using cluster analysis of selected basin physiographic properties. For the model calibration a genetic programming algorithm was employed in both studies. In the Szolgay et al. (2003b) both regional approaches were compared with two single site calibration methods and the trial and error calibration in group of 7 catchments from Hlavčová et al. (2000). It was concluded, that the regionally calibrated model parameters can be used in ungauged basins with similar physiographic conditions. Here a comparison of such a regional calibration method with results of single site calibration is performed using data from all 14 catchments selected for regionalisation in Szolgay et al. (2003b) in order to get a broader basis for conclusions about the performance of the new method.
Keywords: monthly rainfall runoff model, prediction in ungauged catchments, genetic algorithm, regional calibration, cluster analysis.

SCHÄTZUNG MONATLICHER ABFLÜSSE IM NICHT BEOBACHTETEN EINZUGSGEBIETEN

Zusammenfassung: Eine Methode zur Schätzung monatlicher Abflüsse im nicht beobachteten Einzugsgebieten mit Hilfe vom mathematischen Niederschlag-Abflussmodellen wurde vorgeschlagen. Regional gültige Modellparameter wurden in homogenen Regionen mit Hilfe genetischer Algorithmen so optimiert, dass das Modell in allen beobachteten Einzugsgebieten und im ganzen Gebiet zugleich akzeptable Ergebnisse liefert. Die homogenen Gebiete wurden Anhand von Einzugsgebietparameter mittels Clusteranalyse ermittelt.
Schlüsselworte: monatlicher Abflüsse, Niederschlag-Abflussmodelle, nicht beobachtete Einzugsgebiete, regionale Optimierung, genetische Algorithmen, Clusteranalyse

1. Introduction

In ungauged basins classical model calibration through optimisation techniques cannot be used to estimate the parameters of conceptual hydrologic rainfall runoff models. Since parameters of such models cannot be either derived directly from catchment characteristics or measured, methods for their regional estimation were often sought. The main problem in such a case is related to the determination of model parameter values from catchment properties. Many studies have followed the same methodological approach. First, a watershed model was calibrated based on data available at a number of sites in a region. This was followed by the application of a regionalisation method, which attempted to relate the calibrated model parameters to the catchment characteristics. The most common method that was used was bivariate and multivariate regression. Abdulla and Lettenmaier (1997), Sefton and Howarth (1998) and Xu and Singh (1998) provide more detailed reviews of such studies.

In Hlavčová et al. (2000) and Szolgay et al. (2003b) a different methodology was proposed for the regionalisation of watershed model parameters. It involved the concurrent calibration of a model to available sites in the Záhorie region in seven catchments. The objective of the model calibration was to reproduce the behaviour of observed monthly streamflows at individual sites and in the region as a whole as well.

In Hlavčová et al. (2000) it was attempted to select such catchments subjectively, in Szolgay et al. (2003b) cluster analysis based on selected basin physiographic properties was used to pool catchments into homogeneous groups. For the model calibration a genetic programming algorithm was employed in both studies. In Szolgay et al. (2003b) both regional approaches were compared with two single site calibration methods in the group of 7 catchments from Hlavčová et al. (2000). It was shown, that the regionally calibrated model parameters could be than used in ungauged basins with similar physiographic conditions in water resources studies. In this study the performance of such a regional calibration scheme is compared with a single site calibration method in the Záhorie region of West Slovakia using all 14 catchments used for catchment pooling in Szolgay et al. (2003b).

2. Input data and selection of pilot basins

The pilot area (Figure 1) lies in the lower left-side part of the Morava River watershed in the Slovak Republic near the Austrian and Czech border, and it is comprised of the north-western slopes of the Small and White Carpathian Mountains. The transboundary Morava River flows through three countries: the Czech Republic, the Slovak Republic and Austria. The catchment area is 26,580 km^2 and only 8.6 % is situated in the area of the Slovak Republic. The pilot area is drained by the Chvojnica, Myjava, Rudava and Malina rivers to the Morava River.

