Seasonality indices for regionalizing low flows

G. Laaha¹* and G. Blöschl²

¹ Institut für Angewandte Statistik, Universität für Bodenkultur Wien, Gregor Mendel Str. 33, A-1180 Vienna, Austria
 ² Institut für Hydraulik, Gewässerkunde und Wasserwirtschaft, Technische Universität Wien, Karlsplatz 13/223, A-1040 Vienna, Austria

Abstract:

In this study we examine three seasonality indices for their potential in regionalizing low flows. The indices are seasonality histograms (SHs) that represent the monthly distribution of low flows, a cyclic seasonality index (SI) that represents the average timing of low flows within a year, and the seasonality ratio (SR), which is the ratio of summer and winter low flows. The rationale of examining these indices is the recognition that summer and winter low flows are subject to important differences in the underlying hydrological processes. We analyse specific low flow discharges q_{95} , i.e. the specific discharge that is exceeded on 95% of all days at a particular site. Data from 325 subcatchments in Austria, ranging in catchment area from 7 to 963 km², are used in the analysis. In a first step, three seasonality indices are compared. Their spatial patterns can be interpreted well on hydrological grounds. In a second step, the indices are used to classify the catchments into two, three, and eight regions based on different combinations of the indices. In a third step, the value of the seasonality indices for low flow regionalization is examined by comparing the crossvalidation performance of multiple regressions between low flows and catchment characteristics. The regressions make use of the three seasonality-based classifications. The results indicate that grouping the study area into two regions and three regions and separate regressions in each region gives the best performance. A global regression model yields the lowest performance and a global regression model that uses different calibration coefficients in each of the eight regions only performs slightly better. This suggests that separate regression models in each of the regions are to be preferred over a global model in order to represent differences in the way catchment characteristics are related to low flows. Copyright © 2006 John Wiley & Sons, Ltd.

KEY WORDS low flows; regionalization; regional regression; classification; cluster analysis; seasonality index; cross validation; prediction of ungauged basins

INTRODUCTION

Many branches of water resources management need accurate estimates of low flows. If suitable measurements are not available, then the low flow characteristics need to be estimated from regional information by some sort of hydrological regionalization technique. A classification of possible approaches is given in Smakhtin (2001). Regional regression is probably the most widely used technique in low flow estimation at ungauged sites (e.g. Vogel and Kroll, 1992; Dingman and Lawlor, 1995; Schreiber and Demuth, 1997). Examples also include the development of national low flow estimation procedures for the UK (Institute of Hydrology, 1980; Gustard *et al.*, 1992) and for Switzerland (Aschwanden and Kan, 1999). The models usually consist of regression relationships between some characteristic low flow discharge and physical catchment characteristics. Process understanding can be introduced in the models in a number of ways. One frequently used approach to introduce process understanding is to fit separate regression models to hydrologically homogeneous subregions. Nathan and McMahon (1990) compared several multivariate statistical approaches based on physical catchment characteristics to obtain possible groupings of hydrologically similar stations that can serve as a basis for fitting separate regionalization models to data. However, they stated that '… groupings obtained are very sensitive to the initial choice of predictor variables', and hence are highly subjective.

^{*} Correspondence to: G. Laaha, Institut für Angewandte Statistik, Universität für Bodenkultur Wien, Gregor Mendel Str. 33, A-1180 Vienna, Austria. E-mail: gregor.laaha@boku.ac.at

G. LAAHA AND G. BLÖSCHL

Seasonality has attracted a lot of attention in the literature recently to assist in the regionalization of hydrological quantities. Burn (1997) suggested a method that uses the seasonality of flood response as the basis for a similarity measure within the region of influence approach to flood regionalization. The regionalization technique was applied to a set of catchments from the Canadian prairies and was shown to be effective in estimating extreme flow quantiles. Merz et al. (1999) and Piock-Ellena et al. (2000) have illustrated that the seasonality approach is indeed useful in the context of flood frequency regionalization in Austria. They used a cluster analysis based on circular statistics of flood occurrence within the year to identify homogeneous regions and plotted vector maps to visualize the spatial patterns of the seasonalities of floods and other hydrological variables. The interpretation of these seasonality patterns led to an assessment of the main climate-driven flood-producing processes in Austria. Seasonality appears to be a useful indicator of catchment similarity in terms of hydrological processes, and we believe that the analysis of low flow seasonality should be useful for low flow regionalization. An application of a low flow seasonality index (SI) in the UK (Young et al., 2000) suggested that, if the spatial variability of low flow seasonality was rather weak, there is little discriminatory power in this index. It is clear that the usefulness of this method hinges on the existence of clear spatial patterns in low flow seasonality. Laaha (2002) compared two seasonality measures for low flows monitored at 57 stream gauges in Upper Austria and found that both measures were capable of classifying catchments into summer and winter low-flow-dominated subregions.

The natural factors that influence the various aspects of the low-flow regime of the river include the infiltration characteristics of soils, the hydraulic characteristics and extent of the aquifers, the rate, frequency and amount of recharge, the evapotranspiration rates from the basin, distribution of vegetation types, topography and climate. These factors and processes may be grouped into those affecting gains and losses of streamflow during the dry season of the year (Smakhtin, 2001). In highly seasonal climates, such as an alpine climate, low flows in different dry seasons (summer and winter) may be generated by different processes, and rivers will have two distinct low-flow seasons in winter and summer, controlled by different processes. In Austria, summer low flows occur during long-term persistent dry periods when evaporation exceeds precipitation. The consequence is a slow depletion of the soil reservoir in accordance with the recession of discharges. Important low flow generating factors are the distribution of precipitation during the summer season and the storage properties of soil. Winter low flows are affected by freezing processes. Persistent frost leads to the storage of precipitation in the snow cover and to ice formation in the topsoil. Thus, catchment altitude (which is highly correlated with temperature) and aquifer thickness (which affects the fraction of retarded water, as well as the recession of stream flow) seem to be important factors of winter low flows. Because of the fundamental differences of summer and winter processes, regionalization may take advantage of a separation of summer and winter low flows (Tallaksen and Hisdal, 1997; Laaha, 2000a). For the same reasons, seasonality is also potentially useful for regionalizing annual low flows. There are different ways of incorporating seasonality in regionalization models, e.g. by fitting separate models for homogeneous groups, or by adjusting the model to different group means of the low flow characteristic by separate coefficients.

Examples for seasonality analysis in the context of low flow regionalization are, however, rare. Schreiber and Demuth (1997) analysed seasonality of mean annual 10-day minimum MAM(10) of total discharges measured in 169 catchments in southwest Germany. Average occurrence of MAM(10) per month was determined for 10 regions and for the whole study area. The results indicated typical low flow occurrence from September to October for large parts of the study area, apart from the Pre-Alps (Voralpen region), which are dominated by winter low flows (January and February). The differences of low flow seasonality were found to depend mainly on catchment altitude. Aschwanden and Kan (1999) investigated the long-term characteristic seasonal distribution of Q_{95} for representative gauges from 143 headwater catchments in Switzerland, based on the 1935–96 observation period. They found two different typical seasonal distributions of low flows, again depending on catchment altitude. In alpine catchments, low flows occur exclusively from November to March. In the hilly landscapes of Mittelland and Jura, low flows may occur during the whole year, but clearly most frequently during summer and autumn. Dingman and Lawlor (1995) stated that, in the Vermont and New

Hampshire region, annual 7-day minimum flows usually occur in late summer or early fall in response to regional climatic patterns, but they occur in some years during late winter in the more northern and highelevation streams. The mean time of occurrence for annual 7-day minimum flows is in August for Vermont and the Connecticut River basin, in September in the Saco River basin and in August or September in the rest of New Hampshire, except at the highest elevations, where it occurs in February. However, none of these studies explicitly accounted for the seasonal heterogeneity in low flow regionalization. Possible benefits of approaches to include seasonality in the regionalization of low flows are unclear.

The aim of this paper is to investigate the value of seasonality indices for regionalizing low flows. As a regionalization model, we use stepwise multiple regressions based on physical catchment characteristics and seasonality indices. The value of different models that incorporate seasonality by different approaches is assessed by cross-validation, which emulates the prediction of low flows at ungauged catchments. We compare the models for the 95% quantile of specific discharges q_{95} and we also examine the specific low flow discharge of the summer and winter periods (q_{95s} , q_{95w}).

The paper is organized as follows. The next section summarizes the data and the disaggregation method used in this study for calculating specific low flow discharges for residual catchments. The third section presents different seasonality measures and shows how subregions of similar seasonality can be isolated. The value of these seasonality measures for regionalization is investigated in the fourth and fifth sections: the fourth section presents the method of regionalization and cross-validation used in this study and describes how seasonality measures have been considered in regression modelling, and the results are given in the fifth section. A discussion and conclusions then follow in the sixth and seventh sections respectively.

DATA

Study area

The study has been carried out in Austria, which is physiographically quite diverse. There are three main zones in terms of the landscape classification: high Alps in the west, lowlands in the east, and there is hilly terrain in the north (foothills of the Alps and Bohemian Massif) (Figure 1). Elevations range from 117 to 3798 m a.s.l. Geological formations vary significantly, too. Austria has a varied climate with mean annual



Figure 1. Topography and stream gauging network in Austria. Points indicate location of gauges used in this study

G. LAAHA AND G. BLÖSCHL

precipitation ranging from 500 mm in the eastern lowlands up to about 2800 mm in the western alpine regions. Runoff depths range from less than 50 mm per year in the eastern part of the country to about 2000 mm per year in the Alps. Potential evapotranspiration ranges from about 730 mm per year in the lowlands to about 200 mm per year in the high alpine regions. This diversity is reflected in a variety of hydrological regimes (Kresser, 1965) and low flows exhibit important regional differences in terms of their quantity and their seasonal occurrence (Laaha and Blöschl, 2003).

