DISSERTATION

Process based regionalisation of low flows

(Prozessbasierte Regionalisierung von Niederwasserabflüssen)

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines Doktors der technischen Wissenschaften unter der Leitung von

a.o.Univ.Prof. DI. Dr. Günter Blöschl
E222
Institut für Hydraulik, Gewässerkunde und Wasserwirtschaft

und

Univ.Prof. DI. Dr. Harald Strelec
H505
Institut für Angewandte Statistik und EDV

eingereicht an der Technischen Universität Wien
Fakultät für Bauingenieurwesen

von

DI. Gregor Laaha
8940792
Schlöglgasse 45/1/7
1120 Wien

Wien, im Dezember 2003

Gregor Laaha

Kurzfassung

Abflußdaten für 325 österreichische Einzugsgebiete mit einer Gebietsfläche von 7 bis 963 km² werden zur Analyse der Genauigkeit (Kreuzvalidierung) mehrerer Methoden zur Ermittlung von Niederwasserabflüssen Q_{95} an Stellen ohne Abflußbeobachtungen verwendet. Q_{95} entspricht jener Abflußmenge, die an 95% aller Tage der Meßperiode überschritten wurde. Der erste Vergleich zeigt, daß die Verwendung der Niederwassersaisonalität zur Klassifikation der Einzugsgebiete in Regionen eine Verbesserung der Genauigkeit eines Regressionsmodells zwischen Niederwasserabflußspenden q_{95} und Gebietskenngrößen gegenüber einem globalen Modell bewirkt, wenn für jede Region ein getrenntes Regressionsmodell erstellt wird. Der zweite Vergleich zeigt, daß ein regionaler Regressionsansatz, der auf einer Gruppierung der Einzugsgebiete in acht Saisonalitätsregionen basiert, mit einer erklärten räumlichen Varianz von 70% für q_{95} eine wesentlich höhere Genauigkeit erzielt als Regressionsansätze, die auf alternativen Gruppierungen basieren (Residuenmustermethode, gewichtete Clusteranalyse, Regressionsbaum). Eine dritte Analyse erschließt die Information kurzer Abflußreihen für die Schätzung von Q_{95}. Kontinuierliche Abflussbeobachtungen über ein Jahr übertreffen das beste Regionalisierungsverfahren, während Einzelmessungen deutlich ungenauere Werte als das beste Regionalisierungsverfahren ergeben. Die Analysen zeigen, daß Prozeßverständnis jedenfalls zur Regionalisierung von Niederwasserkenngrößen beitragen kann, und damit eine genauere Ermittlung der Niederwasserabflüsse als mittels existierender Standardverfahren möglich ist.

Abstract

Stream flow data from 325 Austrian catchments, ranging in area from 7 to 963 km^2, are used for exploring the predictive (cross validation) performance of a number of methods for estimating Q95 low flows in ungauged catchments. Q95 is the discharge exceeded on 95% of all days of the measurement period. The first comparison suggests that the use of low flow seasonality indices to group catchments into regions improves the predictive performance of a regression model between low flows and catchment characteristics over a global model, provided separate regressions are used in each region. The second comparison suggests that a regional regression approach based on a catchment grouping of eight seasonality regions outperforms regressions based on other catchment groupings including the residual pattern approach, weighted cluster analysis and regression trees, and explains 70% of the spatial variance of q95 specific low flow discharges. A third analysis exploits the information from short stream flow records for estimating Q95. One year of continuous stream flow data outperforms the best regionalisation method but one spot gauging does not outperform the best regionalisation method. The analyses suggest that process understanding can indeed assist in regionalising low flow characteristics more accurately than existing standard methods.

Zusammenfassung

Zahlreiche wasserwirtschaftliche und wasserbauliche Fragestellungen erfordern eine genaue Kenntnis von Niederwasserkenngrößen. Dazu zählen die optimale Nutzung von Wasserressourcen, der qualitative und quantitative Schutz von Gewässern und der Betrieb von Wasserkraftwerken. Für Gewässerstellen, an denen keine ausreichend langen Abflußbeobachtung vorliegen, können Niederwasserkennwerte mittels hydrologischer Regionalisierungsverfahren aus Einzugsgebieten mit Abflussmessungen übertragen werden. Die dieser Dissertation zugrundeliegende These ist, daß Prozeßverständnis, auch in vereinfachtem Maße, zur Regionalisierung von Niederwasserkenngrößen beitragen kann, und hierdurch eine genauere Ermittlung dieser Werte als mittels existierender Standardverfahren möglich ist. Die Arbeit verfolgt zwei Stoßrichtungen – die Analyse niederwasserrelevanter Prozesse auf der regionalen Skale und Vergleiche von Regionalisierungsverfahren zur Ermittlung der für österreichische Verhältnisse am besten geeigneten Methoden. Ein umfangreicher Datensatz wird verwendet, der Österreich zum großen Teil abdeckt. Er besteht aus 325 Einzugsgebieten mit Gebietsflächen zwischen 7 und 963 km². Für alle Gebiete liegen kontinuierliche Abflußbeobachtungen über den Zeitraum 1977 bis 1996 vor. Die betrachtete Niederwasserkenngröße ist das 5% Quantil der Dauerlinie Q_{95}. Sie entspricht jener Abflußmenge, die an 95% aller Tage der Meßperiode überschritten wurde.

In Abschnitt 2 werden drei Saisonalitätsindizes in Hinblick auf ihr Potential für die Regionalisierung von Niederwasserkennwerten untersucht. Die betrachteten Indizes sind das Saisonalitätshistogramm (SH), welches die monatliche Verteilung von Niederwässern beschreibt, ein zyklischer Saisonalitätsindex (SI), welcher das mittlere Auftreten und die Variabilität des Auftretens von Niederwässern beschreibt, sowie die Saisonalitätsrate (SR), die als Quotient aus Sommer- und Winterniederwasserabfluß definiert wird. Die Saisonalitätsanalyse fußt auf der Überlegung, daß Sommer- und Winterniederwässer durch grundlegend unterschiedliche hydrologische Prozesse hervorgerufen werden. In einem ersten Schritt werden die drei Saisonalitätsindizes für das Untersuchungsgebiet verglichen. Ihre räumlichen Muster lassen sich hydrologisch gut interpretieren. In einem zweiten Schritt werden verschiedene Kombinationen der Indizes für eine Klassifikation in zwei, drei und acht Regionen herangezogen. In einem dritten Schritt wird der Wert der Saisonalitätsindizes für die Niederwasserregionalisierung untersucht, indem die drei saisonalitätsbasierten Klassifikationen in die Regionalisierung einbezogen werden. Als Vergleichsmaß dient die mittels Kreuzvalidierung ermittelte Modellgüte der Mehrfachregressionsmodelle zwischen

mittels Kreuzvalidierung ermittelte Modellgüte der Mehrfachregressionsmodelle zwischen Niederwässern und Einzugsgebietsmerkmalen. Die Klassifikation des Untersuchungsgebietes in drei Regionen mit getrennten Regressionen für jede Region erweist sich als am besten geeignet, gefolgt von der Klassifikation des Untersuchungsgebietes in zwei Regionen mit getrennten Regressionen für jede Region. Ein globales Regressionsmodell erzielt die geringste Modellgüte, und ein globales Regressionsmodell mit unterschiedlichen Kalibrierungskoeffizienten für jede der acht Regionen erzielt nur geringfügig bessere Resultate. Die Ergebnisse belegen die Vorteile einer getrennten Regionalisierung in Teilgebieten gegenüber einem globalen Modell, die mit einer besseren Wiedergabe der Beziehungen zwischen Niederwasserabfluß und Einzugsgebietsmerkmalen zusammenhängen.

Um den Wert der Saisonalitätsindizes einzuordnen und die getrennte Regionalisierung von Einzelregionen mittels regionaler Mehrfachregression zu erweitern, werden in Abschnitt 3 vier Methoden zur Gruppenbildung von Einzugsgebieten in Hinblick auf ihren Wert für die Bestimmung von Niederwasserabflußspenden q_{95} in unbeobachteten Gebieten untersucht. Die betrachteten Gruppierungsmethoden sind die Residuenmustermethode, die gewichtete Clusteranalyse, der Regressionsbaum und die Gruppierung in acht Saisonalitätsregionen. Die Regression zwischen q_{95} und Gebietskennwerten erfolgt getrennt für jede Gruppe. Die Güte der einzelnen Methoden wird mittels Kreuzvalidierung verglichen, wodurch eine zuverlässige Angabe der Genauigkeit für Gebiete ohne Abflußmessungen möglich ist. Die Gruppierung auf Basis von Saisonalitätsregionen erweist sich als die beste Methode. Das darauf basierende regionale Regressionsmodell erklärt 70% der räumlichen Varianz von q_{95}. Die hohe Güte dieser Methode dürfte mit den markanten Unterschieden der Niederwasserprozesse im Untersuchungsgebiet zusammenhängen. Winterniederwässer sind eine Folge der Retention von Niederschlägen in der saisonalen Schneedecke, während Sommerniederwässer eine Folge des relativ großen Bodenfeuchtedefizits in Einzugsgebieten des Flachlands im Sommer sind. Die Gruppierung mittels Regressionsbaum erweist sich als zweitbeste Methode (64% erklärte Varianz), die Güte der Residuenmustermethode ist ähnlich (63% erklärte Varianz). Die Gruppierung mittels gewichteter Clusteranalyse erklärt nur 59% der räumlichen Varianz von q_{95} und erzielt somit nur eine geringe Verbesserung gegenüber dem globalen Regressionsmodell, das keinerlei Gruppierung einbezieht (57% erklärte Varianz). Eine Analyse der Residueneigenschaften aller Methoden belegt ebenfalls die Vorteile der Gruppierung auf der Basis von Saisonalitätsregionen, zeigt aber auch, dass alle Methoden zur Unterschätzung der Niederwasserspenden q_{95} in sehr nassen Einzugsgebieten neigen.