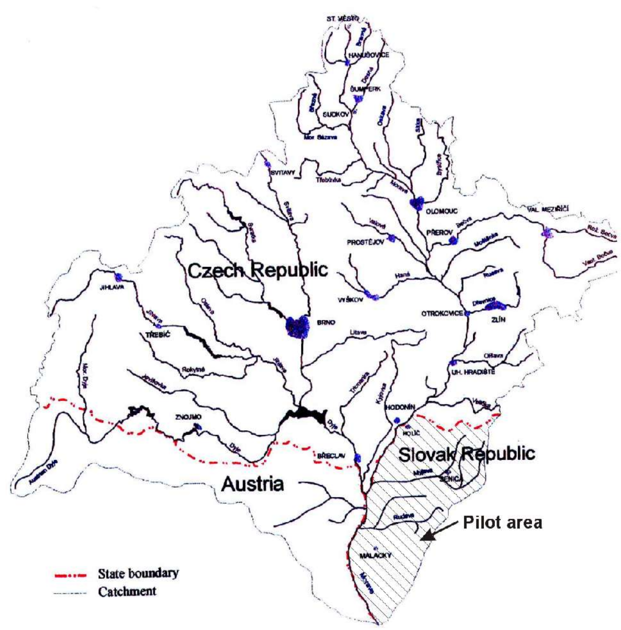

Figure 1. The Morava River watershed and the position of the pilot area in its Slovak part

The relief of the mountains is a typical hillyland relief with low, flat ridges and wide valleys. The highest altitudes of the mountains are up to 800 m a.s.l.; the lowest parts of the area are located to the SSE of the town of Senica with altitudes of 220 - 250 m a.s.l. The values of the mean annual runoff yield are in the range of 2 $l.s^{-1}.km^{-2}$ to 9 $l.s^{-1}.km^{-2}$.

For this study the data series from Szolgay et al. (2003b) was used, so 14 catchments from the analysed territory were available. Table 1 contains the list of the catchments with their identification codes and the length of available runoff data series, location of gauge stations is illustrated in Figure 2. The length of available data series varies from 6 to 17 years, which reflects the usual situation that practitioners are dealing with in water resources studies in the region.