Discharge data

Discharge data used in this study are daily discharge series from 325 stream gauges. These data represent a complete set of gauges for which discharges have been continuously monitored from 1977 to 1996 and where hydrographs have not been seriously affected by abstractions and karst effects during low flow periods (Laaha and Blöschl, 2003). Catchments for which a significant part of the catchment area lies outside Austria have not been included, as no full set of physiographic data was available for them. The catchments used here cover a total area of 49 404 km², which is about 60% of the national territory of Austria. Although a larger number of catchments are monitored in Austria, we have chosen to give priority to a consistent observation period to make all records comparable in terms of climatic variability.

Disaggregation of nested catchments

Nested catchments were split into subcatchments between subsequent stream gauges based on the hierarchical ordering of gauges presented in Laaha and Blöschl (2003). The advantage of using subcatchments rather than complete catchments is that the application of regionalization techniques to small ungauged catchments is more straightforward. Also, discharge characteristics of nested catchments are statistically not independent and disaggregation into subcatchments between subsequent stream gauges makes them more independent. The disadvantage of the disaggregation is that errors may be somewhat larger, as the low flow characteristics are estimated from differences of the stream flow records at two gauges. If the errors of the upstream and downstream gauges are assumed normally distributed and independent, then the error variances are additive. A standard error of 3% (Laaha, 2000a and b) for the low flow characteristics. If the errors are not independent, then the errors would be slightly smaller. These errors are small compared with the regionalization errors to be expected (Laaha and Blöschl, 2005).

Low flow characteristics

Low flows were quantified by the Q_{95} flow quantile, $Pr(Q > Q_{95}) = 0.95$, i.e. the discharge that is exceeded on 95% of all days of the measurement period. This low flow characteristic is widely used in Europe and was chosen because of its relevance for multiple topics of water resources management (e.g. Kresser *et al.*, 1985; Gustard *et al.*, 1992; Smakhtin, 2001). For gauged catchments without an upstream gauge we calculated the Q_{95} low flow quantile directly from the stream flow data. For subcatchments we calculated Q_{95} from the differences of stream flows at the two gauges. To make the low flow characteristic more comparable across scales, we standardized Q_{95} by the catchment area. The resulting specific low flow discharges q_{95} ($1 \text{ s}^{-1} \text{ km}^{-2}$) were considered to be representative of the characteristic unit runoff from the catchment area during sustained dry periods.

A map of specific low flow discharge q_{95} in Austria is presented in Figure 2. The pattern of calculated low flow characteristics q_{95} appears rather smooth and homogeneous over geographically similar regions. The low flows are obviously related to terrain, since the alpine region shows higher values and stronger spatial variability. Here, typical values of q_{95} appear to range from 6 to 20 l s⁻¹ km⁻², whereas regions situated in the southern Alps indicate lower discharges because of drier climatic conditions. On the other hand, typical values of q_{95} for hilly terrain and the lowlands range from 0 to 8 l s⁻¹ km⁻².



Figure 2. Specific low flow discharge q_{95} (1 s⁻¹ km⁻²) from runoff data observed in 325 subcatchments in Austria. Alpine catchments show higher values and a larger variability



Figure 3. Ratio of summer and winter low flow discharges (SR) for 325 subcatchments in Austria. SR > 1 indicates a winter low flow regime; SR < 1 indicates a summer low flow regime

Catchment characteristics

We used 31 physiographic catchment characteristics in the low flow regionalization in this paper (Table I). They relate to catchment area A, topographic elevation H, topographic slope S, precipitation P, geology G, land use L, and drainage density D. All percentage values with the exception of mean slope S_M relate to the area covered by a class relative to the total catchment area. Some of the catchment characteristics

G. LAAHA AND G. BLÖSCHL

had to be adapted from the original sources to make them more useful for regionalization. For instance, the original classification of the metallurgic map used here distinguishes 670 geological classes, from which we derived nine hydrogeological classes we deemed relevant for low flow regionalization. One of them is termed source region, which is the percentage area where the density of springs is large. In a similar vein, we condensed the original Corine Landcover classification (Aubrecht, 1998) into nine land-use classes. The average stream density (i.e. length of a stream by unit area (m km^{-2})) of sub-basins was calculated from the stream density map of the Hydrological Atlas of Austria (Fürst, 2003), which is based on the digital drainage network of Austria at the 1:50000 scale (Behr, 1989). Because of its relationship with infiltration rates of different geological units (e.g. Grayson and Blöschl, 2002), this index may be a useful alternative to geological characteristics in low flow regionalization. Three precipitation characteristics of average annual, summer and winter precipitation from 1977 to 1996 estimated by the regionalization model of Lorenz and Skoda (1999) were used. A number of topographical characteristics were derived from a digital elevation model at a 250 m grid resolution. All characteristics were first compiled on a regular grid and then combined with the subcatchment boundaries of Laaha and Blöschl (2003) and Behr (1989) to obtain the characteristics for each catchment. A statistical summary of the catchment characteristics is given in Table I.

Variable	Variable description	Units	Min.	Mean	Max.
A	Subcatchment area	10^1 km ²	0.70	15.22	96.30
H_0	Altitude of stream gauge	10^{2} m	1.59	5.93	22.15
H_{+}	Maximum altitude	10 ² m	2.98	17.48	37.70
$H_{\rm R}$	Range of altitude	10^{2} m	0.81	11.56	30.06
$H_{\rm M}$	Mean altitude	10 ² m	2.32	10.53	29.45
S _M	Mean slope	%	2.70	24.34	56.00
S _{SL}	Slight slope	%	0.00	28.06	100.00
$S_{\rm MO}$	Moderate slope	%	0.00	46.18	93.00
S _{ST}	Steep slope	%	0.00	25.78	80.00
Ρ	Average annual precipitation	10^2 mm	4.67	10.71	21.03
$P_{\rm S}$	Average summer precipitation	10^2 mm	2.94	6.47	12.08
$P_{\rm W}$	Average winter precipitation	10^2 mm	1.55	4.24	8.95
$G_{\rm B}$	Bohemian Massif	%	0.00	9.70	100.00
G_0	Quaternary sediments	%	0.00	6.22	94.50
G_{T}	Tertiary sediments	%	0.00	15.91	100.00
$G_{ m F}$	Flysch	%	0.00	6.90	100.00
$G_{ m L}$	Limestone	%	0.00	25.21	100.00
$G_{\rm C}$	Crystalline rock	%	0.00	25.44	100.00
$G_{ m GS}$	Shallow groundwater table	%	0.00	1.74	48.00
$G_{ m GD}$	Deep groundwater table	%	0.00	7.51	79.80
$G_{\rm SO}$	Source region	%	0.00	1.23	35.20
$L_{\rm U}$	Urban	%	0.00	0.67	14.50
$L_{\rm A}$	Agriculture	%	0.00	21.37	97.30
L _C	Permanent crop	%	0.00	0.12	20.30
$L_{\rm G}$	Grassland	%	0.00	20.10	71.70
$L_{ m F}$	Forest	%	0.00	47.25	100.00
$L_{\rm R}$	Wasteland (rocks)	%	0.00	8.45	81.20
$L_{\rm WE}$	Wetland	%	0.00	0.10	16.40
$L_{ m WA}$	Water surfaces	%	0.00	0.42	18.20
$L_{\rm GL}$	Glacier	%	0.00	1.37	43.80
D	Stream network density	10^2 m km^{-2}	1.18	8.01	13.98

Table I. Statistical summary of the characteristics of the 325 subcatchments used in this paper. Units were chosen in a way to give similar ranges for all characteristics

SEASONALITY ANALYSIS

Seasonality measures

The seasonality ratio (SR). Summer and winter low flows are subject to important differences in the underlying hydrological processes. Thus, we expect that summer and winter low flows exhibit different spatial patterns caused by the variability of physical catchment properties. This topic can best be addressed by a separate mapping of summer and winter low flows. Daily discharge time-series have been stratified into summer discharge series (from 1 April to 30 November) and winter discharge series (1 December to 31 March). These dates were chosen to capture summer drought processes safely in the Austrian lowlands in the summer period, and frost and snow accumulation processes in alpine areas in the winter period. From winter and summer discharge time-series, characteristic values for summer low flows q_{95s} and winter low flows q_{95w} were calculated for each subcatchment. The SR of q_{95s} and q_{95w} was then calculated:

$$SR = q_{95s}/q_{95w}$$
 (1)

A map of SR for Austria is presented in Figure 3. Values of SR > 1 indicate the presence of a winter low flow regime and values of SR < 1 indicate the presence of a summer low flow regime. The map demonstrates a clear and ordered classification of low flow seasonality in Austria. Alpine regions are dominated by winter low flows, whereas lowlands and hilly terrain in the north and east of Austria are dominated by summer low flows. In between, a transition zone characterized by weak seasonality appears. The plot appears to be useful for visualizing the patterns of summer and winter low flows.

The SI. We use an index similar to Burn (1997) and Young *et al.* (2000) to represent the seasonal distribution of low flow occurrence. The index is based on two parameters, θ and r, which are calculated from the Julian dates of all days of the observation period when discharges are equal or below Q_{95} , by means of circular statistics (Mardia, 1972). The parameter θ is the mean day of occurrence, measured in radians, and is a measure of the average seasonality of low flows. The parameter θ takes values between 0 and 2π : $\theta = 0$ relates to 1 January, $\pi/2$ relates to 1 April, π relates to 1 July, and $3\pi/2$ relates to 1 October. The parameter r is the mean resultant of days of occurrence, which is a dimensionless measure of the variability of low flow seasonality. Possible values of r range from zero to unity, with r = 1 corresponding to strong seasonality (all low flow events occurred on exactly the same day of the year) and r = 0 corresponding to no seasonality (low flow events are uniformly distributed over the year).