Abschnitt 4 untersucht Methoden der Niederwasserregionalisierung bei Vorliegen von kurzen Abflußreihen an der betrachteten Stelle. Verschiedene Methoden zur Korrektur der durch Klimaschwankungen eingetragen Fehler bei der Bestimmung des Niederwasserabflusses Q_{95} mittels kurzer Abflußreihen werden verglichen. Die Methoden zur Klimakorrektur bestehen aus zwei Schritten, Wahl des Referenzpegels und Kennwertkorrektur (engl.: record augmentation), und verwenden Information naheliegender Pegel mit längeren Abflußbeobachtungen. Die Genauigkeit der Methoden wird durch den Vergleich der korrigierten Abflusskennwerte hypothetisch verkürzter Reihen mit den Schätzungen aus den vollständigen 20 jährigen Abflußbeobachtungen am selben Pegel bestimmt. Die Ergebnisse zeigen, dass eine Methode, welche den stromabwärts liegenden Pegel am selben Gewässer verwendet, die genauesten Werte liefert. Die Wahl des Referenzpegels aufgrund ähnlicher Einzugsgebietsmerkmale oder aufgrund der Kreuzkorrelation von Jahresniederwässern führt zu deutlich ungenaueren Ergebnissen. Die Wahl der Korrekturmethode bei einem bestimmten Referenzpegel beeinflusst zwar das Ergebnis, fällt insgesamt aber weniger ins Gewicht als die Wahl des Referenzpegels. Die Genauigkeit der geschätzten Niederwasserkennwerte kann durch die Klimakorrekturmethoden für Abflußreihenlängen unter fünf Jahren drastisch erhöht werden. Das Bestimmtheitsmaß der Schätzung von q_{95} steigt von 63 auf 89% für ein Jahr Abflußbeobachtung und von 86 auf 93% für drei Jahre Abflußbeobachtung wenn die Klimakorrektur mittels stromabliegenden Referenzpegel erfolgt. Für eine Beobachtungsdauer von fünf Jahren oder mehr ist der Wert der Klimakorrektur wesentlich geringer. Ein Verfahren, das Einzelmessungen des Abflusses während Niederwasserperioden verwendet, besitzt nur eine geringfügig größere Genauigkeit bei der Ermittlung von q_{95} als ein einfaches Regionalisierungsverfahren. Vergleiche mit einem in Abschnitt 3 analysierten detaillierten Regionalisierungsverfahren zeigen, daß, im Durchschnitt über das gesamte Untersuchungsgebiet, kontinuierliche Abflussbeobachtungen über ein Jahr selbst das detaillierte Regionalisierungsverfahren klar übertreffen, während Einzelmessungen deutlich ungenauere Werte als das detaillierte Regionalisierungsverfahren ergeben.

Die Analyse von Niederwasserprozessen und der Vergleich von Regionalisierungsverfahren im Rahmen dieser Dissertation belegen, daß Prozeßverständnis jedenfalls zur Regionalisierung von Niederwasserkenngrößen beitragen kann, und damit eine genauere Ermittlung der Niederwasserabflüsse als mittels existierender Standardverfahren möglich ist.

Summary

Accurate estimates of low flow characteristics are needed for a range of purposes in water resources management and engineering including environmental flow requirements, water uses and discharges into streams, and hydropower operation. For sites where no long term stream flow records are available, regionalisation techniques can be used to infer the low flow characteristics from other catchments where stream flow data have been collected. The hypothesis put forward in this thesis is that process understanding, if in a simplified way, can assist in regionalising low flow characteristics to provide more accurate estimates than existing standard methods. The analyses proceed along two main strands, unravelling the processes that drive low flows at the regional scale, and comparing regionalisation methods in order to identify the most suitable method for a setting such as Austria. The comparisons are made on a comprehensive Austrian data set, including 325 sub-catchments that range in area from 7 to 963 km^2. The data set covers a continuous period from 1977 to 1996 for all stream flow records. The low flow characteristic chosen here is the Q_{95} low flow which is the discharge that is exceeded on 95% of all days of the measurement period.

In section 2, three seasonality indices are examined for their potential in regionalising low flows. The indices are seasonality histograms (SH) 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. In a first step, the three seasonality indices are compared for the study region. 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 regionalisation is examined by comparing the cross validation performance of multiple regressions between low flows and catchments characteristics. The regressions make use of the three seasonality based classifications. The results indicate that grouping the study domain into three regions using separate regressions in each region gives the best performance, followed by grouping the study area into two regions and separate regressions in each region. 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.

To put the predictive power of the seasonality indices into context and to extend the analysis to the case of separate regression models in each of the regions, four catchment grouping methods are compared in Section 3 in terms of their performance in predicting specific low flow discharges q_{95}. The grouping methods are the residual pattern approach, weighted cluster analysis, regression trees and the grouping into eight seasonality regions. For each group, a regression model between catchment characteristics and q_{95} is fitted independently. The performance of the methods is assessed by leave-one-out cross-validation of the regression estimates which emulates the case of ungauged catchments. Results indicate that the grouping based on seasonality regions performs best and explains 70% of the spatial variance of q_{95}. The favourable performance of this grouping method is likely related to the striking differences in seasonal low flow processes in the study domain. Winter low flows are associated with the retention of solid precipitation in the seasonal snow pack while summer low flows are related to the relatively large moisture deficits in the lowland catchments during summer. The regression tree grouping performs second best (explained variance of 64%) and the performance of the residual pattern approach is similar (explained variance of 63%). The weighted cluster analysis only explains 59% of the spatial variance of q_{95} which is only a minor improvement over the global regression model, i.e. without using any grouping (explained variance of 57%). An analysis of the sample characteristics of all methods suggests that, again, the grouping method based on the seasonality regions has the most favourable characteristics although all methods tend to underestimate specific low flow discharges in the very wet catchments.

Section 4 explores the low flow regionalisation case where a short stream flow record is available at the site of interest. A number of methods of adjusting Q95 estimates from short stream flow records for climate variability are compared. The climate adjustment methods consist of two steps, donor site selection and record augmentation, and use information from nearby sites with longer stream flow records. The accuracy of the methods is assessed by comparing the adjusted estimates from hypothetically shortened records with estimates from the full 20 year record at the same site. The results indicate that the downstream donor selection method performs best on all scores. The catchment similarity and correlation donor selection methods do not perform as well. The relative performance of the record augmentation methods depends on the donor selection method but, overall, the choice of record augmentation method is less important than the choice of the donor site. The value of the climate adjustment methods is very significant for record lengths shorter than 5 years. The coefficient of determination of q95 specific low flows increases from 63 to 89% for one year records, and from 86 to 93% for three year records when adjusting the estimates by the downstream site method. For five years or more, the value of the climate adjustment methods is much smaller. A method that uses spot gaugings of stream flow during a low flow period only performs slightly better than a simple regionalisation procedure in terms of predicting Q95 at an otherwise ungauged site. Comparisons with more sophisticated regionalisation procedures from Section 3 suggest that, on average over the study region, one year of continuous stream flow data clearly outperforms the more sophisticated regionalisation method while the spot gauging method provides less accurate low flow estimates than the sophisticated regionalisation method.

Danksagung

Mein großer Dank gilt den Betreuern meiner Dissertation, Prof. Günter Blöschl und Prof. Harald Strelec, sowie allen die mich bei meiner Arbeit in vielfältiger Weise unterstützt haben.

Danke!

Table of contents

1. Introduction

Accurate estimates of low flow characteristics are needed for a range of purposes in water resources management and engineering. They are used in water quality management applications including discharge permits, and in siting treatment plants. Low flow estimates are also used in water supply planning to determine allowable water transfers and withdrawals. Other applications of low flow estimates include determination of minimum downstream release requirements from hydropower (Stedinger et al, 1992; Gustard et al., 2004). Low flow characteristics are best estimated from long-term stream flow data but for sites where these data are unavailable hydrological regionalisation techniques can be used to infer them from other catchments where stream flow data have been collected.

