Modelling summer and winter droughts as a basis for estimating river low flows

GREGOR LAAHA
Department of Applied Statistics, University of Agricultural Sciences (BOKU),
Gregor Mendel Strasse 33, A-1180 Vienna, Austria
e-mail: laaha@ezy1.boku.ac.at

Abstract An approach is presented towards modelling low flow characteristics from basin characteristics based on seasonalities which, ultimately, can be used for estimating design values for ungauged sites. The main idea behind this approach is that summer and winter droughts are related to important differences in underlying physical processes. Consistent modelling of low flows therefore should be based on a separation into groups with similar seasonality. Different methods for detecting regional seasonality patterns have been applied to data from 57 monitored basins in Upper Austria. Results indicate a clear spatial separation between stations with typical summer and winter seasonality depending on altitude. For modelling nonlinear relationships, interactions and discontinuous patterns, regression tree modelling is proposed and appears to be more appropriate than multiple regression.

Key words low flow; drought; seasonality; regionalization; regression tree; multiple regression; cluster analysis

INTRODUCTION

The determination of river discharges during floods and droughts is a necessary precondition for many topics related to water resources management, such as the optimized usage of water resources, the qualitative and quantitative protection of water courses and the protection of man against flood damage. For an assessment of water resources or hydrological risks, low flows are commonly summarized by characteristic values which can be calculated from hydrographs or need to be estimated at ungauged sites. There are two reasons why the determination of such characteristics should account for seasonality. First, the interaction between quantity and seasonality of low flows can be a relevant factor. For instance, the resilience of aquatic ecosystems depends on the concentration of pollutants and water temperature. Quantity and seasonality of low flows therefore have to be considered when determining design values of acceptable sewage loads. Second, the consistent determination of low flow characteristics should take the different physical processes of summer and winter droughts into account.

As a consequence, current scientific techniques used to model low flow characteristics for alpine regions in Austria need to be modified in order to manage hydrological risks better. In this paper, a methodology will be elaborated for a representative subregion. The database consists of a daily discharge time series of 57 gauges in Upper Austria monitored during a 20-year standard period, from which the low flow characteristic $Q_{95}$, $Pr(Q > Q_{95} = 0.95)$ has been calculated. In order to avoid
the problem of nested basins, only small basins have been selected. Additionally, the
discharge yield \( \Delta Q_{95} \) of residual basins between two subsequent nested gauges have
been added to the database, based on the assumption that summer low flows occur
simultaneously (see below).

From a geological point of view, the study area consists, from north to south, of
the Bohemian Massif (Palaeozoic granite and gneiss), the Danube Valley (alluvial
sediments), the molasse zone (Tertiary sediments, often superimposed by Quaternary
gravel), the flysch zone (Helvetic loam and sandstone) and the High Alps (limestone
and dolomites, often subjected to karstic influences). This strong geological diversity
gives rise to important variations in topographical characteristics (altitude, aspect,
slope) and land use, but also to differences in climatic conditions (temperature and
seasonal or annual precipitation). Similar to recent studies in Switzerland by
Aschwanden & Kan (1999), and also to studies by the Institute of Hydrology (1980)
and Demuth (1993), the relationships between specific low flow discharge \( q_{95} \)
\( (1 \text{ s}^{-1} \text{ km}^{-2}) \) \( (Pr(q > q_{95}) = 0.95) \) and about 30 basin characteristics will be examined in
this paper. Additionally, a statistical approach based on seasonality is proposed.

**PHYSICAL DIFFERENCES BETWEEN SUMMER AND WINTER DROUGHTS**

Consistent modelling of low flow characteristics requires taking into account the
respective underlying physical processes for summer and winter low flows. Summer
low flows emerge during long-term persistent dry periods when evaporation exceeds
precipitation. This leads to a slow depletion of the soil reservoir, which results in
recession curves of hydrographs. Therefore, summer low flows generally emerge
synchronically over large climatically homogeneous regions and depend on the storage
properties of the soil as well as on the ratio of precipitation and evaporation. Winter
low flow events occur due to freezing of the soil and due to the storage of precipitation
as snow. In western and central Europe, they therefore seem to depend on the altitude
of the hydrological basin. In order to perform a consistent regional analysis, it is
important to separate summer and winter low flows. The approach used in this study is
to find groups of stations with typical seasonal distributions that can be attributed to
summer or winter seasonality but also enables transition regions to be identified.