For each subcatchment, the days on which discharge was smaller than Q_{95} were extracted over the period of record and transformed into Julian dates D_j (i.e. the day of the year ranging from 1 to 365 in ordinary years and 1 to 366 in leap years). D_j represents a cyclic variable that can be displayed as a vector on the unit circle. Its directional angle, in radians, is given by

$$\theta_j = \frac{D_j 2\pi}{365} \tag{2}$$

The arithmetic mean of Cartesian coordinates x_{θ} and y_{θ} of a total of n single days j is defined as

$$x_{\theta} = \frac{1}{n} \sum_{j} \cos(\theta_{j})$$

$$y_{\theta} = \frac{1}{n} \sum_{j} \sin(\theta_{j})$$
(3)

From this, the directional angle of the mean vector was derived by

$$\theta = \arctan\left(\frac{y_{\theta}}{x_{\theta}}\right) \quad 1 \text{ st and 4th quadrants } : x > 0 \tag{4}$$
$$\theta = \arctan\left(\frac{y_{\theta}}{x_{\theta}}\right) + \pi \quad 2 \text{nd and 3rd quadrants: } x < 0$$

The mean day of occurrence is obtained by back-transforming the mean angle to a Julian date:

$$D = \theta \frac{365}{2\pi} \tag{5}$$

The length r of the mean vector is a measure of the variability of low flow days:

$$r = \sqrt{x_{\theta}^2 + y_{\theta}^2} \tag{6}$$

Seasonality indices for each sub-basin were displayed by a vector map (Figure 4), which gives a synoptical representation of the mean day of occurrence and the intensity of seasonality for a large number of catchments. The vector map provides a nice overview of the regional patterns of low flow seasonality in Austria.

Seasonality histogram (SH). The SH (Laaha, 2002) allows a more detailed description of the seasonal distribution of low flows than the SI. Again, this description is based on the Julian date of all days when the discharge of a catchment (or the differential discharge of a subcatchment) falls below the threshold Q_{95} . Histograms based on monthly classes were plotted from these data. Hence, the SH illustrates the occurrence of low flows in each month and provides supplemental information to the SI. In particular, it illustrates which months are affected by low flows and it provides a good representation of the shape of the seasonal distribution, including multimodal and skewed distributions.



Figure 4. SI of 325 subcatchments in Austria. Long arrows indicate strong seasonality and their direction represents the mean day of occurrence of specific low flow discharges less than q_{95}

Delineation of homogeneous regions

Cluster analysis of SH. SHs consist of 12 variables representing the monthly occurrence frequency of low flows (Laaha, 2002). To delineate regions that are homogeneous in terms of seasonality, partitive cluster analysis (partitioning around medoids (PAM); see Kaufmann and Rousseeuw (1990)) was applied to classify SHs automatically. PAM is an exhaustive partitioning method by which the ensemble of catchments is classified into several exclusive subsets. The optimal cluster centres (medoids) were chosen automatically by the algorithm. The number of clusters was optimized by means of the silhouette plot, an ordered representation of the silhouette width (Kaufman and Rousseeuw, 1990) of each histogram, which gives a relative measure of the similarity of one histogram to its allocated cluster centre with respect to its similarity to the next best suitable cluster centre. The maximum average silhouette width among several classifications into different numbers of clusters is related to the optimum number of clusters. We compared partitions of two to eight clusters. The analysis led to an optimal number of two clusters.

The graphical representation of catchments by the first two principal components of SHs (Figure 5, left) indicates that the clusters correspond to two very distinct groups of catchments in terms of seasonality. The first principal component separates catchments into winter and summer types. The second principal component further distinguishes between the timing of low flows within the regime types: negative values correspond to occurrence near spring, and positive values correspond to occurrence in autumn. The overlap of clusters in autumn corresponds to a group of catchments that exhibit no clear summer or winter seasonality.

Two possible classifications of catchments have been derived from the cluster analysis of an SH. The first classification corresponds to the two clusters obtained by the cluster analysis by which catchments are classified into summer and winter regime types. The second classification further distinguishes a third group containing 33 catchments that exhibit mixed seasonality. These catchments were identified by using silhouette width < 0.2 as a criterion (Figure 5, right).

The location of summer- and winter-type catchments can be seen from Figure 6, indicating two contiguous regions of different seasonality. Winter low flows typically occur at higher altitudes in the Alps, and summer low flows typically occur in the lower parts of Austria. The alternative classification into three regime types



Figure 5. Left: graphical representation of cluster membership of catchments (points) by the first two principal components of SHs. The big ellipse contains catchments of the summer-type cluster; the smaller ellipse contains catchments of the winter-type cluster. Right: determination of catchments that exhibit weak or mixed low flow regimes by silhouette width, illustrated by the silhouette plot

is shown in Figure 7. Mixed seasonality typically appears in the transition zone from the high Alps to the foothills of the Alps. Both classifications are generally in accordance with the spatial pattern of the SR (Figure 3); but, instead of the gradual representation of seasonality by the SR, the cluster analysis results in a mutually exclusive classification of catchments. Cluster analysis of an SH appears to be an appropriate basis for regionalizing low flows separately for catchments that exhibit typical summer and winter regimes.

Visual grouping based on different seasonality measures. Based on an interpretation of the SI and SHs, regions of approximately homogeneous seasonality have been identified visually. This approach is more subjective than automatic classification, but allows us to take additional information into account, such as



Figure 6. Classification of 325 subcatchments in Austria into two regime types (summer regime and winter regime)



Figure 7. Classification of 325 subcatchments in Austria into three regime types (summer regime, winter regime, mixed regime)

breaklines of the relief. Moreover, hydrological expert knowledge may be introduced into the classification, e.g. in the interpretation of local anomalies and outliers. This is probably a major advantage over the cluster analysis. The visual grouping approach consists of two steps. In a first step, preliminary regions were detected by synoptical mapping of the SI. In a second step, close inspection of SHs led to a correction and refinement of the preliminary regions. Where boundaries of regions appeared unclear, the digital terrain model was inspected for close-by topographic breaklines to assist in the choice of the boundaries.

Figure 8 presents the seasonality regions so obtained, which correspond to the types of SH presented in Figure 9. Results indicate significant regional differences of low flow seasonalities in Austria. Two zones of clearly contrasting seasonalities exist. One zone represents winter-dominated low flows (seasonality types A-C), which is the alpine region from Vorarlberg to the Wechselregion with a north-south extent from the northern Calcerous Alps to Upper Carinthia. The intensity of seasonality and the mean day of occurrence vary with the elevation of the catchments. Catchments of type A (West-Styria) exhibit mean seasonalities in January, type B (Salzburg and Upper Carinthia) in February and type C (large parts of Tyrol) at the beginning of March. The other zone represents summer-dominated low flows (seasonality types 1-2) and comprises catchments north and east of the Alps (lowlands and hilly terrain with elevations from 117 to about 600 m; in the Mühlviertel region to about 1000 m). Similarly, the regions of type 3 (Innviertel) and type 4 (foothills of the Alps) are summer dominated, although this effect is less clear. The same is true of the regions of type D (Eastern Styria) and type E (northern part of Vorarlberg), which are winter dominated but also exhibit minor summer influences. Finally, Lower Carinthia (type 5) exhibits a very weak seasonality. This seems to be caused by the particular climate of this region. Overall, the classification corresponds well with the patterns of the SR and can be considered a refined classification compared with that obtained by cluster analysis. Since regions appear well interpretable in terms of low flow processes, there is likely some potential for regionalization in the approach.



Figure 8. Regions of approximately homogeneous seasonality in Austria. Letters refer to winter low flow types and numbers to summer low flow types (see Figure 9)



Figure 9. SHs: non-exceedance frequencies of Q_{95} for each month for a typical catchment in each region. Letters relate to winter low flows and numbers relate to summer low flows (see Figure 8)

METHOD OF REGIONALIZATION AND CROSS-VALIDATION

Multiple regression

The regionalization methods used in this study are multiple linear regression models between specific low flow discharge q_{95} and physical catchment characteristics. Physical catchment properties are represented by 31 catchment characteristics, a number that is relatively large compared with other regionalization studies reported in the literature. These catchment characteristics are subject to intercorrelations and multicollinearity, as mentioned above. Rather than performing a selection of the most important variables prior to regionalization, we used a stepwise regression approach. The stepwise regression procedure used Mallow's C_p (Weisberg, 1985: 216) as the criterion of optimality, which was calculated as

$$C_p = \frac{\text{RSS}_p}{\hat{\sigma}^2} + 2p - n \tag{7}$$

The first term is the residual sum of squares of one considered model (RSS_p) with p coefficients divided by the residual error variance $\hat{\sigma}^2$ of the full model and corresponds to the relative optimality in terms of model error. Complexity of models is penalized by the second term, which adds the number of coefficients p minus the number of catchments n. Therefore, C_p is a penalized selection criterion that takes the gain of explained variance as well as the parsimony of models into account and yields models that are optimal in terms of prediction errors. Variable selection starts with one arbitrarily chosen catchment characteristic and subsequently adds variables that minimize the C_p criterion. After each step, whether replacing one of the variables by any of the remaining catchment characteristics will further decrease the criterion is tested. The selection procedure continues until C_p reaches a minimum. The catchment characteristics obtained by the stepwise regression can hence be interpreted as important controls of low flows.

Fitting regression models is often complicated by single extreme values. Elimination of such outliers may apparently improve statistical measures of model quality, leading to overly optimistic results. On the other hand, extreme values may act as leverage points. The effect of such points is to force the fitted model close to the observed value of q_{95} , leading to a small residual for this point. Therefore, regression parameters and residual statistics may be strongly influenced by single extreme values and may not represent the bulk of data. Our approach to this problem is an iterative robustified regression technique. Initial models fitted by stepwise regression were checked for leverage points using Cook's distance (e.g. Weisberg, 1985). Catchments for which Cook's distance was large compared with the remaining catchments were regarded as possible leverage points. These catchments were left out and again stepwise regression was performed until no leverage points remained. Finally, residual diagnostics, including the root-mean-squared error and the coefficient of determination, were calculated for all data, including leverage points.

The regression models so obtained were checked for numerical stability of computation. Since numerical stability is sensitive to different scales of predictors, all catchment characteristics had been scaled by integer powers of 10 to give similar magnitudes in terms of their ranges (see Table I). Since linear regression is scale invariant (Weisberg, 1985: 185), the regression models, including their residual statistics, remain unaffected by the rescaling, but the numerical stability is improved.

Regionalization methods examined

*Regionalization of q*₉₅ *low flows. Global regression:* In a first approach, one global regression model was fitted to all 325 catchments, using the robustified stepwise regression technique. The global model does not account for seasonality; hence, it is a benchmark case against which to test the seasonality-based regionalization methods.