The regionalisation of low flow characteristics is usually based on some sort of regression model between the low flow characteristic of interest and catchment characteristics that are available for ungauged sites (e.g., Vogel and Kroll, 1992; Gustard et al., 1992; Schreiber and Demuth, 1997; Skop and Loaiciga, 1998). If the study domain is large, different low flow processes may prevail in different parts of the domain. It has therefore been suggested to use some representation of the driving processes, if in a simplified way, in regionalising low flows. There are numerous ways of making use of various types of information on the hydrological processes at the regional scale. A number of authors have proposed to split the domain into regions and apply a regression relationship to each of the regions independently (e.g., Gustard and Irving 1994; Clausen and Pearson 1995; Aschwanden and Kan, 1999). In some instances it is clear how to group a domain into regions of approximately uniform low flow processes but, more often, the choice is far from obvious. A number of methods of identifying homogeneous regions have therefore been put forward in the literature in the context of low flow regionalisation (e.g. Hayes, 1992; Nathan and McMahon, 1999; Laaha, 2002). All of these methods use low flow data and most of them use catchment characteristics as well. Among the various indicators to low flow processes, low flow seasonality appears to be particularly promising in a country such as Austria, given that the hydrologic regimes exhibit very apparent spatial patterns (Merz et al.,1999; Piock-Ellena et al., 2000). If stream flow records do exist at the site of interest, but are too short for obtaining reliable local estimates, the regionalisation methods can be extended to make use of these records. While these short records are unlikely to provide the full information of long records it is clear that they do provide some information which may be used in estimating the long term low flow characteristics for these stream gauge locations (Vogel and Kroll, 1991).

Due to climatic variability and other sources of variability that occur over short time scales, low flow characteristics estimated from a few years of stream flow data deviate from the long-term average, and some adjustment for climate variability is needed to make them compatible with low flow characteristics obtained from longer records. These adjustment methods can be thought of as a combination of low flow regionalisation methods with short stream flow records at the site of interest.

While much work has been done in the literature on low flow regionalisation, comprehensive comparisons of process based regionalisation methods in a hydrological setting such as Austria, to the best of my knowledge, do not exist. The thesis put forward here is that process understanding, if in a simplified way, can assist in regionalising low flow characteristics to provide more accurate estimates than existing standard methods. The analyses proceed in two main strands, unravelling the processes that drive low flows at the regional scale, and comparing regionalisation methods in order to identify the most suitable method for a setting such as Austria.

The comparisons will be made on a comprehensive Austrian data set, including 325 catchments with close to natural flow conditions. The data set covers a continuous period from 1977 to 1996 for all stream flow records. The low flow characteristic chosen here is the Q_{95} low flow which is 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 to multiple topics of water resources management (see, e.g., Kresser et al., 1985; Gustard et al., 1992; Smakhtin, 2001).

This thesis consists of three main parts. The seasonality of low flows is analysed in Section 2, in order to gain insights into the most relevant low flow processes and their spatial distribution in Austria. This section also examines the potential of three seasonality indices in low flow regionalisation. Section 3 consists of a comparison of a number of catchment grouping methods that represent the regional heterogeneity of low flow processes. One the groupings makes use of seasonality indices. The predictive performance of the various grouping and regionalisation methods is examined by cross validation which provides a faithful measure of the predictive performance of the methods for the case of ungauged catchments. Section 4 finally explores the case where short stream flow records are available at the site of interest. Various methods of climate adjustment that make use of some of the regionalisation concepts of the previous sections are compared, again, in terms of their predicitive performance in a cross validation mode.

2 Seasonality indices for regionalising low flows

2.1 Introduction

Many branches of water resources management need accurate estimates of low flows. If suitable measurements are not available, the low flow characteristics need to be estimated from regional information by some sort of hydrological regionalisation 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 United Kingdom (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 sub-regions. Nathan and McMahon (1990) compared several multivariate statistical approaches based on physical catchment characteristics to obtain possible groupings of hydrologically similar stations which can serve as a basis for fitting separate regionalisation models to data. However, they stated that "... groupings obtained are very sensitive to the initial choice of predictor variables" and hence are highly subjective.

Seasonality has attracted a lot of attention in the literature recently to assist in the regionalisation 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 regionalisation. The regionalisation 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 regionalisation 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 visualise the spatial patterns of the seasonalities of floods and other hydrologic 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 I believe that the analysis of low flow seasonality should be useful for low flow regionalisation. An application of a low flow seasonality index 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 sub-regions.

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 to 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, regionalisation may take advantage of a separation of summer and winter low flows (Tallaksen and Hisdal, 1997; Laaha, 2000). Because of the same reasons, seasonality is also potentially useful for regionalising annual low flows. There are different ways of incorporating seasonality in regionalisation 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.

The aim of this section is to investigate the value of seasonality indices for regionalising low flows. As a regionalisation model, I 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. I compare the models for the

95% quantile of specific discharges (q95) and I also examine the specific low flow discharge of the summer and winter periods (q95s, q95w).

This section is organised as follows: Section 2.2 summarises the data and the disaggregation method used in this study for calculating specific low flow discharges for residual catchments. Section 2.3 presents different seasonality measures and shows how sub-regions of similar seasonality can be isolated. The value of these seasonality measures for regionalisation is investigated in sections 2.4 and 2.5. Section 2.4 presents the method of regionalisation and cross-validation used in this study and describes how seasonality measures have been considered in regression modelling. The results are given in section 2.5, followed by a discussion and conclusions in sections 2.6 and 2.7, respectively.

2.2 Data

2.2.1 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) (Fig. 2.1). Elevations range from 117 to 3798 m a.s.l.. Geological formations vary significantly, too. Austria has a varied climate with mean annual 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 hydrologic regimes (Kresser, 1965) and low flows exhibit important regional differences in terms of their quantity and their seasonal occurrence (Laaha and Blöschl, 2003).

2.2.2 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, I have chosen to give priority to a consistent observation period to make all records comparable in terms of climatic variability.

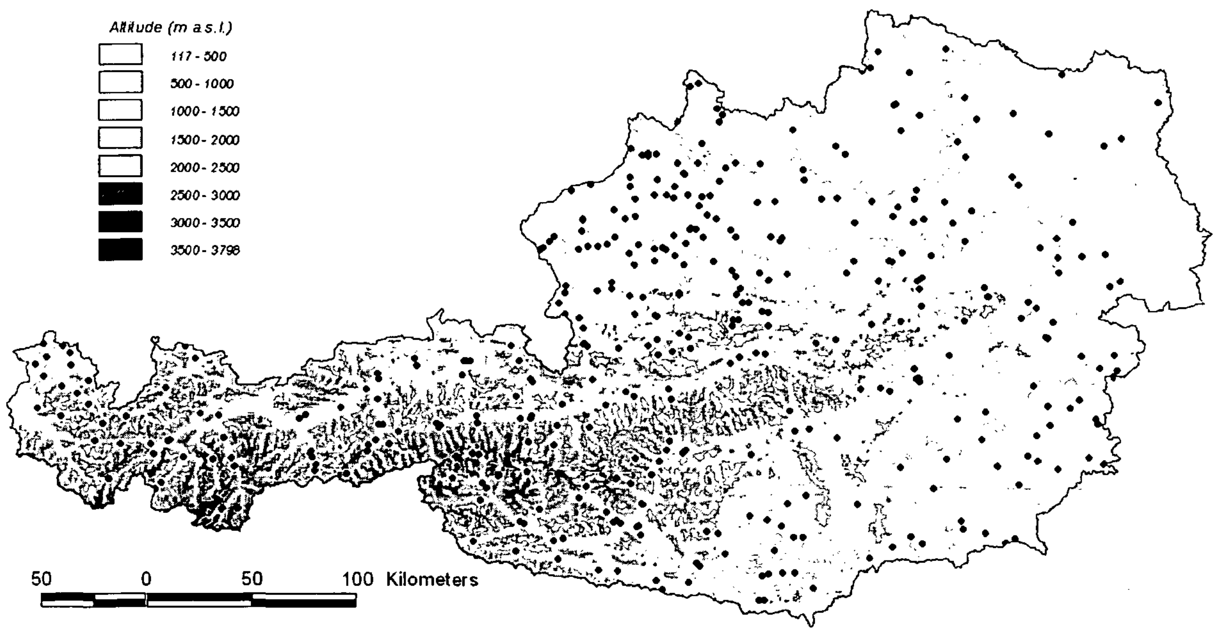

Fig. 2.1. Topography and stream gauging network in Austria. Points indicate location of gauges used in this study.

2.2.3 Disaggregation of nested catchments

Nested catchments were split into sub-catchments between subsequent stream gauges based on the hierarchical ordering of gauges presented in Laaha and Blöschl (2003). The advantage of using sub-catchments rather than complete catchments is that the application of regionalisation techniques to small ungauged catchments is more straightforward. Also, discharge characteristics of nested catchments are statistically not independent and disaggregation into sub-catchments 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.

2.2.4 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 (see, e.g., Kresser et al., 1985; Gustard et al., 1992; Smakhtin, 2001). For gauged catchments without an upstream gauge I calculated the Q_{95} low flow quantile directly from the stream flow data. For sub-catchments I calculated Q_{95} from the differences of stream flows at the two gauges. To make the low flow characteristic more comparable across scales I standardised Q_{95} by the catchment area. The resulting specific low flow discharges q_{95} ($l \cdot s^{-1} \cdot km^{-2}$) 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 q95 in Austria is presented in Fig. 2.2. The pattern of calculated low flow characteristics q95 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 q95 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 q95 for hilly terrain and the lowlands range from 0 to 8 $l.s^{-1}.km^{-2}$.

