Selection of Catchment Descriptors for the Physical Similarity Approach. Part I: Theory

Martin HEŘMANOVSKÝ and Pavel PECH

Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Prague, Czech Republic

Abstract

Heřmanovský M., Pech P. (2013): **Selection of catchment descriptors for the physical similarity approach. Part I: Theory**. Soil & Water Res., **8**: 133–140.

This paper focuses on a description of the method used for the identification of optimal catchment descriptors for the physical similarity approach consisting of a scheme for the identification of optimal catchment descriptors and the procedure for finding hydrologically homogeneous regions using inverse clustering. Andrews' curves are used as the basis for homogeneity checking. The identification of an optimum catchment descriptor is based on the assumption that the addition of an optimal catchment descriptor to a predefined set of catchment descriptors improves the accuracy of model parameter estimation within a set of tested catchments. Two criteria are proposed for the selection of optimal catchment descriptors – a criterion evaluating estimates of model parameters on the basis of different potentially optimal groups of catchment descriptors, *MIN*, and a criterion evaluating the improvement in model parameter estimation after the addition of a potentially optimal catchment descriptor into the group of preliminarily identified optimal catchment descriptors, *MAX*. The proposed method provides an alternative to the trial-and-error method for the identification of optimal catchment descriptors.

Keywords: catchment characteristic; inverse clustering; regionalisation

Conceptual rainfall-runoff models are standard tools that can be used for the prediction of runoff on catchments of interest. These models have a simpler structure than their physically-based alternatives and generally lower demands on input data (Clarke 1973; Dingman 2002; Wagener *et al*. 2004). The problem with using conceptual hydrological models is related to their application in the so-called ungauged catchments, where runoff records are either very short or poor, or do not exist (Sivapalan *et al*. 2003). Hydrological regionalisation often provides a solution to this problem (e.g. in Jarboe & Haan 1974; Magette *et al*. 1976). The regionalisation process can be implemented (according to the applied method) using three approaches: the spatial proximity approach (e.g. Vandewiele & Elias 1994; Oudin *et al*. 2008), the regression approach (e.g. Magette *et al.* 1976; Xu 1999; Wagener & Wheater 2006) and the physical similarity approach (e.g. Acreman & Sinclayr 1986; Burn & Boorman 1993; Parajka *et al*. 2005; Young 2006).

Nowadays, there is a large number of hydrological studies which compare the above-mentioned regionalisation approaches. However, these studies often give different results (e.g. Parajka *et al*. 2005; Kay *et al*. 2006; Oudin *et al*. 2008; Zhang & Chiew 2009). The differences in the results have three main causes. The first one is related to the input data for regionalisation. Different authors used different catchment data sets in their studies (e.g. Merz & Blöschl 2004 – a set of 308 catchments in Austria, Young 2006 – a set of 260 catchments in GB, Oudin *et al*. 2008 – set

of 913 catchments in France, Zhang & Chiew 2009 – a set of 210 catchments in Australia) and different sets of catchment descriptors. The second cause lies in the hydrological model applied (HBV – Merz & Blöschl 2004; Parajka *et al*. 2005; PDM – Kay *et al.* 2006; Young 2006; SIMHID and Xinanjiang – Zhang & Chiew 2009; TOPMO and GR4J – Oudin *et al*. 2008). The third cause is related to the fact that each regionalisation approach employs a different procedure and that individual studies differ in this respect (e.g. Parajka *et al*. 2005 vers. Oudin *et al*. 2008).

The best regionalisation approach (based on the published results) seems to be the spatial proximity approach using an appropriate method of parameter or output estimation. Conversely, the worst regionalisation approach seems to be the regression approach. The physical similarity approach generally provides better results than the regression approach but worse results than the spatial proximity one. Some combinations of these regionalisation approaches were tested in order to improve the regionalisation results. However, these combinations often lead to no more than a slight improvement of the regionalisation results. Zhang and Chiew (2009) combined e.g. the spatial proximity and the physical similarity approach, but they only achieved marginally better results than with the spatial proximity approach alone.

The spatial proximity approach is criticised in the literature (e.g. Acreman & Sinclayr 1986; Wagener & Wheater 2006), because the geographical closeness of catchments does not guarantee their similar hydrological behaviour. Therefore, the physical similarity approach that compares the closeness of catchments on the basis of their characteristics (catchment descriptors) seems to be the most reasonable regionalisation approach.