Grouping into two regions and separate regressions in each region: In the second approach, regionally restricted regression models were each fitted for contiguous regions consisting of summer-dominated and of winter-dominated catchments. This corresponds to the original classification of catchments obtained by the cluster analysis of SHs (Figure 6).

Grouping into three regions and separate regressions in each region: Similar to the second approach, regionally restricted regression models were separately fitted for three groups of catchments, corresponding to summer regime, winter regime and mixed seasonality. This grouping corresponds to the second classification of catchments obtained by the cluster analysis of SHs (Figure 7). As opposed to the classification into two regions, these regions are spatially discontiguous, and prediction of ungauged sites would require some decision rule based on data that are available at both gauged and ungauged sites.

Global regression with different Z parameters in eight regions: In the fourth approach, a global regression model is fitted to the data that explicitly represents group membership of catchments in one of the eight seasonality regions by a coefficient termed Z. The linear model so obtained (a generalization of the multiple regression model for numeric and factor variables) fits a separate coefficient (additive parameter Z) to each seasonality region. This coefficient accounts for differences in the average low flows between seasonality zones. This approach is more parsimonious than fitting separate linear regression models for each region, which may be an advantage if a large number of subregions is used. Regression parameters for catchment characteristics, however, are fitted globally and the model is, therefore, not suitable for non-linear relationships between low flows and catchment characteristics.

G. LAAHA AND G. BLÖSCHL

Regionalization of summer period (q_{95s}) and winter period (q_{95w}) low flows. Global regression: As an alternative approach to consider seasonality in low flow regionalization, specific low flows of the summer period (q_{95s}) and the winter period (q_{95w}) were fitted by two separate global regression models. Since summer and winter low flows are related to different processes, one would expect that representing them separately provides a more realistic representation of spatial low flow variability. Although it is not straightforward to derive annual low flows from the summer and winter low flows, we can expect further insights into the value of accounting for seasonality in the regionalization.

Grouping into two regions and separate regressions in each region: The last approach considered in this paper is a combination of spatial grouping into summer and winter regions (Figure 6) and the separate regionalization of the summer period and the winter period low flows. Models were separately fitted for summer and winter low flows and separately in the summer and winter low-flow-dominated regions, leading to four temporally and regionally restricted submodels. This approach was used to obtain a more precise separation of summer and winter processes than by any of the two underlying methods alone.

Cross-validation

The error of prediction at ungauged sites can be assessed by the average residual squared error. However, this will tend to be too optimistic, as the same data are used for assessing the model as to fit it, so parameter estimates may be fine-tuned to the particular data set. In order to get a more realistic estimate of prediction error, we used leave-one-out cross-validation. The cross-validation estimate of prediction error is given by

$$V_{\rm cv} = \frac{1}{n} \sum_{i=1}^{n} (\hat{q}_{95i}^{(-i)} - q_{95i})^2 \tag{8}$$

where *n* is the total number of catchments, q_{95i} is the observed specific low flow discharge q_{95} for catchment *i* and $\hat{q}_{95i}^{(-i)}$ is the model prediction without using observed low flows from catchment *i*. The root-mean-squared error based on cross-validation is therefore

$$rmse_{cv} = \sqrt{V_{cv}}$$
(9)

and the coefficient of determination based on cross-validation is

$$R_{\rm cv}^2 = \frac{V_q - V_{\rm cv}}{V_q} \tag{10}$$

where V_q is the spatial variance of the observed specific low flow discharges q_{95} . Note that the complete set of catchments, including leveraging points (see 'Multiple regression' section), are incorporated in the cross-validation with the exception of one or two regression outliers in case they were too far from the bulk of the data.

The advantage of cross-validation over other techniques of assessing predictive errors is its robustness and its general applicability to all regionalization models. This is because cross-validation works well even if the regionalization models are far from correct (Efron and Tibshirani, 1993). Cross-validation is hence a full emulation of the case of ungauged sites.

RESULTS

Examining model assumptions

The multiple regression approach is based on two main assumptions, unbiasedness ($E[res_i] = 0$) and homoscedasticity (Var[res_i] = constant), where res_i is the residual of catchment *i*. Normality of residuals



Figure 10. Scatter plots of predicted versus observed specific low flow discharges q_{95} (1 s⁻¹ km⁻²) in the cross-validation mode. Each panel corresponds to a regional regression model and each point corresponds to a catchment. Point markers L indicate leverage points (i.e. catchments that have been left out in model calibration)

is a desirable property if one is interested in interpretable estimates of model performance. In this study, model assumptions are carefully checked by scatter plots of observed versus predicted values, and histograms and normal probability plots of residuals.

Scatter plots of observed versus predicted specific low flow discharges q_{95} (1 s⁻¹ km⁻²) in the crossvalidation mode are presented in Figure 10. Each panel corresponds to one regional regression model and each point corresponds to one catchment. The scatter plots allow a detailed examination of the performance of individual catchments, including the existence of outliers and a potential heteroscedasticity of the observations and the predictions. For all models, the outliers tend to increase with q_{95} , which suggests that the predictions are heteroscedastic. One would usually apply a variance-stabilizing transformation in this case, such as taking logarithms of q_{95} . However, since preliminary analysis indicated little effect on the model parameters, the level of heteroscedasticity was considered acceptable in the context of this paper, as the main focus was on evaluating the potential of seasonality indices on low flow regionalization. The global regression model exhibits the widest scatter among all models. No extreme outliers appear. Grouping into two regions and separate regressions in each region exhibits a somewhat narrower scatter for the bulk of data. Model fitting was complicated by a larger number of leverage points, which clearly appear as outliers of prediction. Model fitting without leverage points obviously led to a stronger selectivity between well-represented catchments and outliers, which might correspond to typical and atypical catchment conditions. Grouping into three regions and separate regressions in each region appears similar to grouping into two regions, but leverage points appear as even stronger outliers. The global regression using different Z parameters in each of the eight regions appears to give a similar performance as the global model without Z parameters. One apparent deficiency of all models is the large scatter and clear bias for very wet catchments. It appears that none of the models can cope very well with these large discharges.

Normal probability plots of cross-validated residuals $(1 \text{ s}^{-1} \text{ km}^{-2})$ are presented in Figure 11. For all regionalization models, residuals appear only approximately normally distributed. Single extreme outliers appear (typically one or two per model). Since such outliers exert a strong influence on second-order statistics, such as sum of squared residuals, they will not be used in the calculation of performance measures (mean squared error, coefficient of determination) in order to represent the bulk of the catchments rather than outliers.

A more detailed assessment of Figure 11 yields that, for all models, only small residuals exhibiting absolute values less than $2 \ 1 \ s^{-1} \ km^{-2}$ (points between dashed lines, representing about two-thirds of all catchments) approximate a normal distribution well. Larger residuals, however, deviate from a normal distribution, and normal probability plots indicate heavy-tailedness, i.e. a higher probability of larger residuals than expected values of normal distribution. The deviation from the normal distribution is stronger for regional restricted



Figure 11. Normal probability plots of cross-validated residuals (l s⁻¹ km⁻²) of regionalization. Each panel corresponds to one regionalization model and each point corresponds to one catchment

models than for global models. This is probably due to the mixture of the residual distributions of submodels. One consequence of the different distributions of residuals from the different models is that standard measures of model performance, such as coefficient of determination or root-mean-squared error, are not exactly comparable between models, due to disproportional influences of large residuals. In addition to mean-squared-error residual statistics (rmse_{cv} and R_{cv}^2), we will, therefore, also assess the performance of models by a classification of catchments according to absolute values of cross-validated residuals. This measure of model quality appears less sensitive to the distribution of cross-validated residuals.

Relative importance of predictor variables

The regression model equations of the four resulting models are presented in Table II. The catchment characteristics have been automatically selected by the stepwise regression algorithm; their order in the regression equation, therefore, corresponds to the relative importance of catchment characteristics in terms of predictive performance. However, the importance for predictive performance may not be seen as a straightforward evaluation of process controls, because of intercorrelations between catchment characteristics, different accuracy of catchment characteristics and remaining influences of single outliers on the stepwise selection procedure.

The global regression model consists of eight catchment characteristics. Range of altitude H_R is of prime importance and has a positive effect on low flows. The proportion of rocks L_R , which is large in mountainous areas, has a negative effect on low flows. From three precipitation characteristics, winter precipitation P_W was selected and has a positive effect. Catchment geology is represented by four parameters: quaternary sediments G_Q and deep groundwater tables G_{GD} have a positive effect on low flows; Flysch G_F and crystalline rocks G_C have a negative effect on low flows.

Grouping into two regions and separate regressions in each region leads to two significantly more parsimonious regression equations. The summer model consists of only two catchment characteristics, i.e. winter precipitation P_W and maximum altitude H_+ , both indicating a positive effect on low flows. The winter model consists of five catchment characteristics. Winter precipitation P_W appears as the most important characteristic, followed by the proportion of quaternary sediments G_Q , both exhibiting positive effects on low flows. Also, mean slope S_M indicates a positive effect on low flows, whereas two land-use characteristics that indicate high-mountainous conditions (proportions of glaciers L_{GL} and proportion of rocks L_R) exhibit negative effects.