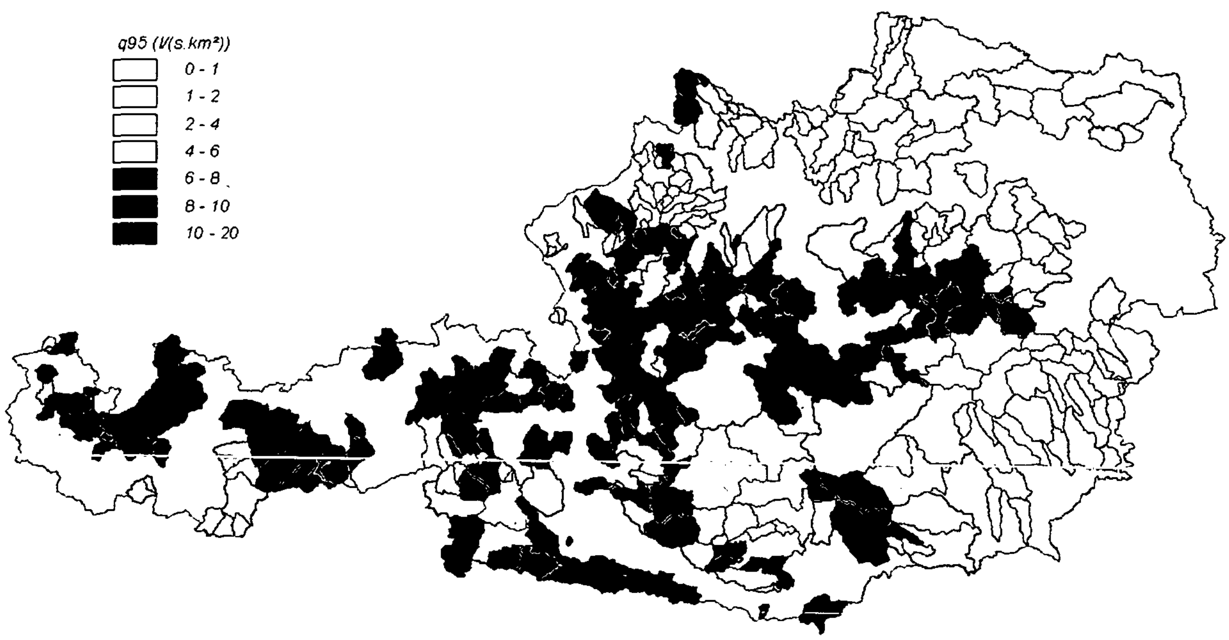

Fig. 2.2: Specific low flow discharge q95 [$l.s^{-1}.km^{-2}$] from runoff data observed in 325 sub-catchments in Austria. Alpine catchments show higher values and a larger variability.

2.2.5 Catchment characteristics

I used 31 physiographic catchment characteristics in the low flow regionalisation in this section (Table 2.1). They relate to catchment area (A), topographic elevation (H), topographic slope (S), precipitation (P), geology (G), land use (L), and drainage density (D). All percent values with the except of mean slope (S_M) relate to the area covered by a class relative to the total catchment area. Some of the catchment characteristics had to be adapted from the original sources to make them more useful for regionalisation. For instance, the original classification of the metallurgic map used here distinguishes 670 geological classes from which I derived 9 hydrogeological classes I deemed relevant for low flow regionalisation. One of them is termed source region which is the percent area where the density of springs is large. In a similar vein, I 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 regionalisation. Three precipitation characteristics of average annual, summer and winter precipitation from 1977 to 1996 estimated by the regionalisation 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 sub-catchment 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 2.1.

Table 2.1. Statistical summary of the characteristics of the 325 sub-catchments used in this section. Units were chosen in a way to give similar ranges for all characteristics.

Acronym	Variable description	Units	Min.	Mean	Max.
A	Sub-catchment area	10^1 km²	0.70	15.22	96.30
H_0	Altitude of streamgauge	10^2 m	1.59	5.93	22.15
H_+	Maximum altitude	10^2 m	2.98	17.48	37.70
H_R	Range of altitude	10^2 m	0.81	11.56	30.06
H_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_{MO}	Moderate slope	%	0.00	46.18	93.00
S_{ST}	Steep slope	%	0.00	25.78	80.00
P	Average annual precipitation	10^2 mm	4.67	10.71	21.03
P_S	Average summer precipitation	10^2 mm	2.94	6.47	12.08
P_W	Average winter precipitation	10^2 mm	1.55	4.24	8.95
G_B	Bohemian Massif	%	0.00	9.70	100.00
G_Q	Quaternary sediments	%	0.00	6.22	94.50
G_T	Tertiary sediments	%	0.00	15.91	100.00
G_F	Flysch	%	0.00	6.90	100.00
G_L	Limestone	%	0.00	25.21	100.00
G_C	Crystalline rock	%	0.00	25.44	100.00
G_{GS}	Shallow groundwater table	%	0.00	1.74	48.00
G_{GD}	Deep groundwater table	%	0.00	7.51	79.80
G_{SO}	Source region	%	0.00	1.23	35.20
L_U	Urban	%	0.00	0.67	14.50
L_A	Agriculture	%	0.00	21.37	97.30
L_C	Permanent crop	%	0.00	0.12	20.30
L_G	Grassland	%	0.00	20.10	71.70
L_F	Forest	%	0.00	47.25	100.00
L_R	Wasteland (rocks)	%	0.00	8.45	81.20
L_{WE}	Wetland	%	0.00	0.10	16.40
L_{WA}	Water surfaces	%	0.00	0.42	18.20
L_{GL}	Glacier	%	0.00	1.37	43.80
D	Stream network density	10^2 m/km²	1.18	8.01	13.98

2.3 *Seasonality analysis*

2.3.1 Seasonality measures

2.3.1.1 Seasonality ratio (SR)

The first approach addressing low flow seasonality presented in this study is based on the recognition that summer and winter low flows are subject to important differences in the underlying hydrological processes. Thus, I 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 April 1st to November 30th) and winter discharge series (December 1st to March 31st) and characteristic values for summer low flows (q95s) and winter low flows (q95w) were calculated for each sub-catchment. From this, the ratio *SR* of q95s and q95w was calculated:

$$SR = {q95s}/{q95w} \tag{1}$$

A map of SR for Austria is presented in Fig. 2.3. 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 seasonalities 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 characterised by weak seasonality appears. The plot appears to be useful for visualising the patterns of summer and winter low flows.

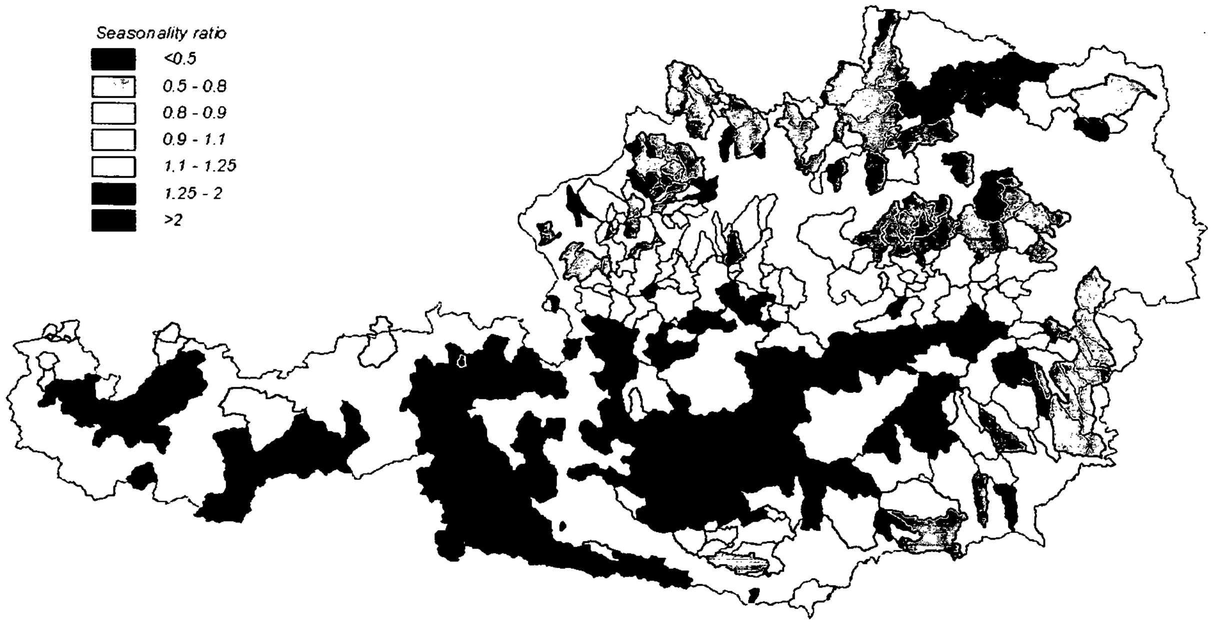

Fig. 2.3: Ratio of summer and winter low flow discharges (SR) for 235 sub-catchments in Austria. SR > 1 indicates a winter low flow regime, SR < 1 indicates a summer low flow regime.

2.3.1.2 Seasonality index (SI)

I 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 Q95, by means of circular statistics (Mardia, 1972). The first parameter, θ, is the mean day of occurrence, measured in radians, and is a measure of the average seasonality of low flows. θ takes values between 0 to 2π; 0 relating to January 1st, $\pi/2$ relating to April 1st, π relating to July 1st, and $3\pi/2$ relating to October 1st. The second 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 0 corresponding to no seasonality (low flow events are uniformly distributed over the year).