In the physical similarity approach, a complete parameter set can be transferred from one gauged catchment with the catchment descriptors closest to the ungauged catchment (single donor approach) or, alternatively, the parameter set for the ungauged catchment can be estimated using parameter sets of a few most similar gauged catchments (multiple donor approach) (e.g. Burn & Boorman 1993; Merz & Blöschl 2004; Parajka *et al*. 2005; Oudin *et al*. 2008). The reason for the selection of catchments with the most similar catchment descriptors is based on the assumption that the hydrological behaviour of such catchments should be very similar to that of the ungauged catchment.

Various methods based on the physical similarity approach, ranging from relatively simple to complex, were described and tested in the literature (e.g. Burn 1990; Burn & Boorman 1993; Zrinji & Burn 1994; Parajka *et al*. 2005). All these methods include three steps: (1) selection of catchment descriptors, (2) identification of homogeneous hydrological regions (multiple donor approach) or identification of single most similar gauged catchments (simple donor approach), (3) estimation of chosen hydrological characteristics for the ungauged catchments (e.g. Nathan & McMahon 1990; Burn & Boorman 1993).

Two basic catchment descriptor groups can be used in the physical similarity approach – hydrological descriptors (i.e. the descriptors derived on the basis of rainfall-runoff records) on the one hand and physiographic and climatic descriptors on the other hand. Burn (1990), Zrinji and Burn (1994) and Yadav *et al*. (2007) recommend the application of hydrological descriptors, as these descriptors may describe the hydrological behaviour of a catchment much more efficiently. However, derivation of the hydrological descriptors in ungauged catchments is impossible. Therefore, physiographic and climatic descriptors are frequently used. Important groups of such descriptors are soil descriptors (e.g. Wagener *et al*. 2004), land cover descriptors, climatic descriptors (Young 2006) and morphological descriptors (e.g. Parajka *et al*. 2005; Laaha & Blöschl 2006; Oudin *et al*. 2008).

A set of catchment descriptors (and their optimal number) is often selected using the trial-and-error method, considering their various combinations (e.g. in Zrinji & Burn 1994; Parajka *et al*. 2005; Laaha & Blöschl 2006; Oudin *et al*. 2008). However, the choice of optimal catchment descriptors depends on the catchment set investigated and the hydrological characteristics to be estimated (Oudin *et al*. 2008). Nathan and McMahon (1990) suggested an alternative approach. They selected catchment descriptors on the basis of stepwise regression. The catchment descriptors which occurred most frequently in regression equations were used in the catchment grouping process.

This paper proposes an improved method for optimal selection of physiographic and climatic descriptors. The method contains an algorithm for progressive hierarchical delimitation of hydrologically homogeneous regions and an inverse clustering method for the catchment grouping process.

The aim of this paper is to provide an alternative to the trial-and-error method of selecting optimal catchment descriptors.

In the next chapters, a method for the catchment grouping process and an algorithm for progressive hierarchical delimitation of hydrologically homogeneous regions are presented and their advantages and disadvantages are discussed. In the second paper of this series, the newly proposed method will be illustrated on the example of identifying optimal catchment descriptors for the SAC-SMA model parameters (Burnash 1995) for catchments from the MOPEX (Duan *et al*. 2006) catchment set.

MATERIAL AND METHODS

Catchment descriptors. The catchment descriptor is any measurable or estimable catchment property characterizing a given catchment. There are two main groups of catchment descriptors – the hydrological descriptors and the physiographic and climatic descriptors. Hydrological descriptors will not be considered in the following text, because their derivation in ungauged catchments is impossible. Therefore, the term catchment descriptor will only refer to physiographic and climatic descriptors.

Physiographic and climatic descriptors can be divided into five basic groups: (1) soil characteristics (e.g. soil hydraulic characteristics, such as saturated hydraulic conductivity, wilting point, porosity or soil type, used e.g. by Parajka *et al*. 2005 or Young 2006), (2) geological descriptors (e.g. areal fractions of certain rock types as used in Yokoo *et al*. 2001), (3) land cover descriptors (e.g. the fractions of forest cover and urban areas used by Xu 1999 or Oudin *et al*. 2008), (4) morphological descriptors (e.g. catchment area, mean slope and river network density as used in Parajka *et al*. 2005 or Wagener & Wheater 2006) and (5) climatic descriptors (e.g. mean annual precipitation, mean annual evaporation or aridity index used by Parajka *et al*. 2005; Young 2006 and Oudin *et al*. 2008).