Grouping into three regions and separate regressions in each region leads to three regression equations. Again, the summer model consists of only two catchment characteristics, i.e. annual precipitation P and

Group	Ν	$R_{\rm cv}^2$ (%)	rmse _{cv}	Model
Global	325	57	2.62	$\hat{q}_{95} = -2.04 + 0.23H_{\rm R} - 0.08L_{\rm R} - 0.04G_{\rm F} + 1.29P_{\rm W} + 0.04G_{\rm Q} + 0.04S_{\rm ST} + 0.03G_{\rm GD} - 0.01G_{\rm C}$
Two regions total summer winter	325 182 143	59 60 47	2·58 2·60 2·55	$\hat{q}_{95} = -3.57 + 1.71P_{\rm W} + 0.18H_+$ $\hat{q}_{95} = -0.39 + 0.86P_{\rm W} + 0.20G_{\rm Q} - 0.05L_{\rm GL} + 0.12S_{\rm M} - 0.06L_{\rm R}$
Three regions total summer winter mixed	325 177 115 33	59 66 46 40	2.57 2.43 2.72 2.74	$\begin{aligned} \hat{q}_{95} &= -6.56 + 1.16P + 0.12H_{\rm R} \\ \hat{q}_{95} &= -2.40 + 1.04P_{\rm S} + 0.08S_{\rm ST} + 0.46L_{\rm WA} - 0.06L_{\rm R} \\ \hat{q}_{95} &= 14.90 + 0.04G_{\rm L} - 0.44H_{\rm M} - 0.08G_{\rm T} \end{aligned}$
Eight regions	325	58	2.61	$\hat{q}_{95} = -2.17 + Z + 1.18P_{\rm W} + 0.23H_{\rm R} - 0.08L_{\rm R} + 0.04G_{\rm Q} + 0.07S_{\rm M} - 0.03G_{\rm F} - 0.02G_{\rm C} - 0.02G_{\rm T}$

Table II. Performance and coefficients of regional regression models for q_{95} low flows^a

^aN is the number of catchments within one group; q_{95} units are 1 s⁻¹ km⁻²; units of catchment characteristics, see Table I.

range of altitude H_R , both indicating a positive effect on low flows; therefore, this is similar to the summer model for the grouping into two regions. The winter model exhibits four catchment characteristics and is, therefore, more parsimonious than the winter model of the grouping into two regions. Again, one precipitation characteristic P_S and one slope characteristic S_{ST} exhibit positive effects on low flows, and the proportion of rocks L_R exhibits a negative effect on low flows; the proportion of water surfaces L_{WA} appears as one further, positive, effect. The model for the transition zone (mixed regimes) consists of three characteristics. Catchment geology is represented by two parameters; limestone G_L exhibits positive effects on low flows, and tertiary sediments G_T exhibit negative effects on low flows. A further negative effect on low flows is given by mean altitude H_M .

Global regression, but different Z parameters in each of the eight regions, yields a model similar to the benchmark global regression model. The model exhibits exactly the same coefficients for H_R , L_R , G_Q , and only slightly modified coefficients for P_W , G_F and G_C . The percentage of steep slope S_{ST} is changed to mean slope S_M and the proportion of deep groundwater tables G_{GD} is replaced by the proportion of tertiary sediments G_T .

Overall performance of models

Table II presents two measures of model performance, the coefficient of determination R_{cv}^2 and the rootmean-squared error rmse_{cv}. Both are obtained from cross-validated residuals and, therefore, are representative of the prediction of low flows in ungauged catchments. Global regression exhibits a relative performance of $R_{cv}^2 = 57\%$, corresponding to rmse_{cv} = 2.62 1 s⁻¹ km⁻². Grouping catchments into two regions and separate regressions in each region improves the overall model performance to $R_{cv}^2 = 59\%$, rmse_{cv} = 2.58 1 s⁻¹ km⁻². The summer low-flow-dominated region exhibits better performance ($R_{cv}^2 = 60\%$) than the winter-dominated region ($R_{cv}^2 = 47\%$). All parameters except for the intercept of the model for the winter-dominated region are significant at the 0.05 significance level (Table III). Grouping catchments into three regions and separate regressions in each region yields an overall performance of $R_{cv}^2 = 59\%$, rmse_{cv} = 2.57 1 s⁻¹ km⁻², indicating no further total improvement over grouping into two regions. Again, the submodel for summer-dominated catchments exhibits a better performance ($R_{cv}^2 = 66\%$) than the submodel for summer-dominated catchments exhibits a better performance ($R_{cv}^2 = 66\%$) than the submodel for summer-dominated catchments exhibits a better performance ($R_{cv}^2 = 66\%$) than the submodel for summer-dominated catchments exhibits a better performance ($R_{cv}^2 = 66\%$) than the submodel for summer-dominated catchments exhibits a better performance ($R_{cv}^2 = 66\%$) than the submodel for summer-dominated region ($R_{cv}^2 = 46\%$), and the submodel for two regions. The global regression using different Z parameters in each of the eight regions (Tables II and IV) exhibits a moderate performance of $R_{cv}^2 = 58\%$, corresponding to rmse_{cv} = 2.61 1 s⁻¹ km⁻², i.e. it is very similar to the global regression model. Overall, the cross-validated coefficients of determination correspond well with the relative scatter of the

A similar effect was observed for the summer period low flows q_{95s} and winter period low flows q_{95w} (Table V). The global model for summer-period low flows ($R_{cv}^2 = 65\%$) clearly performs better than the

Class upper limit (1 s ⁻¹ km ⁻²) Estimation performance	0.5 Excellent	1 Very good	2 Good	3 Sufficient	5 Poor	10 Very poor	> 10 Outliers
Global model	19.4	40.3	63.7	78.5	92.6	100.0	100.0
Two regions	18.8	40.0	68.0	80.3	94.2	99.7	100.0
Three regions	24.0	44.9	65.8	79.4	93.2	99.1	100.0
Eight regions	22.8	41.5	64.0	80.9	93.8	100.0	100.0

Table VI. Cumulative frequency (%) of catchments classified by absolute values of cross-validated residuals

Copyright © 2006 John Wiley & Sons, Ltd.

1 1	C			()
Parameter	Value	SE	t value	$\Pr(> t)$
Summer				
Intercept	-3.57	0.52	-6.90	0.000
$P_{\rm W}$	1.71	0.14	11.91	0.000
H_+	0.18	0.03	5.26	0.000
Winter				
Intercept	-0.39	0.90	-0.43	0.667
$P_{\rm W}$	0.86	0.13	6.89	0.000
G_0	0.20	0.04	4.61	0.000
L _{GL}	-0.05	0.02	-2.05	0.043
$S_{\rm M}$	0.12	0.02	5.08	0.000
$L_{\rm R}$	-0.06	0.01	-4.47	0.000

Table III. Standard error (SE) and significance (*t*-test) of the parameters for the regional regression model based on two regions. The column Pr(>|t|) presents empirical significance levels of two-sided Student *t*-tests (*t* value)

Table IV. Z parameters for each of the eight regions. p is the

Group	Region	$Z (l \ s^{-1} \ km^{-2})$	<i>p</i> -value ^a
A–C	Alps	0.44	0.52
1	Flatland and hilly terrain (N and E of Austria)	0.38	0.95
2	Bohemian Massif	-0.32	0.26
3	Foothills of Alps (Upper Austria)	2.26	< 0.01
4	Flysch zone	0.22	0.40
5	Lower Carinthia	-1.33	0.07
D	Pre-Alps (Styria)	-0.42	0.33
Е	Pre-Alps (Vorarlberg)	-1.24	0.22

^a Significance level.

Table V. Performance of regional regression models for summer-period low flows q_{95s} and winter-period low flows q_{95w}

Group	Ν	$R_{\rm cv}^2$	rmse _{cv}	Model ^a
Summer period, global	325	65	3.05	$\hat{q}_{95s} = -3.94 + 0.18S_{\rm M} + 0.48P - 0.16L_{\rm GL} + 0.03G_{\rm GD} + 0.04G_{\rm Q} - 0.04G_{\rm F} + 0.25H_{+} - 0.08L_{\rm R} - 0.02G_{\rm C} - 0.02L_{\rm F}$
Winter period, global	325	49	2.87	$\hat{q}_{95w} = -0.54 + 0.91P_{\rm W} + 0.07G_{\rm Q} - 0.07L_{\rm R} - 0.04G_{\rm F} + 0.04G_{\rm GD} + 0.14S_{\rm M} - 0.02G_{\rm C}$
Summer period, summer region	182	60	2.73	$\hat{q}_{958} = -3.66 + 1.43P_{\rm W} + 0.27H_{+} - 0.04G_{\rm F} + 0.03G_{\rm GD}$
Summer period, winter region	143	50	3.07	$\hat{q}_{95s} = -0.30 + 0.52P + 0.09S_{sT} - 0.14L_{G} + 0.21G_{O}$
Winter period, summer region	182	55	2.89	$\hat{q}_{95w} = -0.39 + 2.84P_{W} - 0.04G_{F} + 0.15H_{+} - 1.15P_{S} + 0.04G_{O}$
Winter period, winter region	143	44	2.39	$\hat{q}_{95w} = 0.99 + 0.68P_{\rm W} + 0.19G_{\rm Q} + 0.14S_{\rm M} - 0.18H_{\rm M} - 0.03L_{\rm R}$

^a q_{95} units are 1 s⁻¹ km⁻²; units of catchment characteristics, see Table I.

global model for winter-period low flows ($R_{cv}^2 = 49\%$). This is also true if a grouping into two regions and separate regressions in each region is performed for summer-period (q_{95s}) and winter-period (q_{95w}) low flows. In the summer region, the model for q_{95s} ($R_{cv}^2 = 60\%$) performs better than the model for q_{95w} ($R_{cv}^2 = 55\%$), and in the winter region the model for q_{95s} ($R_{cv}^2 = 50\%$) performs better than the model for q_{95w} in the winter region ($R_{cv}^2 = 44\%$). The finding that three of these four regionally restricted models for the summer and

winter periods perform poorer than the global models for the summer and winter periods indicates that the grouping into summer- and winter-dominated regions is not suitable for summer- and winter-period low flows.

Estimation performance of individual catchments

One further assessment of estimation performance was achieved by the classification of catchments according to absolute values of cross-validated residuals presented in Table VI. Values $< 0.5 \ 1 \ s^{-1} \ km^{-2}$ correspond to excellent estimation performance, $0.5 - 1 \ 1 \ s^{-1} \ km^{-2}$ are graded very good, $1 - 2 \ 1 \ s^{-1} \ km^{-2}$ are good, $2 - 3 \ 1 \ s^{-1} \ km^{-2}$ are sufficient, $3 - 5 \ 1 \ s^{-1} \ km^{-2}$ are poor, $5 - 10 \ 1 \ s^{-1} \ km^{-2}$ are very poor, and values $> 10 \ 1 \ s^{-1} \ km^{-2}$ represent extreme outliers.