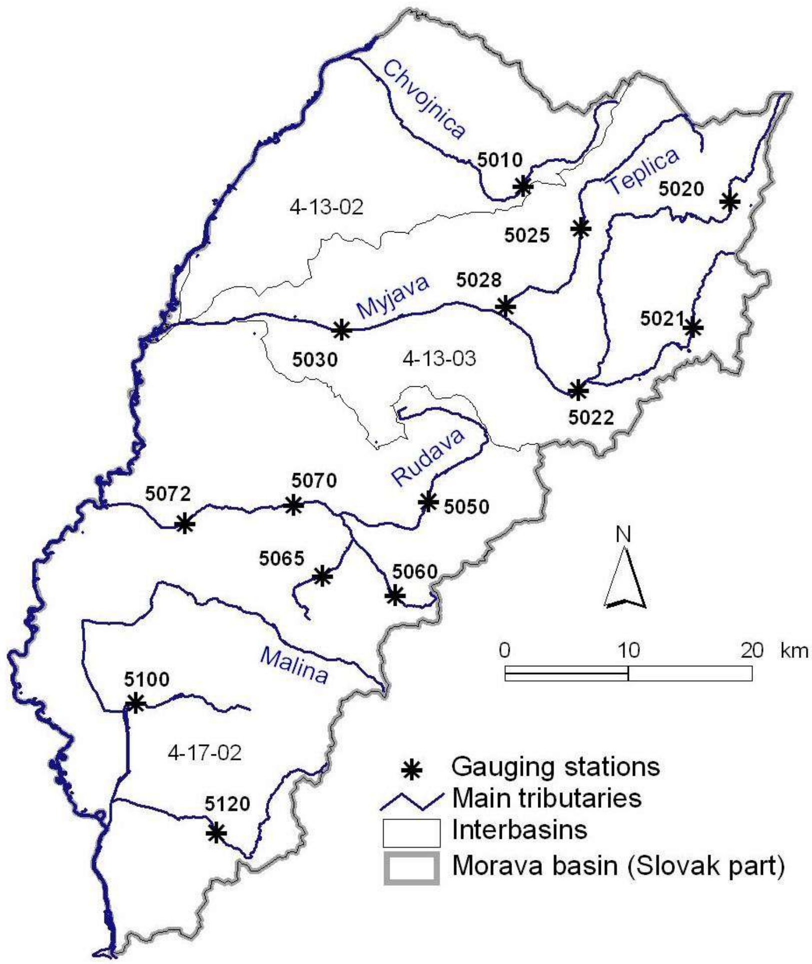

Figure 2. Location of selected gauging stations in the pilot area

Several climatic and physiographic catchment characteristics were derived from a set of digitised hydrological maps and a 100 m resolution raster DEM with the help of GIS methods for this study: the catchment area (*F,* km^2), the gauge datum (*NV*, m a.s.l.), the mean catchment slope (*SL*, %), the mean catchment elevation (*Hpr,* m a.s.l.), the mean hang orientation (*ASP*, degree), the long-term mean annual runoff from a runoff yield map of 1930-1960 (*ODT*, $l.s^{-1}.km^{-2}$), the long-term mean annual precipitation amounts (*Zr*, mm) and the long-term mean annual evapotranspiration (*E*, mm) from the period 1931-80, the long-term mean annual runoff (*ODT-bil*, mm) as computed from the water balance equation from the period 1931-80 as *ODT-bil* = *Zr* – *E*, and the long-term mean annual temperature (*T*, °C). The catchment characteristics for all catchments are presented in Table 2.

Table 1. List of selected catchments with their identification codes and the length of available runoff data

Identification code	Gauging station	River	Observation period
5010	Lopašov	Chvojnica	1981-1997
5020	Myjava	Myjava	1981-1997
5021	Brezová pod Bradlom	Brezovský potok	1985-1997
5022	Jablonica	Myjava	1981-1997
5025	Sobotište	Teplica	1981-1997
5028	Senica	Teplica	1992-1997
5030	Šaštín – Stráže	Myjava	1981-1997
5050	Plavecký Mikuláš	Rudava	1981-1993
5060	Sološnica	Sološnický Potok	1981-1993
5065	Rohožník	Rudávka	1981-1997
5070	Studienka	Rudava	1981-1997
5072	Veľké Leváre	Rudava	1981-1997
5100	Láb	Močiarka	1981-1997
5120	Borinka	Stupávka	1982-1997

Table 2. Physiographic and climatic characteristics of the selected catchments

Id	*F* (km^2)	*NV* (m a.s.l.)	*Hpr* (m a.s.l.)	*ASP* (degrees)	*SL* (%)	*ODT* (l.s^{-1}.km^{-2})	*Zr* (mm)	*E* (mm)	*ODT-bil* (mm)	*T* (^{o}C)
5010	31.13	272.7	445.2	195.8	8.72	6.03	623	483	140	7.4
5020	31.39	324.3	425.9	216.9	12.44	8.77	656	465	191	6.7
5021	35.86	259.3	366.2	164.3	1.43	6.14	667	487	180	7.9
5022	238.45	230.5	425.9	216.9	12.40	8.77	657	491	166	7.8
5025	85.58	236.2	401.6	172.2	8.23	4.21	438	335	103	5.1
5028	152.01	188.5	254.7	155.1	6.49	4.28	472	308	164	5.8
5030	644.89	164.2	282.0	178.4	3.25	4.92	595	472	123	7.8
5050	98.55	180.0	239.9	225.3	2.44	5.48	669	500	169	8.2
5060	10.38	245.3	397.4	304.5	6.23	7.50	754	500	254	7.3
5065	26.10	192.5	314.5	200.0	6.42	6.35	702	500	202	8.3
5070	280.32	170.8	271.4	161.3	4.54	5.42	672	500	172	8.3
5072	304.41	152.8	262.6	217.6	1.99	5.16	665	499	166	8.4
5100	47.10	144.3	247.1	213.9	2.96	5.60	670	498	172	8.7
5120	33.76	216.7	473.9	188.3	7.91	7.50	752	501	251	7.8

3. Single site calibration of the WatBal model

In Slovakia the conceptual spatially-lumped hydrological rainfall-runoff model, the WatBal (Yates, 1994) was extensively used for modelling river runoff in a monthly time step in water resources studies and climate change impact studies. (see e.g. Hlavčová, Čunderlík 1998, Hlavčová, et al. 1999, Szolgay, et al. 2003a). Other rainfall runoff model applications in water resource and climate change studies were reported in Halmová (2000), Kostka and Holko (2000, 2001), Majerčáková (2000), Majerčáková and Takáčová (2001), Petrovič (2000), Štekauerová and Nagy (2001), Takáč (2001).