Catchment descriptors can be determined directly on the basis of field measurements or specialized maps and databases (e.g. catchment area, land-cover descriptors). Some can also be derived indirectly from other characteristics; using predefined relationships, e.g. soil hydraulic characteristics from STATSGO (Clapp & Hornberger 1978; Cosby *et al*. 1984) or mean annual precipitation from PRISM (Daly *et al*. 1994) for MOPEX data. These indirect methods of obtaining catchment descriptors can obviously be affected by larger errors (Wagener *et al*. 2004).

Inverse clustering. The inverse clustering method is based on finding hydrologically homogeneous regions around ungauged catchments. The catchment descriptors (CD) of an ungauged catchment define the location of a region centroid in the catchment descriptor space. The total number of such regions is given by the number of the tested ungauged catchments. Some of the regions may overlap, because CD values of several ungauged catchments may be similar to each other. In this context, the inverse clustering method is derived from the region of influence (ROI) method presented by Burn (1989) in relation to the regional flood frequency analysis. The gauged catchments are assigned to particular clusters using their Euclidean distance from particular ungauged catchments in the catchment descriptor space:

$$d_{UG} = \left[\sum_{A=1}^{N} \left(S_A^U - S_A^G \right)^2 \right]^{\frac{1}{2}} \quad (1)$$

where:

d_{UG} – Euclidean distance between the ungauged catchment U in the region centroid and the gauged catchment G assigned

S_A^U or S_A^G – standardised CD value of the descriptor A for the ungauged catchment U (or the gauged catchment G)

The standardisation of CDs is performed according to the equation:

$$S_A = \frac{X_A - \overline{X}}{s} \quad (2)$$

where:

S_A – value of a standardised catchment descriptor A

X_A – value of a non-standardised catchment descriptor A

$\overline{X}, s$ – the mean and the standard deviation for the catchment descriptor A calculated for all catchments in the data set, respectively

CD values of ungauged catchments are standardised together with the CD values of gauged catchments.

Andrews' curves are used in this paper for assessing the regions' homogeneity. The same homogeneity checking procedure was used by

Laaha and Blöschl (2006) and Nathan and McMahon (1990). These curves are derived on the basis of CDs of particular (either ungauged or gauged) catchments. Andrews' curves are points (representing catchments) with coordinates equal to the standardised catchment descriptors values $x = [S_1, S_2, S_3, ..., S_n]$ in a multi-dimensional space of functions. Each Andrews' curves is defined as a function of the argument b in the following form (Andrews 1972):

$$f(b) = \frac{S_1}{\sqrt{2}} + S_2 \times \sin(b) + S_3 \times \cos(b) + S_4 \times \sin(2b) + \cdots \quad (3)$$

where:

the argument varies within the range $-\pi \le b \le \pi$

$S_1, S_2, S_3, ..., S_n$ – standardised CD values for a particular catchment

The use of Andrews' curves relies on the orthogonality of terms in the Fouries series. The standardised CD value S_1 appears in the constant term of Eq. (3), S_2 and S_3 are associated with the lowest-frequency terms; S_4 and S_5 are associated with the double frequency terms and so on. The distance between two such curves, expressed as an integral of squared differences over the range $-\pi \le b \le \pi$, is proportional to the square of the Euclidean distance metric Eq. (1). Thus, the multi-dimensional points which are located close to each other in the Euclidean space of CDs will yield similar curves. Catchments with Andrews' curves located close to each other (and therefore catchments with similar values of the catchment descriptors) are supposed to have a similar hydrological behaviour. Using Andrews' curves for homogeneity checking is equivalent to using the Euclidean distance.

The proximity of Andrews' curves of the ungauged catchments to those of the assigned gauged catchments can be evaluated using the integrals of squares of their distances as mentioned above. As an alternative, we used the following coefficient of determination:

$$r^2 = 1 - \frac{\sum_{b=1}^{B} \left(f_U(b) - f_G(b)\right)^2}{\sum_{b=1}^{B} \left(f_U(b) - \bar{f}_U\right)^2} \quad (4)$$

where:

$f_U(b)$ – functional value of an Andrews' curve of the ungauged catchment U at the point b

$f_G(b)$ – functional value of an Andrews' curve of the gauged catchment G at the same point b

$\bar{f}_U$ – mean of Andrews' curve values of the ungauged catchment U over the interval ‹–π; π›, which is always equal to $S_1/\sqrt{2}$, and B is the total number of points

The values of $f(b)$ for the ungauged catchment ($f_U(b)$) and for the gauged catchment ($f_G(b)$) were calculated in this study in the interval ‹–π; π› with a step π/500, i.e. B = 1001.