According to the normal probability plots presented earlier ('Examining model assumptions' section), residuals less than $2 \, 1 \, s^{-1} \, km^{-2}$ (at least good estimation performance) are approximately normally distributed and correspond to about two-thirds of all catchments. Hence, their cumulative frequency appears to be one reasonable index of model quality. The model based on separate regression of two regions exhibits the highest proportion of at least well-estimated catchments (68.0%) and, therefore, appears best, followed by grouping into three regions and separate regression for each region (65.8%). The global model, which includes different Z parameters for eight seasonality regions (64.0%), is only slightly better than the global model without different Z parameters (63.7%). The results correspond well to the assessment of overall model performance by coefficients of determination in the cross-validation mode ('Overall performance of model' section).

A closer look at the classification according to estimation performance leads to a different assessment. Separate regression for three regions exhibits the highest number of catchments exhibiting excellent (24.0%) and, at least, very good estimation performance (44.9%). On the other hand, this model also exhibits the highest number of extreme outliers (0.9%). Both measures indicate the highest selectivity between well-represented and poorly represented situations among all models. From this assessment, the model is clearly more adequate than separate regression for two regions.

A similar, albeit weaker, effect was identified for the global model, which includes different Z parameters for eight seasonality regions. Despite the relative poor number of at least well-estimated catchments (64.0%), the model exhibits the second highest number of excellent represented catchments (22.8%) among all models, which indicates an adjustment of the model for the eight seasonality regions to some extent. The overall gain, however, is relatively slim.

Regional performance of models

The regional performance of models was assessed by maps of residuals between predicted and observed values of specific low flow discharge q_{95} (l s⁻¹ km⁻²) for each subcatchment (Figure 12). In this figure, each panel corresponds to one model. Positive residuals indicate overestimation by the model, and negative residuals indicate underestimation. Overall, the residual maps show a much stronger differentiation of models than the scatter plots or coefficients of determination. The global model, which serves as the benchmark against which to test the models that account for seasonality, generally exhibits a rather random residual pattern, indicating that the model performs equally well for a range of hydrological conditions in Austria. Closer inspection, however, indicates that the model fits somewhat better for the northern part of Austria than for the remaining area. Large negative residuals occur in the northern Calcareous Alps, which corresponds to the bias of very wet catchments as described in the 'Examining model assumptions' section. Large residuals occur in the southern part of Upper Austria, which exhibits a specific hydrological situation that would require a separate regionalization model.

Compared with the global model, both the regional regression models that fit separate models for two and three regions obtained from seasonality analysis indicate clearly different residual patterns. The grouping into two regions and separate regressions in each region leads to an improved model performance in the southeast of Austria, most of which is part of the summer-dominated region. Alpine catchments that correspond to the winter region are unchanged. The model exhibits a higher selectivity between well-represented and poorly



Figure 12. Residuals of predicted minus observed specific low flows q_{95} (1 s⁻¹ km⁻²) in cross-validation mode. Each panel corresponds to one regionalization model

represented situations, apparently corresponding to the hydrological complexity of catchments. The model, therefore, is clearly more suitable than the global model, which does not represent these situations well. A similar, albeit even stronger, effect may be observed for the regional regression based on the grouping into three regions. Regionalization is clearly improved for a large area from the east (Wechsel region) to the southwest of Austria (East-Tyrol). This means that there is a clear improvement in summer- and winter-dominated regions. On the other hand, the number of outliers increases, and these outliers exhibit clearly larger residuals. Both effects correspond to a higher selectivity of the regionalization model between well-represented and poorly represented situations, indicating the best performance for predicting low flows among all models considered.

The global regression using different Z parameters in each of the eight regions exhibits a pattern of crossvalidated residuals very similar to the global model, which is in accordance with the small change of both the regression coefficients and the coefficient of determination. Average low flows for each region are practically identical (see Table IV), and seasonality has no significant influence on the quantity of low flows.

DISCUSSION

Performance of regionalization methods as compared to the literature

The global regression model obtained in this study uses seven catchment characteristics as predictors. These are H_R (range of altitude), L_R (fraction of wasteland or rocks), G_F (fraction of Flysch), P_W (average winter precipitation), G_Q (quaternary sediments), S_{ST} (fraction of steep slope), G_{GD} (fraction of aquifers with deep groundwater table) and G_C (fraction of crystalline rocks). The global model exhibits a cross-validated coefficient of determination of $R_{cv}^2 = 57\%$, and explains 60% of the variance in q_{95} (goodness of fit, R^2). It is interesting to compare this result with studies in the literature that used a similar number of

catchments as in this paper (325 catchments) and examined q_{95} specific discharges as in this paper, rather than discharges. We generally used cross-validated coefficients of determination in this study, since they appear more appropriate for comparisons among models that differ in both the number of catchments and the number of model parameters. Since the cross-validated coefficients are not generally used in the literature, comparisons with the literature will be based on the goodness-of-fit coefficient of determination obtained here. Gustard *et al.* (1992) obtained $R^2 = 57\%$ between Q_{95} standardized by the mean flow and portion of hydrologically defined soil classes for 694 catchments in the UK. Schreiber and Demuth (1997) obtained $R^2 = 56\%$ between specific mean annual 10-day minimum discharge MAM(10) and a number of catchment characteristics for 169 catchments in southwest Germany, and Aschwanden and Kan (1999) obtained $R^2 = 51\%$ between specific discharge q_{95} and a number of catchment characteristics for 143 headwater catchments in Switzerland. The R^2 values obtained in this study are slightly larger than those from the literature. It is likely that the difference is related to the hydrological heterogeneity of Austria with clear regional differences in low flows. The better goodness of fit in this study may also be related to using subcatchments rather than complete catchments, which may make the catchment characteristics more relevant to low flow regionalization.

Most other studies in the literature used discharge rather than specific discharge, and so are not directly comparable to the results in this paper. As catchment size usually explains around 80–90% of the variability of low flow discharges (e.g. see Dingman and Lawlor (1995) and Vogel and Kroll (1992)), it is clear that the R^2 values for discharges will be much larger than the R^2 values for specific discharges, particularly if there are significant variations in catchment size within the sample. Dingman and Lawlor (1995) and Vogel and Kroll (1995) and Vogel and Kroll (1995), for example, reported R^2 values of more than 90%.

Relative performance of regionalization methods

In a first step, we compared the overall performance of the models. The coefficients of determination in the cross-validation mode and scatter plots of observed versus predicted low flows q_{95} indicate only slight differences among the models ('Examining model assumptions" and 'Overall performance of model' sections). Fitting separate regression models for groups of catchments that are similar in terms of low flow seasonality increases the predictive performance compared with global regression, whereas the effect of incorporating different Z parameters in each of the eight regions in the global regression model was very small.

In a second step, we compared the performance of submodels for catchments that represent summer and winter low-flow-dominated regimes. As can be seen from the cross-validated coefficients of determination of the submodels, models for summer-dominated catchments exhibit a better performance than the models for the winter-dominated catchments (Table II), and the models for the summer-period low flows perform better than the models for the winter-period low flows (Table V). Summer-dominated catchments correspond to lowlands where the hydrological situation is relatively simple, and winter-dominated catchments correspond to alpine catchments where the hydrological situation tends to be complex. However, it is not quite clear whether the different performances are due to different complexities of the processes or different complexities of the regions. Separate global models for summer-period low flows and winter-period low flows, however, clearly represent different seasonal processes (summer and winter processes) occurring in the catchments. Hence, the relatively poor performance of the model for the winter-period low flows indicates that winter processes are not represented as well as the summer processes. This might point to both a higher complexity of winter processes and a poorer representation of winter processes by the available catchment characteristics.

Overall, models for summer-dominated catchments exhibit a better performance than models for winterdominated catchments, and models for summer-period low flows are better than models for winter-period low flows. Hence, the grouping based on seasonality has a clear effect on the relative performance of the models, which one would expect to translate into a more accurate regionalization. As this effect was not apparent in the performance measures discussed above, it appeared that a more detailed assessment of residuals was needed.

In a third step, therefore, we classified catchments according to estimation performance and compared cumulative frequencies of well and poorly estimated catchments among models (Table VI). The assessment

indicated a clearer effect of including seasonality in regionalization of low flows and pointed to a superior performance of separate regression for three regions that represent summer, winter and mixed seasonality. We also compared maps of residuals between predicted and observed values of specific low flow discharge q_{95} (1 s⁻¹ km⁻²) for each subcatchment (Figure 12). Also, the assessment of models by maps of crossvalidated residuals resulted in a somewhat different assessment from that of the coefficient of determination of cross-validated residuals. We believe that the main reason for this deviation of results is caused by the high sensitivity of the coefficient of determination to outliers. Regional regression performed for two and three regions led to a better representation for the bulk of catchments but increased the number of outliers at the same time, particularly within the winter-dominated region. This effect has not been apparent from the scatter plots and residual histograms for the ensemble of catchments due to the large number of points, but it is evident in the residual maps. As the coefficient of determination is very sensitive to outliers, it does not capture such a situation very well. From the residual maps, however, one can see the better performance of methods that incorporate seasonality by separate modelling of different zones of homogeneous seasonality. And residual patterns of the model based on three regions would even indicate a further separation of regions, which would, finally, lead to the classification of zones that are homogeneous in terms of low flow seasonality ('Visual grouping based on different seasonality measures' section). An alternative consideration of these eight regions in a complex regional regression model that fits separate models for each region similar to the more successful models for two and three zones may, therefore, be a promising approach.

One apparent deficiency of all models is the large scatter and clear bias for very wet catchments (Figure 10). The low flows are consistently underestimated in catchments where observed specific low flow discharges are more than about $12 \text{ l s}^{-1} \text{ km}^{-2}$, and the random prediction error is also rather large. It appears that none of the models can cope very well with these large discharges. Part of the errors may be related to biases in the observed values. A specific discharge of $12 \text{ l s}^{-1} \text{ km}^{-2}$ corresponds to 378 mm of low flow depth per year, which is a relatively large value for Austrian conditions. It should also be noted that it is not uncommon for regionalization models to have a tendency for underestimating large values. For example, the flood regionalization analysis of Merz and Blöschl (2005) showed that flood quantiles in the same study area were consistently underestimated by their method for catchments with above-average specific flood discharges.