For each sub-catchment, the days on which discharge was smaller than Q95 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 which can be displayed as a vector on the unit circle. Its directional angle, in radians, is given by:

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

The arithmetic mean of Cartesian coordinates x_θ and y_θ of the single days j is defined as:

$$x_\theta = \frac{1}{n}\sum_j \cos(\theta_j) \tag{3}$$

$$y_\theta = \frac{1}{n}\sum_j \sin(\theta_j)$$

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

$$\theta = \arctan\left(\frac{y_\theta}{x_\theta}\right) \qquad 1^{st} \text{ and } 4^{th} \text{ quadrant: } x>0 \tag{4}$$

$$\theta = \arctan\left(\frac{y_\theta}{x_\theta}\right) + \pi \qquad 2^{nd} \text{ and } 3^{rd} \text{ quadrant: } x<0$$

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

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

The length r of the mean vector is as 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 (Fig. 2.4) 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.

2.3.1.3 Seasonality histogram (SH)

The seasonality histogram (Laaha, 2002) allows a more detailed description of the seasonal distribution of low flows than the seasonality index. Again, this description is based on the Julian date of all days when the discharge of a catchment (or the differential discharge of a sub-catchment) falls below the threshold Q95. Histograms based on monthly classes were plotted from these data. Hence, the seasonality histogram illustrates the occurrence of low flows in each month and provides supplemental information to the seasonality index. 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.

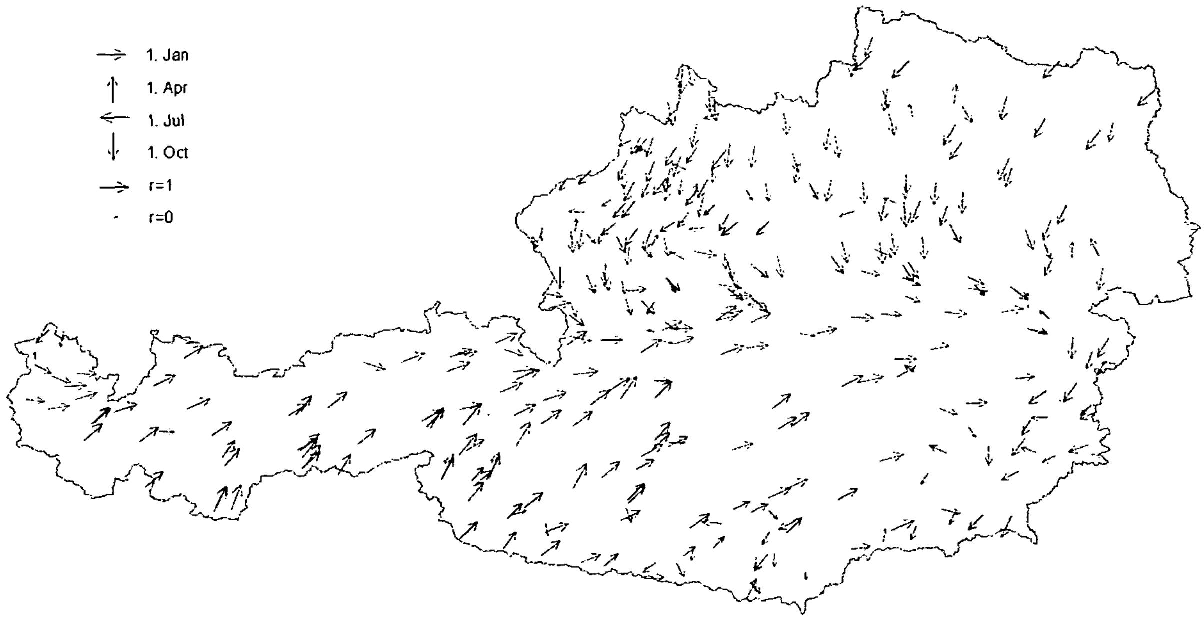

Fig. 2.4: Seasonality index of 325 sub-catchments in Austria. Long arrows indicate strong seasonality and their direction represents the mean day of occurrence of specific low flow discharges less than q95.

2.3.2 Delineation of homogeneous regions

2.3.2.1 Cluster analysis of SH

Saisonality histograms may be regarded as a multidimensional measure of low flow seasonality consisting of 12 variables, each of them representing the monthly occurrence frequency of low flows (Laaha, 2002). To delineate regions that are homogeneous in terms of seasonality, partitive cluster analysis PAM (partitioning around medoids, see Kaufmann and Rousseeuw, 1990) was applied to classify seasonality histograms automatically. PAM is an exhaustive partitioning method by which the ensemble of catchments is classified into several exclusive subsets. Compared to the classical k-means approach to exhaustive partitioning (Hartigan, 1975), PAM has the following useful features: (a) cluster centres (medoids) are automatically chosen by the algorithm; (b) it is also adapted to more robust distance metrics than Euclidean distances; (c) it provides a novel graphical display, the silhouette plot. This is an ordered representation of the silhouette width (Kaufman and Rousseeuw, 1990) of each histogram, which gives a relative measure of the similarity of each histogram to the cluster centre with respect to the similarity to the next suitable cluster centre. The silhouette plot can be applied to select the optimum number of clusters, which is related to the maximum average silhouette width among classifications into different numbers of clusters.

The analysis led to an optimal number of two clusters. The graphical representation of catchments by the first two principal components of seasonality histograms (Fig. 2.5 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 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 (Fig. 2.5 right).

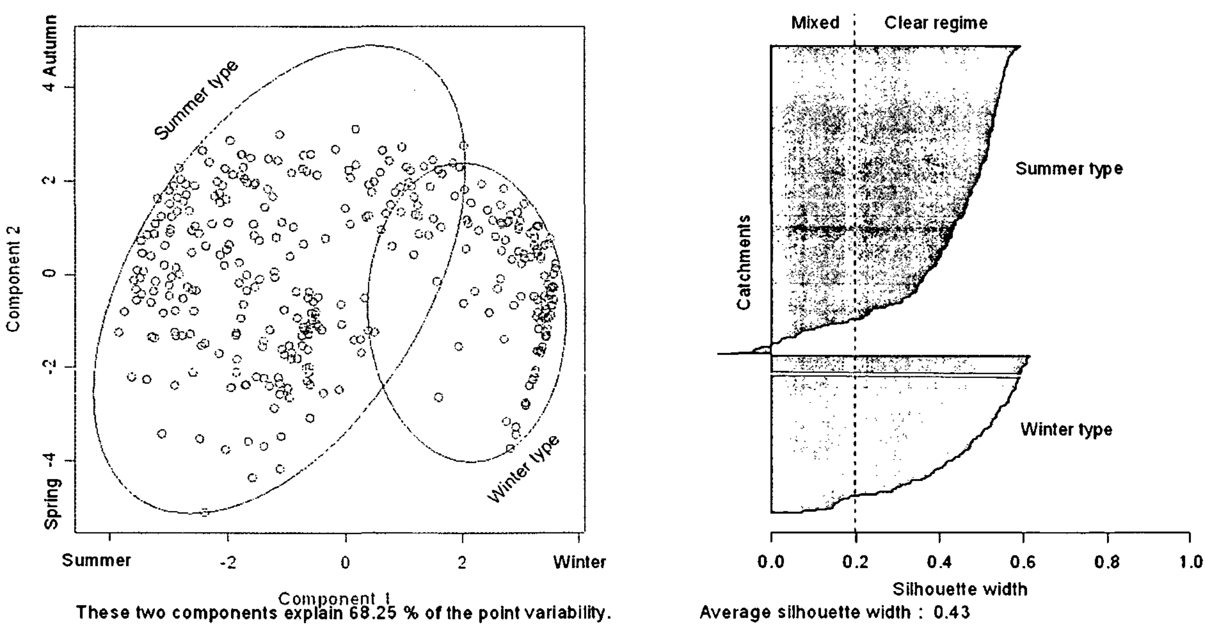

Fig. 2.5: Left panel: graphical representation of cluster membership of catchments (points) by the first two principal components of seasonality histograms. The big ellipse contains catchments of the summer type cluster, the smaller ellipse contains catchments of the winter type cluster. Right panel: Determination of catchments that exhibit weak or mixed low flow regimes by silhouette width, illustrated by the silhouette plot.

The location of summer and winter type catchments can be seen from Fig. 2.6, indicating two contiguous regions of different seasonality. Winter low flows typically occur in higher altitudes of the Alps, summer low flows typically occur in the lower parts of Austria. The alternative classification into three regime types is shown in Fig. 2.7. 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 seasonality ratio (Fig. 2.3), but instead of the gradual representation of seasonality by the seasonality ratio, the cluster analysis results in a mutually exclusive classification of catchments. Cluster analysis of SH appears to be an appropriate basis for regionalising low flows separately for catchments that exhibit typical summer and winter regimes.