**REGIONAL ANALYSIS OF THE SEASONALITY OF LOW FLOW EVENTS**

**Methods**

The *seasonality index* presented by Burn (1997) has been applied to describe the
typical seasonality of the low flow events by the mean and standard deviation of their
date of occurrence (Fig. 1) using circular statistics (Mardia, 1972). This method has
proved useful in several meteorological and hydrological studies, e.g. Bartels (1943),
Nobilis (1986), Merz *et al.* (1999), Young *et al.* (2000). However, from a statistical
point of view, the description of seasonal distributions by their mean and standard
deviation requires that the occurrences follow a unimodal distribution. If the
distributions are multimodal, as may result from a mixing of hydrological processes in transition zones, the seasonality index may be ambiguous or not physically meaningful. Therefore, it could be argued that the seasonality index only provides a rough description of seasonal distributions.

In order to obtain a more detailed description, a cluster analysis of the monthly low flow occurrence frequencies has been elaborated. Low flows are assumed to occur when the mean daily streamflow falls below a specific threshold (e.g. Q95). Histograms of low flow occurrence frequencies for each month describe the distribution of low flow events of a specific gauging station during the year. Clearly, the information about the seasonality of low flow events indicated by the shape of the histogram goes beyond the sole use of mean and standard deviation of the dated occurrence. Cluster analysis (partitioning around medoids, see Kaufmann & Rousseeuw, 1990) has been used for an automatic classification of measurement stations. In addition to clustering, this method calculates the silhouette width—a statistical measure representing the dissimilarity between the low flow occurrence histograms and the corresponding histogram of the related cluster centre. The silhouette width has been used to rank histograms by similarities. The graphical diagnostics obtained have proved useful for a temporal and regional analysis of low flow seasonality. Furthermore, the optimal number of clusters is related to the maximum average silhouette width among different possible partitions. Therefore, no a priori assumptions about the number of different seasonal distributions need be made.
Application

The analysis of 57 gauges in Upper Austria resulted in two groups of stations showing similar seasonal distributions of days when low flows occurred. In the first group, consisting of 47 gauges, 90% of the events occurred between July and November. Only a small number of events began in May or June or persisted until December. Exhibiting this typical seasonality, these stations will be referred to as summer low flow stations. Similarly, the second group consists of eight winter low flow stations. When examining the study area for spatial patterns of similar stations, it turned out that these groups form two homogeneous subregions. These also show significant differences in their geographical characteristics since the winter low flow stations are situated in the alpine region, while the summer low flow stations are situated in the lower parts of the area (Fig. 2). This clear emergence of two spatially separated regions due to different underlying low flow processes has been used for a phenomenon-based model of low flow characteristics as discussed below.

MODELLING OF WINTER AND SUMMER LOW FLOWS

Winter low flows

Searching for an indicator of the stations' seasonality, a clear dependency between seasonality and maximum basin altitude has been found. Small basins reaching up to more than 1870 m exhibit winter low flows while summer low flows occur in basins lower than 1790 m. In between these limits, mixed seasonality occurs. Obviously
arising from lower temperatures at high altitudes, this relationship can be applied to
assign typical seasonalities to small ungauged basins and therefore can be used to
select the valid model from two seasonal restricted models for estimating low flow
characteristics at ungauged sites.

Analysing the influence of basin characteristics on winter low flows, a significant
positive correlation between \( q_{95} \) and the mean altitude has been found, indicating that
higher basins experience less severe low flows. This result seems surprising, given that
higher basins are associated with lower temperature, which should lead to severer
winter droughts. However, all gauges are situated in valleys between 400 and 600 m
altitude and mean altitude therefore reflects a measure of the average range of altitudes
within catchments. Hence, one possible interpretation is that this average range of
altitude is just a surrogate for the subsurface storage volume in high mountains that
compensates low flows. In this case, it should be considered as an important predictor
in modelling specific discharges of mountainous regions.

Finally, a temporal analysis of winter low flows has been performed for each year.
Results indicate that for basins exhibiting typical winter seasonality, extreme low flow
events already start in late summer due to a lack of precipitation. Therefore, it can be
argued that summer precipitation should be included in the modelling of winter low
flow processes. The definite elaboration of the winter model is the subject of ongoing
research.