Relative performance of approaches to seasonality analysis

We used different approaches to seasonality analysis in this paper. We first compared three seasonality measures, the SR, the SI and the SH, which use a different number of parameters for representing the seasonal distribution of low flows. The SR consists of one parameter, which is the ratio of summer-period low flow characteristic q_{95s} and winter-period low flow characteristic q_{95s} . The SI consists of two parameters, the mean and the variability of days when discharges fall below Q_{95} . An SH consists of 12 parameters that represent the monthly frequency of days when discharges fall below Q_{95} . The SH, therefore, provides the most detailed information about low flow seasonality, the SI contains less information, and the SR contains the least information. However, one drawback is that the synoptical interpretation of seasonality measures from a large number of catchments becomes more difficult with an increasing number of parameters. Whereas general patterns are best visible for the SR, the determination of patterns by the SI requires closer inspection. There is too much information in the SH for a synoptical interpretation at the regional scale, so some classification technique is needed to exploit the higher information content.

Hence, we compared two classification techniques in order to exploit the information of the SH in the best possible way. The first technique, a partitioning approach to cluster analysis that also provides information about the optimal number of clusters, led to a classification of Austrian catchments into two contiguous regions that correspond to summer and winter low flow regimes. An alternative classification was obtained by allocating catchments that exhibit strong dissimilarities to both clusters to a third group, which hence corresponds to weak seasonality or a mixed low flow regime. Both clusters correspond well with patterns of SR and are plausible, since they are consistent with landscape regions in Austria. Although the effect on the

overall model performance was small, closer inspection of the performance of submodels for the regions and the residual pattern maps showed that the classification increases the performance of regionalization and there is some value in the seasonality grouping. The second technique is the visual grouping of the SH, which uses the large-scale pattern of the SI to obtain a preliminary classification that is then refined by supplemental information provided by the SH. Topographic information was further used to infer boundaries of the regions. The contiguous regions so obtained are homogeneous in terms of seasonality, and they generally correspond well with patterns of SR and both groupings obtained from cluster analysis, and exhibit a finer classification than the regions obtained by cluster analysis and a visual inspection of the patterns of SR and SI alone. Although this classification did not give a significant improvement over the global model, we believe there is some potential in this classification provided that more sophisticated regionalization models are used.

Most important controls in comparison with controls in the literature

Controls of low flow seasonality. The results of the seasonality analysis allow an interpretation of processes of low flow generation in Austria at the regional scale. Catchment altitude exerts a very important influence on seasonality. Essentially by altitude, Austria is divided into an alpine region where low flows are dominated by winter processes and into flatlands and hilly terrain where low flows are dominated by summer processes. The changeover between these two regions is restricted to rather small transition zones (foothills of the Alps along the Flysch zone and in Eastern Styria, and Lower Carinthia). Smaller differences of mean seasonality between subregions may be explained by different climatic influences, such as influences of Atlantic, continental, pannonic and interalpine climate and, perhaps, foehn effects. The seasonality patterns appear independent of the patterns of q_{95} . This suggests that seasonality is an indicator of processes rather than an indicator of the magnitude of specific low flows.

It is interesting to compare the results of the seasonality analysis of Austrian catchments with results from geographically similar regions. Schreiber and Demuth (1997) analysed seasonality of MAM(10) of total discharges of 169 catchments in southwest Germany. The study area is situated close to the northwest of Austria. Large parts of southwest Germany exhibit hilly landscapes that are very similar to the northern part of Austria, apart from one region termed Südwestdeutsches Schichtstufenland, which exhibits different geological conditions. Generally, the low flow seasonality of southwest Germany corresponds well to seasonality types of the northern part of Austria (type 3 and type 4), although seasonality in German catchments appears somewhat retarded compared with Austrian catchments. This might be due to greater influence of Atlantic climate, which typically shows a higher tendency of rainfall events in summer season than a continental climate and, therefore, should cause slower baseflow recessions. A further reason might be found in the higher altitude of German catchments compared with Austrian summer-type catchments, which result in lower evaporation rates, which also may lead to slower baseflow recession. Both studies highlight the dominant influence on catchment altitude on low flow seasonality in a temperate climate. However, the low flow seasonalities of the Pre-Alps differ strongly in the two study areas. We believe that this difference is caused by the different designs of the two studies. Although seasonality measures calculated in this study are representative for subcatchments, the seasonality measures calculated in southwest Germany are representative for total discharges of entire catchments. The probable reason for this difference is that the main rivers in the Pre-Alps frequently arise in alpine catchments, and the seasonality of total discharges as calculated for southwest Germany reflects, in this specific case, the clear influence of the alpine flow regime on low flows and, therefore, is not comparable to seasonality of residual discharges of subcatchments as calculated in this study.

Aschwanden and Kan (1999) investigated the long-term characteristic seasonal distribution of Q_{95} for different regions in Switzerland, based on the 1935–96 observation period. Switzerland is situated next to the southwestern border of Austria and is geographically very similar to Austria, apart from the higher altitude of the Swiss Alps. The Swiss authors found two typical seasonal distributions of low flows in Switzerland, depending on catchment altitude. In alpine catchments, low flows occur exclusively from November to March.

In hilly landscapes, such as Mittelland and Jura, low flows may occur during the whole year, but clearly most frequently during summer and autumn. Seasonality patterns in Austria are generally in line with seasonality types in Switzerland, since alpine catchments exhibit strong winter seasonality in both countries, and hilly landscapes exhibit weaker seasonality, along with mean occurrence in summer and autumn. The occurrence of low flows during all month in the Swiss hilly landscapes, which has not been observed for Austrian catchments, probably reflects the aggregation of discharge from alpine catchments and hilly landscapes. Our approach, which is based on residual flows of nested catchments, proved appropriate to eliminate such aggregation effects. Landscapes in the northeast and east of Austria, however, exhibit a lower altitude and a more continental climate than Swiss landscapes and indicate much stronger summer seasonality than Swiss landscapes.

Controls of specific low flow discharge. Further insights into important controls of low flows were gained from a comparison of the different regional regression models (Table II). Although a large number of catchment characteristics were supplied to the stepwise variable selection algorithm, only 17 catchment characteristics have been selected. Most of them occur in several models. Many regression models contain similar parameters, although the models for the summer and the winter regions differ significantly. Overall, most of the regression coefficients are plausible, as their signs and some of the catchment characteristics can be interpreted on hydrological grounds. We believe, therefore, that the interpretation of the regression coefficients provides useful insights into the important process controls of low flows in Austria.

Apart from the model for the heterogeneous group of catchments that exhibit mixed seasonality (a small band along the Alps exhibiting low variability of precipitation), precipitation is significant for all models. Also, the fact that one precipitation characteristic is frequently selected first by the stepwise regression algorithm supports the finding that precipitation is one of the most important controls of low flows in Austria. Precipitation has a positive effect for summer and winter low flows, as one would expect. The positive effect on winter low flows may be related to a tendency of precipitation periods to be generally warmer than dry winter periods. This may be a result of different atmospheric circulation patterns in cold and wet winter periods. However, the effect of annual and seasonal precipitation remains unclear, since winter precipitation is most frequently selected by stepwise regression and also appears in the model for one summer region, whereas summer precipitation appears in the model for one winter region. All precipitation characteristics exhibit very similar spatial patterns, and the selection of one characteristic appears rather random.

Catchment relief is represented in all equations, generally by one altitude parameter $(H_M, H_+ \text{ or } H_R)$ or by one slope parameter $(S_{ST} \text{ or } S_M)$, in the case of both global models concurrently by altitude and slope. The relief, further, has a strong influence on precipitation. Hence, catchment topography (altitude and slope) appears as an equally important control of low flows in Austria as precipitation. Altitude has a positive effect on summer low flows (less evaporation) and a negative effect on winter low flows (lower temperature). Slope generally has a positive effect on low flows; it is possibly correlated with storage volume in high mountains.

Catchment geology is represented in many regression models, generally by multiple parameters, but usually selected after topographical or precipitation characteristics. Hence, it represents the third important control of low flows in Austria. Six out of nine geological characteristics are significant in explaining low flows in Austria. Three geological classes that represent low permeability aquifer conditions, i.e. proportion of crystalline rock G_C , Tertiary sediments G_T and Flysch G_F , exhibit negative effects on low flows. Two geological classes that indicate high permeability aquifers, i.e. proportion of Quaternary sediments G_Q and deep groundwater table aquifers G_{GD} , exhibit positive effects on low flows. The large number of geological characteristics is probably due to the use of geological classes separately instead of one geological index, as sometimes used in the literature.

Land use plays a minor role. The proportion of rocks L_R is the sole land-use characteristic that appears frequently, but it is associated with high mountain conditions rather than an independent indicator of land use. Two further land-use characteristics were selected only once, i.e. the proportion of glaciers L_{GL} (which has a negative effect on winter low flows) and the proportion of water surfaces L_{WA} (which has a positive effect on winter low flows), but they also appear as indicators of high mountain conditions (L_{GL}) and hydrological conditions (L_{WA}), rather than as independent indicators of land use. However, the true land-use characteristics, such as urban areas, agriculture, permanent crop, grassland and forest, were never selected by the stepwise selection technique and do not appear to be important controls.

Stream density D never occurred in the regression equations, and does not appear to be a significant indicator of low flows. Similarly, subcatchment area A was never selected by the stepwise regression algorithm, so very little influence of catchment scale on specific low flows appears to exist. This is likely due to the large spatial scales of drought events, which exceed the spatial scale of the subcatchments used in this study.

In the final step, we compare process controls of Austria with those recorded in the literature. Smakhtin (2001) provided a comprehensive overview of catchment descriptors used in regional estimation models. Basin and climate characteristics that are most commonly related to low-flow indices include catchment area, mean annual precipitation, channel and/or catchment slope, stream frequency and/or density, percentage of lakes and forested areas, various soil and geology indices, length of the main stream, catchment shape and watershed perimeter and mean catchment elevation. The frequency of different categories of catchment descriptors in 120 low flow estimation models was assessed by Demuth and Young (2004): 73% of all catchment characteristics used are physiographic descriptors, 22% are climatic descriptors and 5% are hydrological descriptors. Among the physiographic descriptors, morphometric descriptors make up the highest proportion (46%), followed by surface cover (17%), geology and soil (10% each). However, we believe that the frequency of the catchment characteristics used depends largely on the availability and quality of the data, so a general assessment of the importance of catchment characteristics from this comparison is difficult. Such an assessment is only meaningful among studies that exhibit similar hydrological conditions and similar study designs.