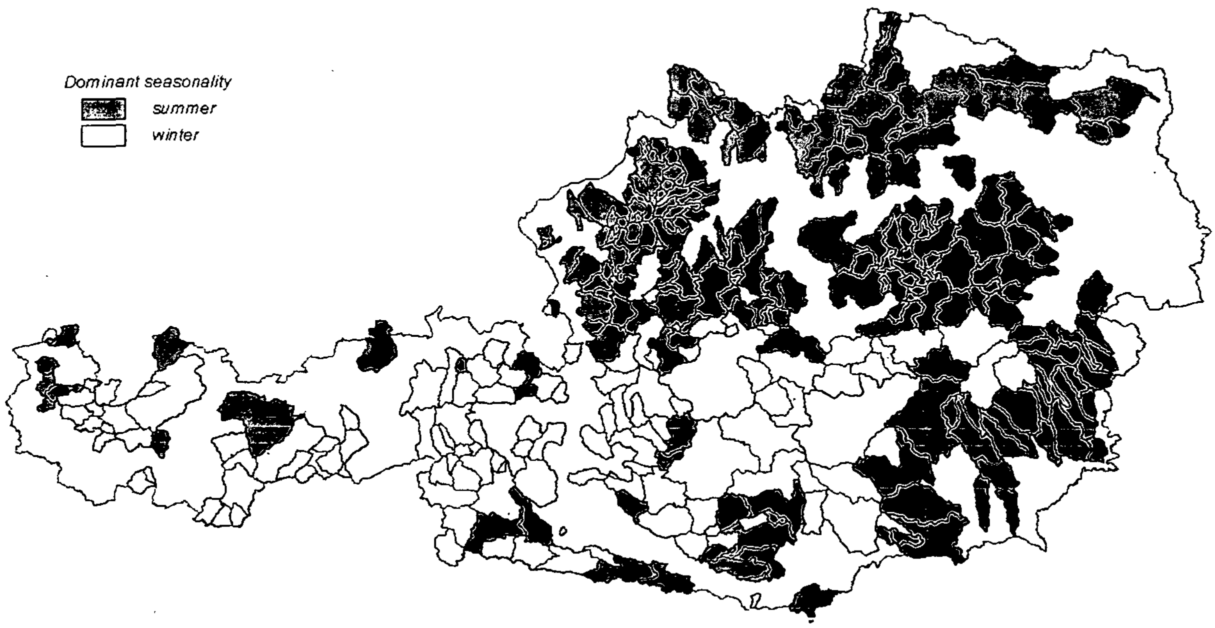

Fig. 2.6: Classification of 325 sub-catchments in Austria into two regime types (summer regime and winter regime).

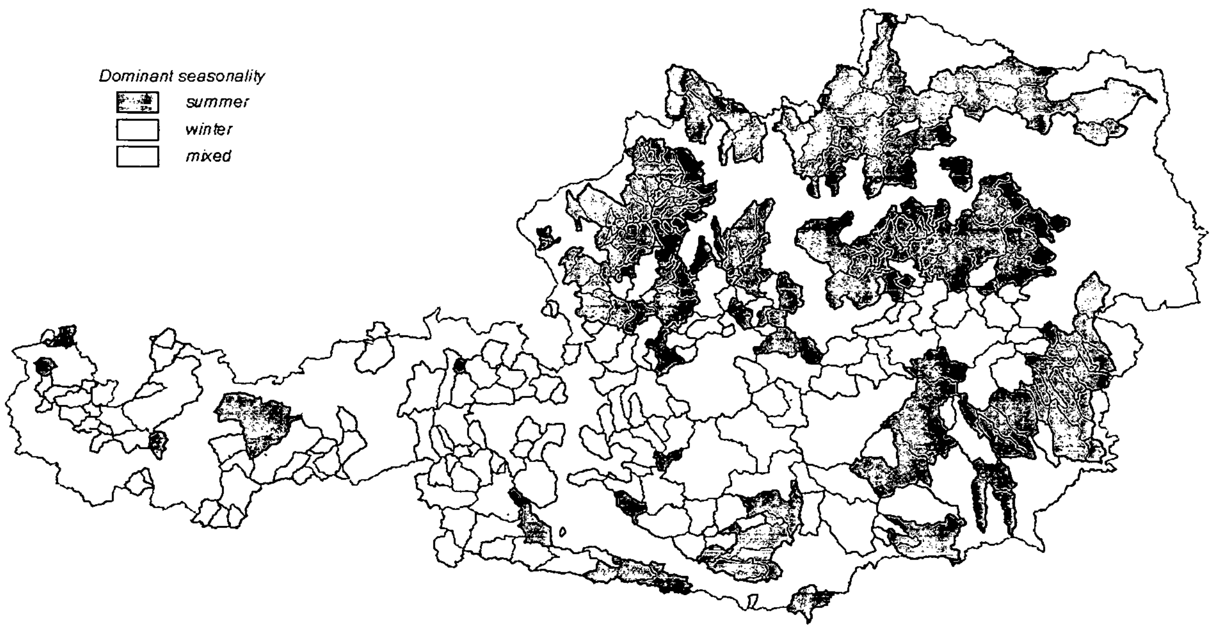

Fig. 2.7: Classification of 325 sub-catchments in Austria into three regime types (summer regime, winter regime, mixed regime).

2.3.2.2 Visual grouping based on different seasonality measures

Based on an interpretation of the seasonality index and seasonality histograms, regions of approximately homogeneous seasonality have been identified visually. This approach is more subjective than automatic classification, but allows me to take additional information, such as breaklines of the relief, into account. 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 seasonality histograms led to a correction and refinement of 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.

Fig. 2.8 presents the seasonality regions so obtained, which correspond to the types of seasonality histograms presented in Fig. 2.9. 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 and mean seasonality vary with the elevation of 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 seasonality ratio and can be considered a refined classification compared to that obtained by cluster analysis. Since regions appear well interpretable in terms of low flow processes, there is likely some potential for regionalisation in the approach.

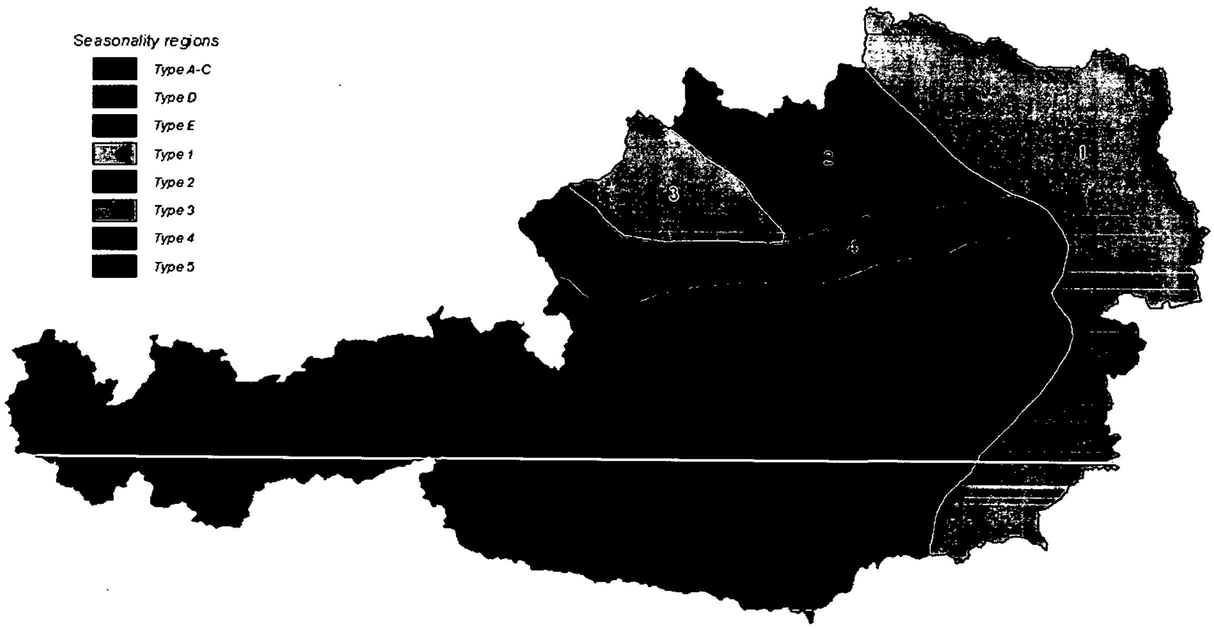

Fig. 2.8: Regions of approximately homogeneous seasonality in Austria. Letters refer to winter low flow types, numbers to summer low flow types (see Fig. 2.9).