The values of coefficient of determination calculated according to Eq. (4) fall within the interval ‹–∞; 1› and depend on the number of CDs used. Higher r^2 means a higher similarity between the catchments. For a perfect match of two Andrews' curves, the value of coefficient of determination is equal to 1. In this case, catchments have the same values of CDs.

The gauged catchments whose Andrews' curves are close to the Andrews' curve of a particular ungauged catchment make a hydrologically homogeneous region assigned to this ungauged catchment. The degree of similarity between the Andrews' curve of a gauged catchment and the ungauged catchment in question is expressed by the coefficient of determination according to Eq. (4). It is then compared with a predefined threshold value of the coefficient of determination (r_t^2). This threshold value determines the size and boundaries of the hydrologically homogeneous regions of gauged catchments around particular ungauged catchments.

Model parameters of the ungauged catchment are then estimated as weighted averages of the parameters of the assigned gauged catchments:

$$\theta_i^U = \frac{\sum_{p=1}^{N} \left(w_p \times \theta_{ip}^G\right)}{\sum_{p=1}^{N} w_p} \quad (5)$$

where:

θ_i^U – the i-th model parameter on the ungauged catchment

θ_{ip}^G – the known i-th model parameter on the p-th assigned gauged catchment

w_p – weight applied to the p-th gauged catchment

The inverse Euclidean distances defined by Eq. (1) are used as weights.

Algorithm for selection of optimal catchment descriptors. In this section, an algorithm for selection of an optimal set of catchment descriptors will be presented. It is related to the influence of the tested CDs on the catchment grouping process and, thereby, on the estimation of model parameters in

the ungauged catchments. As a result, we obtain a set of optimal CDs that can be used in the search for a set of gauged catchments that are most similar to a particular ungauged catchment.

As a first step, the available catchment descriptors are grouped into categories (e.g. climatic descriptors, geological descriptors, soil properties and soil types, land-cover descriptors and morphological descriptors). In the next step, a decision about the hierarchy of CDs must be made, regarding their importance for predicting model parameters of ungauged catchments. This is done by ranking their categories and then by ranking individual CDs within particular CD categories. This decision represents a subjective input to the algorithm. It may be based e.g. on the assumption that CDs from superior categories influence the CDs from inferior categories (e.g. the land-cover CDs will probably be influenced by climatic descriptors and soil properties).

The core of the algorithm relies on the existence of a training set of catchments such that for each of them the optimum model parameters are known. Then we explore how well these parameters can be re-estimated according to Eq. (5) for a particular catchment (regarded as ungauged for a moment) from the parameters of similar catchments within the same training set.

Two criteria are used in parallel for selection of an optimal group of CDs. The *MIN* (minimum error) criterion evaluates the accuracy of estimation of all parameters in all ungauged catchments for a given set of *n* selected CDs:

$$MIN = \sum_{i=1}^{N} w_i \left[\left(\frac{\tilde{D}(\theta_i)}{\min(\tilde{D}(\theta_i))} \right) \times 100 - 100 \right] \quad (6)$$

where:

$\tilde{D}(\theta_i)$ – median of absolute values of relative deviations (in %) of the re-estimated model parameter θ_i from its optimal value $\theta_{opt,i}$

$\min(\tilde{D}(\theta_i))$ – minimum value obtained among all values $\tilde{D}(\theta_i)$ for the model parameter θ_i

w_i – weight applied to the model parameter θ_i

N – total number of model parameters

The median is taken over all possible ungauged catchments for a particular set of *n* CDs and the minimum among such medians is taken over all possible sets of *n* CDs (with the constraints explained below). The weights in Eq. (6), assigned to parameters, are different from the weights in Eq. (5), assigned to catchments. The choice of weights in Eq. (6) is specific for each study. *MIN* values for all tested sets of *n* catchment descriptors are then compared. The set of *n* CDs with the lowest *MIN* value is considered the best. In practice this usually means that an optimum set of (*n*–1) CDs have already been selected in the previous step of the procedure and now various additional CDs are added to them, one or another, to find which of them will reduce the value of *MIN* most.