The regionalization study of Switzerland presented by Aschwanden and Kan (1999) is similar to this study in both respects. The regionalization of q_{95} specific low flows in Switzerland resulted in seven regional regression models for different regions in Switzerland. From all regression equations, precipitation and catchment topography appeared as the most important control of low flows. This is fully consistent with the results of this study. A number of pedologic and hydrogeologic parameters were also significant in Aschwanden and Kan (1999). Land use played an important role, but the characteristics were proportion of horticulture and vineyards and the proportion of pre-alpine farming structures. These characteristics are representative of typical landscapes rather than land cover per se. Overall, the results obtained in this study, therefore, are in line with the regionalization study of Switzerland. Precipitation, relief, and hydrogeological conditions appear to be the most important controls of low flows in both countries. The way these controls are related to low flows, however, depends on whether a summer or winter low flow regime is present.

CONCLUSIONS

The objective of this study was to examine the value of different seasonality indices for low flow regionalization. In a first step, three seasonality indices were compared. The main difference between the indices is the information content of low flow seasonality. SHs are the most detailed indices, but classification techniques are needed to compare seasonality among a large number of catchments. The cyclic SI is a more compact index, and the spatial patterns can be delineated by visual inspection of a vector map of SI. The SR is the most condensed index, and the spatial patterns are clearly discernable when plotted on a map. The patterns of the indices obtained for Austria correspond well with the main landscape units of Alps, lowlands and hilly landscapes. In a second step, three catchment classification into two regions corresponding to summer low-flow-dominated and winter low-flow-dominated regimes. The second classification into three regions singles out an additional zone of mixed seasonality. The third classification consists of eight zones that correspond to catchments that exhibit similar typical seasonal distributions of low flows. In a third step, the value of seasonality indices for low flow regionalization was examined by comparing three multiple regression

approaches (which include the seasonality classifications in different ways) with the global regression model (which does not include seasonality). The overall coefficient of determination in cross-validation mode does not change much between the seasonality approaches. Fitting separate models to two regions (summer and winter seasonality) and three regions (summer, winter and mixed seasonality) performed best ($R_{cv}^2 = 59\%$). Including different calibration coefficients in each of the eight seasonality regions resulted in $R_{cv}^2 = 58\%$ and, hence, performs only slightly better than the global regression model ($R_{cv}^2 = 57\%$). The models for the summer regions ($R_{cv}^2 = 66\%$ and 60%), however, clearly perform better than the models for the winter regions ($R_{cv}^2 = 47\%$ and 46\%). The model for the catchments that exhibit mixed seasonality ($R_{cv}^2 = 40\%$) does not perform as well. A classification of catchments according to estimation performance and the residual maps of predicted minus observed q_{95} low flows indicates a clearer difference between models than suggested by the overall coefficients of determination. They allow a better discrimination between well-represented situations and outliers that occur in hydrologically complex parts of the study area. Separate regressions for three and two regions give smaller residuals than the global model. Including different calibration coefficients for each of the eight seasonality regions did not reduce the residuals significantly. This suggests that using separate regression models in different seasonality zones may be a promising approach, which will be examined in detail in a subsequent study.

ACKNOWLEDGEMENTS

Analyses are based on data provided by the Geological Survey of Austria, by the Austrian Federal Ministry of Agriculture, Forestry, Environment and Water Management, and by the Austrian Hydrographical Service. We are very grateful to H. Strelec and two anonymous reviewers for many valuable comments on the manuscript.

REFERENCES

- Aschwanden H, Kan C. 1999. Le débit d'étiage Q347—Etat de la question. Communications hydrologiques 27. Service hydrologique et géologique national: Berne.
- Aubrecht P. 1998. Corine Landcover Österreich. Vom Satellitenbild zum digitalen Bodenbedeckungsdatensatz. Monographien, Band 93. Umweltbundesamt: Wien.
- Behr O. 1989. Digitales Modell des Oberflächenentwässerungssystems von Österreich. Forschungsbericht 11. Institut für Hydraulik, Gewässerkunde und Wasserwirtschaft, Technische Universität Wien.
- Burn DH. 1997. Regionalisation of catchments for regional flood frequency analysis. Journal of Hydrologic Engineering 2(2): 76-82.
- Demuth S, Young AH. 2004. Regionalization procedures. In *Hydrological Drought: Processes and Estimation Methods for Streamflow and Groundwater*, Tallaksen LM, van Lanen HAJ (eds). Developments in Water Science 48. Elsevier.
- Dingman SL, Lawlor SC. 1995. Estimating low flow quantiles from drainage-basin characteristics in New Hampshire and Vermont. *Water Resources Bulletin* **31**(2): 243–256.
- Efron B, Tibshirani RJ. 1993. An Introduction to the Bootstrap. Monographs on Statistics and Applied Probability 57. Chapman and Hall: New York.
- Fürst J. 2003. Gewässernetzdichte [Stream density]. In *Hydrologischer Atlas Österreichs*, BMLFUW (ed.). Österreichischer Kunst- und Kulturverlag: Wien.
- Grayson R, Blöschl G (eds). 2002. Spatial Patterns in Catchment Hydrology-Observations and Modelling. Cambridge University Press.
- Gustard A, Bullock A, Dixon JM. 1992. Low flow estimation in the United Kingdom. Report No. 108, Institute of Hydrology, Wallingford.
- Institute of Hydrology (ed.). 1980. Low flow studies report. Institute of Hydrology: Crowmarsh Gifford, Wallingford, Oxon, UK.
- Kaufman L, Rousseeuw PJ. 1990. Finding Groups in Data: An Introduction to Cluster Analysis. Wiley: New York.

Kresser W. 1965. Österreichs Wasserbilanz. Österreichische Wasserwirtschaft 17: 213-221.

- Kresser W, Kirnbauer R, Nobilis F. 1985. Überlegungen zur Ermittlung von Niederwasserkenngrößen. Mitteilungsblatt des Hydrographischen Dienstes 54. Bundesministerium für Land- und Forstwirtschaft: Wien.
- Laaha G. 2000a. Statistik-Experten-Tools für Niederwasserfragestellungen. Forschungsbericht beim BMLFUW, Abt. IV A 3.
- Laaha G. 2000b. Zur Beurteilung der Genauigkeit von Niederwasserkennwerten [On evaluating the accuracy of low flow characteristics]. Mitteilungsblatt des Hydrographischen Dienstes in Österreich 80. Bundesministerium für Land- und Forstwirtschaft: Wien; 61–68.
- Laaha G. 2002. Modelling summer and winter droughts as a basis for estimating river low flows. In *FRIEND 2002—Regional Hydrology: Bridging the Gap between Research and Practice*, van Lanen HAJ, Demuth S (eds). IAHS Publication No. 274. IAHS Press: Wallingford; 289–295.
- Laaha G, Blöschl G. 2003. Saisonalität von Niederwasserspenden in Österreich. Mitteilungsblatt des Hydrographischen Dienstes in Österreich 82. Bundesministerium für Land- und Forstwirtschaft: Wien.

G. LAAHA AND G. BLÖSCHL

- Laaha G, Blöschl G. 2005. Low flow estimates from short stream flow records—a comparison of methods. *Journal of Hydrology* **306**: 264–286.
- Lorenz P, Skoda G. 1999. Ermittlung von Flächenmitteln des Niederschlags aus punktuellen Messungen. Mitteilungsblatt des Hydrographischen Dienstes in Österreich 78. Bundesministerium für Land- und Forstwirtschaft: Wien; 47–66.
- Mardia K. 1972. Statistics of Directional Data: Probability and Mathematical Statistics. Academic Press: London.
- Merz R, Blöschl G. 2005. Flood frequency regionalisation—spatial proximity vs. catchment attributes. *Journal of Hydrology* **302**(1-4): 283-306.
- Merz R, Piock-Ellena U, Blöschl G, Gutknecht D. 1999. Seasonality of flood processes in Austria. In *Hydrological Extremes: Understanding, Predicting, Mitigating*, Gottschalk L, Olivry J-C, Reed D, Rosbjerg D (eds). IAHS Publication No. 255. IAHS Press: Wallingford; 273–278.
- Nathan RJ, McMahon TA. 1990. Identification of homogeneous regions for the purpose of regionalisation. *Journal of Hydrology* **121**: 217–238.
- Piock-Ellena U, Pfaundler M, Blöschl G, Burlando P, Merz R. 2000. Saisonalitätsanalyse als Basis für die Regionalisierung von Hochwässern. Wasser, Energie, Luft 92. Jahrgang, Heft 1/2, Baden; 13–21.

Schreiber P, Demuth S. 1997. Regionalisation of low flows in southwest Germany. Hydrological Sciences Journal 42(6): 845-858.

Smakhtin VU. 2001. Low flow hydrology: a review. Journal of Hydrology 240: 147-186.

- Tallaksen LM, Hisdal H. 1997. Regional analysis of extreme streamflow drought duration and deficit volume. In *FRIEND'97—Regional Hydrology: Concepts and Models for Sustainable Water Resource Management*, Gustard A, Blazkova S, Brilly M, Demuth S, Dixon J, van Lanen H, Llasat C, Mkhandi S, Servat E (eds). IAHS Publication No. 246. IAHS Press: Wallingford; 141–150.
- Vogel RM, Kroll CN. 1992. Regional geohydrologic-geomorphic relationships for the estimation of low-flow statistics. *Water Resources Research* 28(9): 2451–2458.

Weisberg S. 1985. Applied Linear Regression, second edition. Wiley: New York.

Young AR, Round CE, Gustard A. 2000. Spatial and temporal variations in the occurrence of low flow events in the UK. *Hydrology and Earth System Sciences* 4: 35-45.