**Summer low flows**

**Situation** Prior to the statistical modelling, the relationships between low flows
and basin characteristics had been investigated. The correlation analysis indicated a
significant correlation between \( q_{95} \) and summer precipitation while other basin
characteristics did not exhibit significant correlations. The scatterplot matrix and
principal component analysis pointed to the existence of nonlinearity and
heterogeneity. Furthermore, inter-correlations of different topographic parameters, as
well as correlations between altitude, precipitation and land use appeared. Finally, the
presence of exact collinearity within four groups of variables (land use, geology,
aspect and slope) should be mentioned. For this situation, classical multiple linear
regression is not very well suited. A different approach given by the regression tree
model is therefore proposed.

**Methodology** Regression tree modelling is an exploratory technique for finding
homogeneous regions among predictor variables with respect to a target variable and
can be said to carry out a clustering of similar landscapes with respect to low flows.
Modelling is done by binary recursive partitioning (Clark & Pregibon, 1991), by which
groups of stations are subsequently subdivided by binary conditions (e.g. IF
precipitation < 560 mm THEN ... ELSE ...), starting from the most important process
parameters and proceeding to the least important parameters. The subregions so
obtained are simply modelled by their mean specific low flow discharge or may serve
as a basis for more detailed modelling. This concept is similar to classification trees
which have been applied in hydrology to classify satellite images of snow cover, e.g.
Rosenthal & Dozier (1996) and Blöschl (1999). It appears that one of the advantages of the regression tree is that it can better handle collinearity and inter-correlations than can multiple regressions. This is due to the stepwise procedure of the approach.

**Results** From the regression tree, Upper Austria is subdivided into six regions (Fig. 3). The lowest specific low flow discharges \( q_{95} \) occur not only in the relatively dry molasse zone and in the lower part of the Bohemian Massif but also in the more humid flysch zone where—probably due to reduced storage capacities of loamy soils and due to moderate precipitation—\( q_{95} \) varies between 2.1 and 3.2 \( 1 \text{s}^{-1} \text{km}^{-2} \). Larger low flow discharges of 4.5 and 6.5 \( 1 \text{s}^{-1} \text{km}^{-2} \) occur in the southern part of the molasse zone (strong precipitation and good storage properties due to a deep layer of Quaternary gravel) as well as in important aquifers of the Danube and Traun rivers. The highest specific low flow discharges of about 11.0 \( 1 \text{s}^{-1} \text{km}^{-2} \) are assigned to alpine basins and the mountainous part of the Bohemian Massif. These results underline the importance of summer precipitation and geology in explaining regional variations of specific summer low flow discharges.

![Fig. 3 Homogeneous low flow regions in Upper Austria obtained by regression tree modelling. Numbers indicate specific discharge \( q_{95} \) (1 s\(^{-1}\) km\(^{-2}\)).](image)

**DISCUSSION**

The approach presented here appears appropriate to address two topics of the Austrian low flow project. One topic consists of providing an overall description of low flow discharges at the scale of the whole of Austria by thematic maps of the hydrological atlas (to appear within the next two years). Modelling can also provide one important source of information to establish design values. Both topics require estimation procedures for ungauged sites. Within the presented methodology, basins must be of
pure seasonality to obtain consistent estimations. For big, nested basins, this can be achieved by aggregating the estimated discharge yield $A\Delta 95$ of residual basins of pure seasonality. It is worth noting that this approach seems to be inappropriate for regions dominated by mixed seasonality where a temporal separation might improve consistency of modelling.

CONCLUSIONS

Consistent modelling of low flow characteristics should be based on seasonality. Graphical diagnostics derived from a cluster analysis of monthly low flow frequency proved useful for investigating regional seasonality patterns. Results from a case study of one partially mountainous region in Austria suggested a spatial separation of summer and winter droughts. For geographically highly discontinuous regions a regression-tree-based approach appears to be more appropriate than classical multiple regression. The model can be used to estimate low flow characteristics for ungauged sites after splitting basins into residual basins of pure seasonality.

Acknowledgements The work presented in this paper is part of the research project "Statistical-Expert-Tools for Low Flow Analysis" funded by the Austrian Hydrographical Service. Analyses were based on digital data provided by the Geological Survey of Austria, by the federal ministry BMLFUW and by the Austrian Hydrographical Service. I am very grateful to G. Blöschl and the NE-FRIEND low flow group for many valuable comments on the manuscript.

REFERENCES