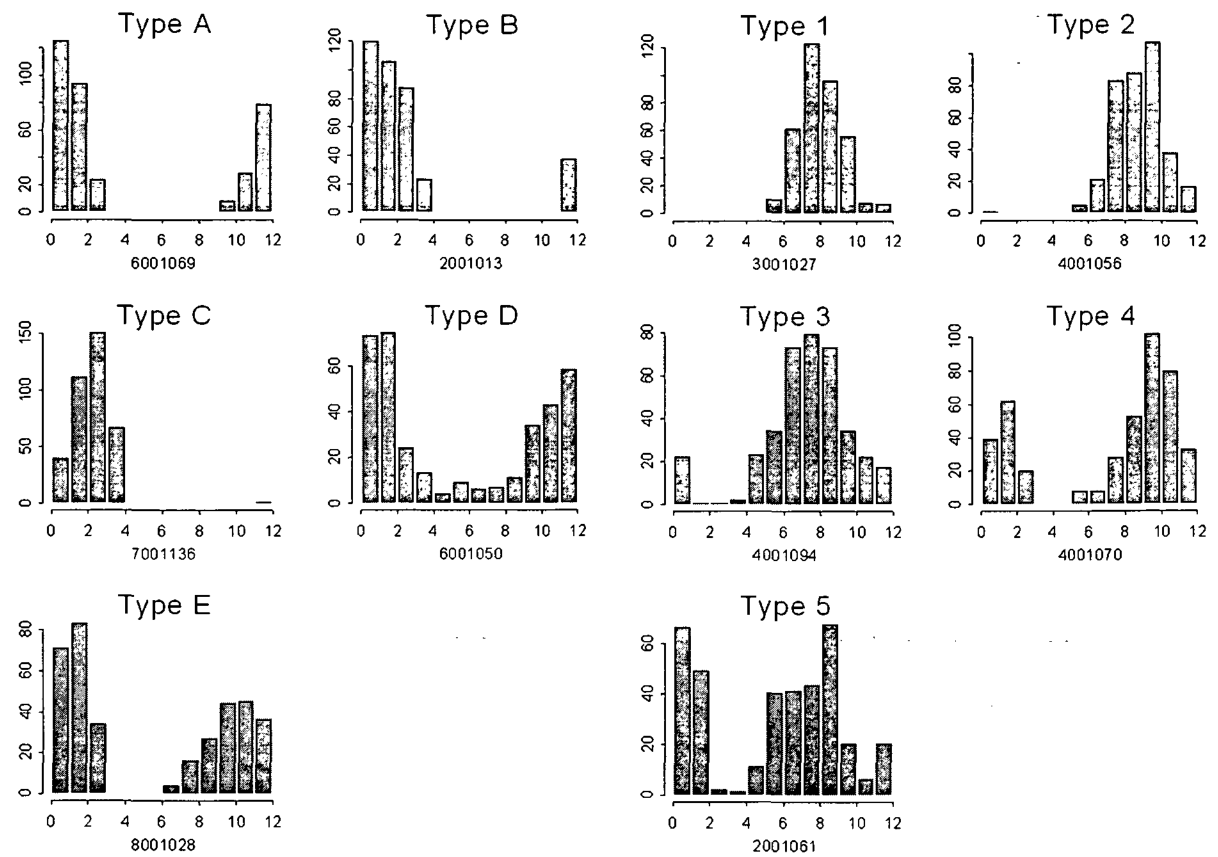

Fig. 2.9: Seasonality histograms: Non-exceedance frequencies of Q_{95} for each month for a typical catchment in each region. Letters relate to winter low flows, numbers relate to summer low flows (see Fig. 2.8).

2.4 Method of regionalisation and cross-validation

2.4.1 Multiple regression

The regionalisation methods used in this study are multiple linear regression models between specific low flow discharge q95 and physical catchment characteristics. Physical catchment properties are represented by 31 catchment characteristics - a number which is relatively large compared to other regionalisation studies reported in the literature. These catchment characteristics are subject to inter-correlations and multicollinearity as mentioned above. Rather than performing a selection of the most important variables prior to regionalisation, I used a stepwise regression approach. The stepwise regression procedure used *Mallow's* C_p (Weisberg, 1985, p. 216) as the criterion of optimality, which was calculated as:

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

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 penalised by the second term which adds the number of coefficients p minus the number of catchments n. C_p is therefore a penalised selection criterion which 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 minimise the C_p criterion. After each step it is tested if replacing one of the variables by any remaining catchment characteristic will further decrease the criterion. 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 values and may not represent the bulk of data. My 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). These leverage points 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 ten to give similar magnitudes in terms of their ranges (see Table 2.1). Since linear regression is scale invariant (Weisberg, 1985, p.185) the regression models, including their residual statistics, remain unaffected by the rescaling but the numerical stability is improved.

2.4.2 Regionalisation methods examined

2.4.2.1 4.2.1 Regionalisation of q95 low flows

2.4.2.1.1 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 bench mark case against which to test the seasonality based regionalisation methods.

2.4.2.1.2 Grouping into two regions and separate regressions in each region

In the second approach, two regionally restricted regression models were separately fitted for two contiguous regions consisting of summer-dominated and winter-dominated catchments, respectively. This corresponds to the original classification of catchments obtained by the cluster analysis of seasonality histograms (Fig. 2.6).

2.4.2.1.3 Grouping into three regions and separate regressions in each region

Similarly to the latter 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 seasonality histograms (Fig. 2.7). 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.

2.4.2.1.4 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 generalisation of the multiple regression model for numeric and factor variables) fits a separate coefficient (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 sub-regions 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.

2.4.2.2 Regionalisation of summer period (q95s) and winter period (q95w) low flows

2.4.2.2.1 Global regression

As a side issue, specific low flows of the summer period (q95s) and the winter period (q95w) 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, I can expect further insights into the value of accounting for seasonality in the regionalisation.

2.4.2.2.2 Grouping into two regions and separate regressions in each region

The last approach is a combination of spatial grouping into summer and winter regions (Fig. 2.6) and the separate regionalisation 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 sub-models. This approach was used to get a more precise separation of summer and winter processes than by any of the two underlying methods alone.

2.4.3 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, I used leave-one-out cross-validation. The cross-validation estimate of prediction error is given by:

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

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

$$rmse = \sqrt{V_{cv}} \tag{9}$$

and the coefficient of determination based on cross-validation is:

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

where V_q is the spatial variance of the observed specific low flow discharges q_{95}. Note that most of the leveraging points (section 2.4.1) are included in the cross validation with the exception of one or two 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 regionalisation models. This is because cross-validation works well even if the regionalisation models are far from correct (Efron and Tibshirani, 1993). Cross-validation is hence a full emulation of the case of ungauged sites.

2.5 Results

2.5.1 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 is a desirable property if one is interested in interpretable estimates of model performance. In this study, model assumptions are carefully checked by scatterplots of observed vs. predicted values and histograms of residuals.

Scatter plots of observed vs. predicted specific low flow discharges q_{95} [$l.s^{-1}.km^{-2}$] in the cross-validation mode are presented in Fig. 2.10. 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-stabilising transformation in this case, such as taking the logarithms of q_{95}, but preliminary analyses showed that this transformation improved the heteroscedasticity of the transformed data but did not improve the heteroscedasticity of the residuals of the back-transformed predictions. 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. In catchments where observed specific low flow discharges are more than about 12 $l \cdot s^{-1} \cdot km^{-2}$ the low flows are consistently underestimated, 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 $l \cdot s^{-1} \cdot km^{-2}$ 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 regionalisation models to have a tendency for underestimating large values. For example, the flood regionalisation analysis of Merz and Blöschl (2004a) showed that flood quantiles in the same study area were consistently underestimated by their method for catchments with above-average specific flood discharges.

Histograms of cross-validated residuals ($l.s^{-1}.km^{-2}$) are presented in Fig. 2.11. For all regionalisation models, residuals appear similar and approximately normally distributed. Single extreme outliers (typically one or two per model) appear. 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 in order to get estimates for the bulk of the catchments.

2.5.2 Relative importance of predictor variables

The regression model equations of the four resulting models are presented in Table 2.2. Positive signs indicate higher discharges, negative signs indicate lower discharges with increasing values of each catchment characteristic. The catchment characteristics have been automatically selected by the stepwise regression algorithm. The order of the catchment characteristics 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 inter-correlations between catchment characteristics, different accuracy of catchment characteristics and influences of single values.

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 ground water tables (G_{GD}) have a positive effect, 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 results in two regression equations. The summer model consists of eight parameters and appears similar to the global regression model. Four catchment characteristics (S_{ST}, G_{GD}, L_R, G_F) are again represented and exhibit even the same coefficients, but the order of the selection is changed. Other parameters are only slightly modified (H_R is replaced by H_+, and the coefficient of P_W is changed which is due to different precipitation sums in the summer region and the entire region). G_C is missing in the selection, because Crystalline rocks do not occur in the summer-dominated region. The winter model consists of six parameters. Negative effects on low flows exhibit G_C and G_T (both indicate low permeability aquifer conditions) and H_M (correlates with temperature and appears to be an important control for winter low flows). G_Q (high permeability aquifer conditions), P_W and S_{ST} all exhibit positive effects on the low flows.

Grouping into three regions and separate regressions in each region leads to three significantly more parsimonious regression equations. The summer model consists of only two parameters, annual precipitation (P) and range of altitude (H_R), both indicating a positive effect on low flows. The winter model exhibits five parameters. Positive effects on low flows exhibit, again, P and S_{ST}. Negative effects on low flows exhibit L_A (landuse agriculture) and two landuse characteristics that indicate high-mountainous conditions (proportion of glaciers L_{GL} and proportion of rocks L_R). The model for the transition zone (mixed regimes) exhibits three characteristics H_M, G_T and L_F, which all have a negative effect on low flows.

Global regression but different Z parameters in each of the eight regions exhibits a regression equation that is very similar to the bench mark global regression model. H_R, L_R, G_Q are represented by the same regression coefficients, and P_W, G_F and G_C are represented by slightly modified coefficients. Percentage of steep slope S_{ST} is changed to mean slope S_M and proportion of tertiary sediments G_T replaces the proportion of deep groundwater tables (G_{GD}).