MAX (maximum improvement) is the second criterion. This criterion evaluates the improvement in estimation of a model parameter after the addition of an additional CD:

$$MAX = \sum_{i=1}^{N} w_i \left(\tilde{D}^{n-1}(\theta_i) - \tilde{D}^{n}(\theta_i) \right) \quad (7)$$

where:

$\tilde{D}^{n-1}(\theta_i)$ – median of absolute values of relative deviations (in %) of the estimated model parameter θ_i from the optimal parameter $\theta_{opt,i}$, where the estimates are based on an optimum set of (*n*–1) CDs

$\tilde{D}^{n}(\theta_i)$ – similar median of absolute values of relative deviations (in %) of the estimated model parameter θ_i from the optimal parameter $\theta_{opt,i}$, where the estimates are based on a set of *n* CDs made by adding an additional CD to the set that produced $\tilde{D}^{n-1}(\theta_i)$

N – total number of model parameters

w_i – weight applied to each model parameter θ_i (the weights in Eq. (7) are the same as the weights in Eq. (6))

The medians are taken over all possible ungauged catchments. The maximum is taken over all possible additional CDs (with the constraints explained below). *MAX* values for various CDs added to the existing set of (*n*–1) CDs are compared. The set of *n* CDs with the highest *MAX* value is considered optimal. In the case where two different sets of *n* CDs are indicated as the best, one according to its smallest *MIN* value and the other according to its highest *MAX* value,, the set with the highest *MAX* value is considered optimal, i.e. the criterion *MAX* is regarded more relevant.

The proposed algorithm can be summarised as follows:

(1) Creating all possible pairs from CDs of the first (superior) category and their testing, i.e., for each such pair, forming regions of similar catchments around each ungauged catchment, estimation of the ungauged catchment's model parameters in every region, calculation

of deviations between the estimated and the optimal parameters and computing *MIN*. The *MIN* values obtained in this way for all pairs of CDs are compared and the optimal pair of CDs is selected as the one that yields smallest differences between the estimated model parameters and the optimal ones.

(2) Creating all possible groups of three CDs composed of the optimal pair identified in step 1 and any remaining CD of the first category or the second (inferior) category, their testing (following the same procedure as in step 1 but, in addition, calculating also the *MAX* values for each group of three CDs) and the selection of the optimal group of three CDs on the basis of their *MIN* and *MAX* values.

(3) Creating all possible groups of four CDs, composed of the optimal group of three CDs identified in step 2 and any remaining CD of the first category or the second category (inferior to the first category) or the third category (inferior to the second category), their testing (the same procedure as in step 2) and selection of the optimal group of four CDs on the basis of their *MIN* and *MAX* values.

(4) Following the same line and looking for optimal groups of five, six etc. CDs, composed of the optimal CDs found in previous steps to which further CDs of the same categories or of other categories of lower superiority are added.

The procedure is repeated as long as the added catchment descriptors go on refining the model parameters estimates. This is indicated by positive *MAX* values. The algorithm starts with the groups of two CDs rather than with single CDs, because Andrews' curves are used for homogeneity checking within each region and the coefficient of determination, which indicates the nearness of particular Andrews' curves, could not be calculated, if only single CDs were used; the Andrews' curves would in this case become lines parallel to the x-axis and the denominator of the second term in the coefficient of determination would be zero.

Two auxiliary rules are applied in the algorithm described above. The first one excludes any further CDs of the superior category from consideration as soon as a CD of any inferior category has been identified as belonging to the optimal set. This condition relies on the idea that the previously selected CDs of the superior category may already contain, at that moment, virtually all relevant information inherent to this category and, therefore, adding any further CDs of this category would be useless. The second rule prefers a CD of an inferior category to a CD of a superior category, if both give very similar results (in terms of *MIN* and *MAX* values) and if the group of previously selected $n-1$ optimal CDs already contains at least two CDs of the superior category. This condition is substantiated by the same reasoning as the first one and, in addition, by the requirement that the set of optimal CDs should not be unnecessarily large and the search for a region of gauged catchments similar to the ungauged one should not be unnecessarily complicated.

It is worth mentioning that the roles of the catchment descriptors and the model parameters are interchangeable. These two categories of "parameters" (in a more general sense of the word) only differ in their availability or unavailability for ungauged catchments. A second difference is that the number of model parameters (for a particular model) is constant, while the number of catchment descriptors is variable. However, not all model parameters must necessarily be subject of estimation.