WatBal is a conceptual lumped rainfall-runoff model, which simplifies a river basin into a single nonlinear reservoir. The model simulates water accumulation in the catchment, snowmelt, evapotranspiration, runoff from impermeable areas in the basin, surface and subsurface runoff and baseflow. The inputs required for water balance modelling when using a monthly time step are: the mean monthly precipitation for the basin, the mean monthly river discharges in the closing profile of the basin and the mean monthly potential evapotranspiration (PET). If the PET data is not available, the model uses either the Thornthwaite or the Priestly-Taylor method. The Priestly-Taylor method was used in this study. It requires additional data: the mean long-term monthly hours of sunshine, the mean long-term monthly values of relative air humidity, and the mean monthly air temperature values. For full description of the model see Szolgay et al. (2003b).

In the present study a genetic algorithms (GA) was applied to calibrate the model to at site data. GA are stochastic search methods that simulate the process of the natural selection and the mechanism of population genetics, detailed description of the GA can be found, e.g., in Bäck (1996). It mimics the mechanism of biological evolution, i.e., natural selection, inheritance and mutation. GA combines selection, crossover, and mutation operators with the goal of finding the best solution to an optimisation problem. The GA searches for the optimal solution until a specified termination criterion is met. It searches from a population of decision variable sets, and it includes a random factor, which should help to avoid stopping at local optima. In recent years it has been widely applied in a number of fields.

Kuczera (1997) warned that a genetic algorithm using traditional crossover tended to flounder near the optimum and could not be relied upon to locate the global optimum. Together with Franchini, et al. (1998), he found, that the Shuffled Complex Evolution algorithm approximates better than other tested approaches. For these reasons in this study the so called Generational GA (GGA) was used. The GGA is a type of genetic algorithm in which the entire population is replaced in each iteration. This method of progression for a genetic algorithm has proven to work well for a wide variety of problems. It tends to be a little slower than some other of its modifications, but it also tends avoid local minima. For the calibration the Nash-Sutcliffe criterion, which is widely used in modelling studies, was used. :

$$R^2 = \frac{F_0 - F}{F_0} \tag{1}$$

with:

$$F_0 = \sum_{i=1}^{N} \left(Q_{Obs_i} - \overline{Q}_{Obs} \right)^2 \tag{2}$$

$$F = \sum_{i=1}^{N} \left(Q_{Obs_i} - Q_{Comp_i} \right)^2 \tag{3}$$

$$\overline{Q}_{Obs} = \left(\sum_{i=1}^{N} Q_{Obs_i} \right) / N \tag{4}$$

where N is number of data, Q_{Obs} are the observed and Q_{Comp} the simulated flows, and $\overline{Q}_{Obs}$ and $\overline{Q}_{Comp}$ are their respective mean values.

4. Regional calibration of the Watbal model using genetic algorithm

To arrive at a set of regionally feasible parameters to be used in ungauged catchments, the idea of the simultaneous calibration of the model on all the available data from a group of similar catchments was therefore adopted. Here, following the idea of Burn and Boorman (1992), cluster analysis using selected physiographic catchments characteristics as discrimination variables was used for the grouping similar catchments. Here the results from

Szolgay et al. (2003b) were overtaken. The similarity measure of catchments was defined in the space of selected physiographic catchment parameters. K-means clustering with Euclidean metrics was selected for the analysis. Given the small number of available catchments for this study, in the first step the number of resulting clusters was restricted to be between 2 to 4 in the given case. Several sets consisting of two to four variables were (subjectively) chosen from the available set of those catchment characteristics. These were believed to have significant influence on the runoff formation in the catchments and were not correlated with each other; the degree of low correlation was set to 0.4 in this study. The quality of each clustering trial was tested through the visual inspection of the cluster profile plots. Other methods usually employed in regional frequency analysis (such as homogeneity test) were not performed here due to the restricted amount of available data.

Finally grouping of catchments based on the catchment area (*F,* km^2), the mean catchment slope (*SL*, %), and the long-term mean annual precipitation amounts (*Zr*, mm) were selected as feasible. These variables are believed to be relevant for the runoff formation in the given region, they are not mutually correlated and indicate a reasonable selective grouping of the catchments involved. Table 3 contains the catchment indicator numbers of the catchment in the respective groups.