DISCUSSION

The proposed method of identification of optimal CDs has some advantages but also some disadvantages. The first advantage is the reduction of the total number of CD groups tested during the identification process in comparison with the testing of all CD combinations in the trial-and-error method (e.g. Parajka *et al*. 2005 or Oudin *et al*. 2008).

The next advantage of the proposed method consists in obtaining, as interim results, some information about the positive or negative influence of every added CD on a particular model parameter or a group of parameters. The positive influence means a more accurate estimate of the parameter or a group of parameters after adding a CD to the previously identified set of optimal CDs. It is expressed as the positive partial *MAX* value (obtained from an equation similar to Eq. (7) but taken separately for a particular parameter) or the positive partial *MAX* value for a group of parameters (if the medians are taken over all parameters of that group only). The negative influence of adding a CD to a previously identified set of optimal CDs is indicated by the increase of $\tilde{D}^n(\theta_i)$ for a particular

parameter, which leads to a negative partial *MAX* value for this parameter. A similar negative influence can be observed for a group of parameters (or, in the extreme, for all parameters), too. This information may indicate internal relationships among CDs and model parameters. The question is whether this information may be also useful for estimation of parameters of other models with a similar structure, using the same or a similar set of catchments. It is well-known that regionalisation results also depend on the structure of the model used (Kay *et al*. 2006). Therefore, the transfer of information regarding internal relationships between the CDs and the model parameters among models of similar structure could be possible.

The first disadvantage relates to the fact that the identified group of optimal CDs may not be truly optimal. The groups of CDs identified as the best may be different from those found by the trial-and-error method in which all combinations of CDs are tested.

The second disadvantage relates to the initial phase of the algorithm, when a decision about the hierarchy of CD categories must be made. It is a highly subjective decision and is also affected by the availability or unavailability of particular CD categories. If a training set is available with all CDs and all model parameters already known, it may principally be possible to identify the most relevant CDs by a sort of principal direction analysis. Until this is done, the subjectivity may affect the selection of optimal CDs and thus also the final estimates of model parameters in ungauged catchments. A suitable categorisation of CDs for large and very heterogeneous catchment sets could be based e.g. on the effect of human activities. The superior categories of CDs could be those that are least influenced by human activities (e.g. climatic descriptors and geological descriptors), while the categories of CDs more influenced by human activities (e.g. land-cover descriptors) could possibly be regarded as inferior.

The above-mentioned auxiliary rules applied during the algorithm execution may also affect the choice of optimal CDs. These rules may unfavourably affect the choice of optimal CDs because some (not yet considered) CDs from the superior category may contain further information important for accurate estimation of model parameters. On the other hand, the fact that a CD from the inferior category has been identified as optimal in a previous step may indeed indicate that little unused information is contained in the remaining CDs of the superior category.

Acknowledgements. Supported by the Ministry of Interior of the Czech Republic, Project No. BV II/2-VS. The authors would like to thank reviewers for their constructive suggestions, which helped them to improve the text.

References

Acreman A.C., Sinclayr C.D. (1986): Classification of drainage basins according to their physical characteristics; an application for flood frequency analysis in Scotland. Journal of Hydrology, **84**: 365–380.

Andrews D.F. (1972): Plots of high-dimensional data. Biometrics, **28**: 125–136.

Burn D.H. (1989): Cluster analysis as applied to regional flood frequency. Journal of Water Resources Planning and Management, **115**: 567–582.

Burn D.H. (1990): Evaluation of regional flood frequency analysis with a region of influence approach. Water Resources Research, **26**: 2257–2265.

Burn D.H., Boorman D.B. (1993): Estimation of hydrological parameters at ungauged catchments. Journal of Hydrology, **143**: 311–366.

Burnash R.J.C. (1995): The NWS river forecast system – catchment modelling. In: Singh V.P. (ed.): Computer Models of Watershed Hydrology. Water Resources Publication, Highlands Ranch, 69–118.

Clapp R.B., Hornberger G.M. (1978): Empirical equations for some soil hydraulic properties. Water Resources Research, **14**: 601–604.

Clarke R.T. (1973): A review of some mathematical models used in hydrology, with observations on their calibration and use. Journal of Hydrology, **19**: 1–20.

Cosby B.J., Hornberger G.M., Clapp R.B., Ginn T.R. (1984): A statistical relationship of soil moisture characteristics to the physical properties of soils. Water Resources Research, **20**: 682–690.