Table 3. Catchment identification numbers of the catchment pooled in the respective groups

Pooling group	Identification numbers of catchments
1	5021,5050,5060,5065,5100,5120
2	5010,5020,5022,5025,5028
3	5030,5070,5072

In Hlavčová et al. (2000) and Szolgay et al. (2003b) also an objective function for regional calibration of models was suggested. Its version from the second study, which made it more flexible by adding an additional weight coefficient, was adopted here in this study. The objective function (fitness) was minimized over the entire data set in each individual catchment and in all catchments. The objective function is based on the sum of absolute values of the differences between computed and measured runoff (hydrograph fit) and on the absolute values of the differences between the sums of modelled and observed runoff (the hydrological balance of the catchment). It combines both difference values in a weighted sum; the user can determine the two weight coefficients empirically. Thus the objective of the calibration was to reproduce the behaviour of observed monthly streamflows at individual sites and in the region as a whole as well as.

5. Results

Since all the available data had to be used in the regional calibration of the model, no independent data set was available in the region to test the performance of the clustering process and that of the regional calibration method. Instead, it was decided to compare the performance of the proposed method with the single site calibration. Model performance was measured by the Nash- Sutcliffe criterion.

Table 4 contains the values of both performance criteria for the different calibration methods in the 14 catchments. The regional genetic algorithm calibration methods degraded the model performance in individual catchments when compared to the single site GA approach. Given the usual practical limitations of water resources studies (such as short data series, poor spatial coverage of rainfall data, changing quality of data etc.) it is believed, that it was possible to demonstrate here, that the proposed method performed within limits set by such conditions and be regarded as satisfactory for planning studies especially if there is no option available, which could improve data availability or quality in the region of interest.

Table 4. Comparison of Nash-Sutcliffe model performance criteria for both methods GA single and GA regional for 3 groups of catchments selected cluster analysis

Pooling group	Catchment identification number	Single site calibration	Regional calibration
1	5021	0,614	0,583
	5050	0,505	0,435
	5060	0,352	0,201
	5065	0,399	0,360
	5100	0,323	0,235
	5120	0,695	0,480
2	5010	0,567	0,424
	5020	0,628	0,349
	5022	0,635	0,581
	5025	0,531	0,463
	5028	0,556	0,433
3	5030	0,584	0,546
	5070	0,632	0,489
	5072	0,528	0,528

6. Conclusions

In the case of the absence of measured runoff series single site calibration techniques cannot be used to estimate the parameters of rainfall-runoff models, regional estimation methods are to be used instead. Usually regression methods were suggested to be used for relating the model parameters to the catchment characteristics in a given region in such cases. For the modelling of monthly discharges at an ungauged site this commonly used methodology was not attempted here but a different method was proposed. This method involves the regional calibration of a monthly water balance model to several similar gauged catchments in a given region simultaneously. These catchments were pooled together using cluster analysis of selected basin physiographic properties. For the model calibration a genetic programming algorithm was employed. A specific fitness a function was applied which is believed to be more flexible and appropriate for water resources studies. It is based on the sum of absolute values of the differences between computed and measured runoff (hydrograph fit) and on the absolute values of the differences between the sums of modelled and observed runoff (the hydrological balance of the catchment) and it combines both values in a weighted sum. The model performance of the proposed regional calibration method was found satisfactory; also a degradation of the performance occurred when compared with the on site results using the GA approach. However since short data series, poor spatial coverage of rainfall data, changing quality of data etc. are usual practical limitations of water resources studies, it is believed that it was possible to demonstrate here, that the proposed method performed within acceptable limits.

7. Acknowledgements

The research presented in this paper has been supported by the Science and Technology Assistance Agency under the contract No. APVT-51-006502 and the Slovak Grant Agency under the projects 1/1145/04 and 2/3085/23. All supports are gratefully acknowledged.

8. References

Abdulla, F.A., Lettenmaier, D.P. (1997): *Development of regional parameter estimation equations for a macroscale hydrologic model.* Journal of Hydrology, 197, 230-257.