Daly C., Neilson R.P., Phillips D.L. (1994): A statistical-topographic model for mapping climatological precipitation over mountainous terrain. Journal of Applied Meteorology, **33**: 140–158.

Dingman L.S. (2002): Physical Hydrology. Prentice Hall, New Jersey.

Duan Q., Schaake J., Andréassian V., Franks S., Goteti G., Gupta H.V., Gusev Y.M., Habets F., Hall A., Hay L., Hogue T., Huang M., Leavesley G., Liang X., Nasonova O.N., Noilhan J., Oudin L., Sorooshian S., Wagener T., Wood E.F. (2006): Model Parameter Estimation Experiment (MOPEX): An overview of sci-

ence strategy and major results from the second and third workshops. Journal of Hydrology, **320**: 3–17.

Jarboe J.E., Haan C.T. (1974): Calibrating a water yield model for small ungauged watersheds. Water Resources Research, **10**: 256–262.

Kay A.L., Jones D.A., Crooks S.M., Calver A., Reynard N.S. (2006): A comparison of three approaches to spatial generalization of rainfall–runoff models. Hydrological Processes, **20**: 3953–3973.

Laaha G., Bloschl G. (2006): A comparison of low flow regionalisation methods – catchment grouping. Journal of Hydrology, **323**: 193–214.

Magette W.L., Shanholtz V.O., Carr J.C. (1976): Estimating selected parameters for the Kentucky Watershed Model from watershed characteristics. Water Resources Research, **12**: 472–476.

Merz R., Bloschl G. (2004): Regionalisation of catchment model parameters. Journal of Hydrology, **287**: 95–123.

Nathan R.J., McMahon T.A. (1990): Identification of homogeneous regions for the purposes of regionalisation. Journal of Hydrology, **121**: 217–238.

Oudin L., Andréassian V., Perrin C., Michel C., Le Moine M. (2008): Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments. Water Resources Research, **44**, W03413, doi:10.1029/2007WR006240.

Parajka J., Merz R., Bloschl G. (2005): A comparison of regionalisation methods for catchment model parameters. Hydrology and Earth System Sciences, **9**: 157–171.

Sivapalan M., Takeuschi K., Franks S., Gupta V.K., Karambiri H., Lakshmi V., Liang X., McDonnel J.J., Mendiondo E.M., O'Connel P.E., Oki T., Pomeroy J.W., Schertzer D., Uhlenbrook S., Zehe E. (2003): IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences. Hydrological Sciences, **48**: 857–880.

Vandewiele G.L., Ellias A. (1994): Monthly water balance of ungauged catchments obtained by geographical regionalization. Journal of Hydrology, **170**: 277–291.

Wagener T., Wheater H.S. (2006): Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty. Journal of Hydrology, **320**: 132–154.

Wagener T., Wheater H.S., Gupta H.V. (2004): Rainfall-runoff Modelling in Gauged and Ungauged Catchments. Imperial College Press, London.

Xu C.Y. (1999): Estimation of parameters of a conceptual water balance model for ungauged catchments. Water Resources Management, **13**: 353–368.

Yadav M., Wagener T., Gupta H. (2007): Regionalization of constraints on expected watershed response behaviour for improved predictions in ungauged basins. Advances in Water Resources, **30**: 1756–1774.

Yokoo Y., Kazama S., Sawamoto M., Nishimura H. (2001): Regionalization of lumped water balance model parameters based on multiple regression. Journal of Hydrology, **246**: 209–222.

Young A.R. (2006): Stream flow simulation within UK ungauged catchments using a daily rainfall-runoff model. Journal of Hydrology, **320**: 155–172.

Zhang Y., Chiew H.S.F. (2009): Relative merits of different methods for runoff predictions in ungauged catchments. Water Resources Research, **45**: W07412, doi:10.1029/2008WR007504.

Zrinji Z., Burn D.H. (1994): Flood frequency analysis for ungauged sites using a region of influence approach. Journal of Hydrology, **153**: 1–21.

Received for publication April 25, 2012
Accepted after corrections May 23, 2013

Corresponding author:

Ing. Martin Heřmanovský, Česká zemědělská univerzita v Praze, Fakulta životního prostředí, Kamýcká 129, 165 21 Praha 6-Suchdol, Česká republika
e-mail: hermanovsky@fzp.czu.cz