Bäck, T. (1992): *Evolutionary Algorithms in Theory and Practice.* New York: Oxford University Press.
Burn, D.H., Boorman, D.B. (1992): *Catchment Classification Applied to the Estimation of hydrological Parameters at Ungauged Catchments.* Report No. 118. Institute of Hydrology, Wallingford, 71 pp.
Franchini, M., Galeati, G., Berra, S. (1998): *Global optimisation techniques for the calibration of conceptual rainfall-runoff models.* Hydrological Sciences Journal, 43, No. 3, 43-458.
Halmová, D. (2000): *Impact of the climate change upon multipurpose water reservoir Orava*. Acta Hydrologica Slovaca, Vol. 1, No. 2, 3-12.
Hlavčová, K., Čunderlík, J. (1998): *Impact of climate change on the seasonal distribution of runoff in mountainous basins in Slovakia*. Hydrology, Water Resources and Ecology in Headwaters. IAHS Publ. No. 248, 39 - 46.
Hlavčová, K., Szolgay, J., Čunderlík, J., Parajka, J., Lapin, M. (1999): *Impact of climate change on the hydrological regime in Slovakia.* Publication of the Slovak Committee for Hydrology No. 3. STU a SVH MHP UNESCO, Bratislava, 101 pp. (ISBN 80-227-1296-5)
Hlavčová, K., Szolgay, J., Čistý, M., Kohnová, S., Kalaš, M. (2000): *Estimation of mean monthly flows in small ungauged catchments.* Slovak Journal of Civil Engineering, Vol. VIII, No. 4, 21-29.
Kostka, Z., Holko, L. (2000): *Impact of climate change on runoff in a small mountain catchment*. Publication of National Climate Programme of the Slovak Republic, No. 8., 91-109, Slovak Ministry of the Environment, Bratislava.
Kostka, Z., Holko, L. (2001): *Vegetation change and hydrological regime of a mountain catchment.* Publication of National Climate Programme of the Slovak Republic, No. 10., 82-93, Slovak Ministry of the Environment, Bratislava.
Kuczera, G. (1997): *Efficient subspace probabilistic parameter optimisation for catchment models.* Water Res. Res., Vol. 33, No.1, 177-185.
Majerčáková, O. (2000): *Hydrological monitoring and modelling of possible runoff changes.* Publication of National Climate Programme of the Slovak Republic, No. 9., 5-14, Slovak Ministry of the Environment, Bratislava.
Majerčáková, O., Takáčová, D. (2001): *The possible impact of climate change upon the groundwater levels in alluvia.* Publication of National Climate Programme of the Slovak Republic, No. 11., 31-49, Slovak Ministry of the Environment, Bratislava.
Petrovič, P. (2000): *Climate change impact on hydrological regime in the Nitra river basin*. Publication of National Climate Programme of the Slovak Republic, No. 9., 58-73, Slovak Ministry of the Environment, Bratislava.
Sefton, C.E.M., Howarth, S.M. (1998): *Relationships between dynamic response characteristics and physical descriptors of catchments in England and Wales*. Journal of Hydrology, 211, 1-16.
Štekauerová, V., Nagy, V. (2001): *Impact of climate conditions on water supply for plants in localities Báč and Bodíky.* Acta Hydrologica Slovaca, Vol. 2, No. 1, 58-63.
Szolgay, J., Hlavčová, K., Lapin, M., Danihlík, R. (2003a): *Impact of climate change on mean monthly of runoff in Slovakia.* Meteorological Journal, Vol. 6, 2003, No. 3.
Szolgay, J., Hlavčová, K., Kohnová, S., Danihlík, R. (2003b): *Regional estimation of parameters of a monthly balance model.* J. Hydrol. Hydromech., roč. 51, 2003, č. 4, 256-273.
Takáč, J. (2001): *Climate change impacts on water balance in the agricultural land*. Publication of National Climate Programme of the Slovak Republic, No. 10., 16-26, Slovak Ministry of the Environment, Bratislava.
Xu, C.Y., Singh, V.P. (1998): *A review of monthly water balance models for water resource investigation.* Water Resources Management, 12, 31-50.
Yates, D. (1994): *WatBal – Integrated Water Balance Model for Climate Impact Assessment of River Basin Runoff.* Working Paper 94-64, IIASA.