2.5.3 Relative performance of models

Table 2.2 presents two measures of model performance, the coefficient of determination R^2_{CV} and the root mean 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^2_{CV} = 57%, corresponding to rmse = 2.62 $l.s^{-1}.km^{-2}$. Grouping catchments into two regions and separate regressions in each region improves the overall model performance to R^2_{CV} = 59%, rmse=2.56 $l.s^{-1}.km^{-2}$. The summer low flow dominated region exhibits better performance (R^2_{CV} = 60%) than the winter-dominated region (R^2_{CV} = 51%). Grouping catchments into three regions and separate regressions in each region yields a further slight improvement of R^2_{CV} = 60%, rmse=2.55 $l.s^{-1}.km^{-2}$. Again, the sub-model for summer-dominated catchments exhibits a better performance (R^2_{CV} = 66%) than the sub-model for winter-dominated catchments (R^2_{CV} = 51%). The sub-model for mixed regimes indicates the poorest performance (R^2_{CV} = 35%). The global regression using different Z parameters in each of the eight regions exhibits a moderate performance of R^2_{CV} = 58%, corresponding to rmse = 2.61 $l.s^{-1}.km^{-2}$, 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 methods (Fig. 2.10). Regional regressions based on sub-regions tend to increase model performance, although the overall gain of performance is slim. One significant effect of seasonality based regional regression is that models for summer-dominated regions clearly perform better than models for winter-dominated regions.

A similar effect was observed for the summer period low flows q95s and winter period low flows q95w (Tab. 4). The global model for summer period low flows (R^2_{CV} = 65%) performs clearly better than the global model for winter period low flows (R^2_{CV} = 49%). However, this is not generally true if a grouping into two regions and separate regressions in each region is performed for summer period (q95s) and winter period (q95w) low flows. Here, the model for q95w in the summer region (R^2_{CV} = 43%) performs better than the model for q95w in the winter region (R^2_{CV} = 37%), but the model for q95s in the summer region (R^2_{CV} = 46%) exhibits a poorer performance than the model for q95s in the winter region (R^2_{CV} = 58%). The finding that all 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.

Table 2.2: Performance and coefficients of regional regression models for q95 low flows. Units of q95 are ($l.s^{-1}.km^{-2}$), units of catchment characteristics see Table 2.1.

Group	N	R^2_{CV}	$rmse_{cv}$	Model
Global	325	57	2.62	$\hat{q}_{95}$ = -2.04 + 0.23*H_R - 0.08*L_R - 0.04*G_F + 1.29*P_W + 0.04*G_Q + 0.04*S_{ST} + 0.03*G_{GD} - 0.01*G_C
2 regions total	325	59	2.56	
summer	215	60	2.74	$\hat{q}_{95}$ = -3.08 + 0.04*S_{ST} - 0.04*G_F + 0.03*G_{GD} + 1.53*P_W + 0.03*G_Q - 0.08*L_R + 0.14*H_+
winter	110	51	2.16	$\hat{q}_{95}$ = 9.66 - 0.02*G_C + 0.14*G_Q + 0.43*P_W + 0.05*S_{ST} - 0.36*H_M - 0.41*G_T
3 regions total	325	60	2.55	
summer	210	66	2.43	$\hat{q}_{95}$ = -6.56 + 1.16*P + 0.12*H_R
winter	82	51	2.60	$\hat{q}_{95}$ = -0.72 + 0.62*P - 0.03*L_A - 0.07*L_{GL} + 0.05*S_{ST} - 0.04*L_R
mixed	33	35	2.93	$\hat{q}_{95}$ = 38.28 - 1.29*H_M - 0.22*G_T - 0.21*L_F
8 regions	325	58	2.61	$\hat{q}_{95}$ = -2.17 + Z + 1.18*P_W + 0.23*H_R - 0.08*L_R + 0.04*G_Q + 0.07*S_M - 0.03*G_F - 0.02*G_C - 0.02*G_T

Table 2.3: Z-parameters ($l.s^{-1}.km^{-2}$) for each of the 8 regions. p is the significance level.

Group	**Region**	**Z**	**p-value**
A-C	Alps	+0.437	0.52
1	Flatland & hilly terrain (N,E of Austria)	+0.378	0.95
2	Bohemian Massif	-0.317	0.26
3	Foothills of Alps (Upper Austria)	+2.260	<0.01
4	Flyschzone	+0.222	0.39
5	Lower Carinthia	-1.327	0.07
D	Pre-Alps (Styria)	-0.417	0.33
E	Pre-Alps (Vorarlberg)	-1.235	0.22

Table 2.4: Performance of regional regression models for summer period low flows (q95s) and winter period low flows (q95w). Units of q95 are ($l.s^{-1}.km^{-2}$), units of catchment characteristics see Table 2.2.

Group	N	R^2_{CV}	$rmse_{cv}$	Model
Summer period global	325	65	3.05	$\hat{q}95s$ = -3.94 + 0.18*S_M + 0.48*P - 0.16*L_{GL} + 0.03*G_{GD} + 0.04*G_Q - 0.04*G_F + 0.25*H_+ - 0.08*L_R - 0.02*G_C - 0.02*L_F
Winter period global	325	49	2.87	$\hat{q}95w$ = -0.54 + 0.91*P_W + 0.07*G_Q - 0.07*L_R - 0.04*G_F + 0.04*G_D + 0.14*S_M - 0.02*G_C
Summer period | summer region	215	46	2.71	$\hat{q}95s$ = -3.56 + 0.63*P + 0.07*S_{ST} + 0.03*G_{GD} + 0.14*H_+
Summer period | winter region	110	58	3.20	$\hat{q}95s$ = 5.90 - 0.02*G_C + 0.12*S_{ST} - 0.34*H_M + 0.74*P_S + 0.18*G_Q
Winter period | summer region	215	43	3.08	$\hat{q}95w$ = - 2.21 + 2.01*P_W
Winter period | winter region	110	37	2.15	$\hat{q}95w$ = 7.37 - 0.02*G_C + 0.17*G_Q - 0.21*H_0 - 0.34*G_T + 0.34*P_W

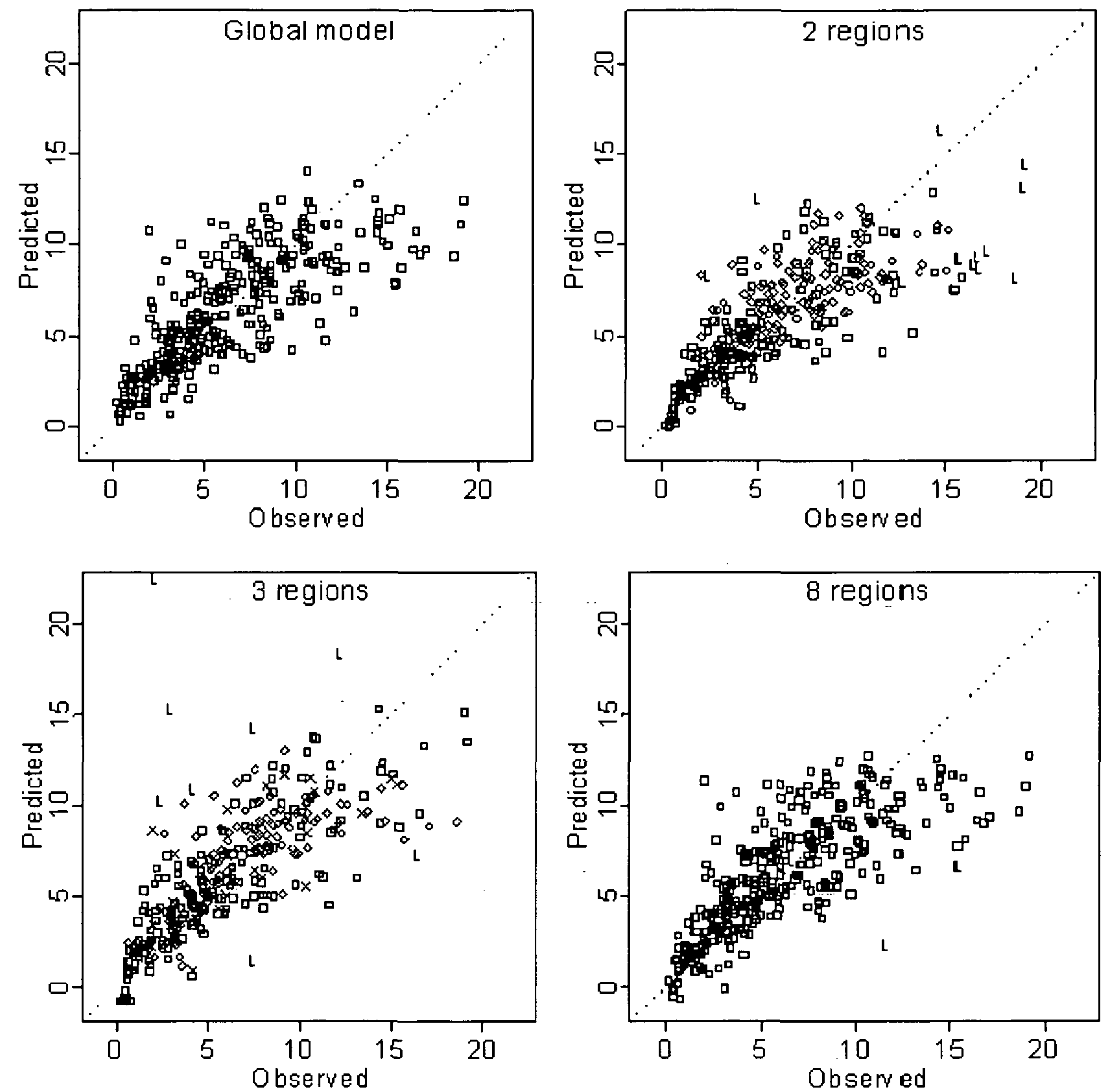

Fig. 2.10: Scatter plots of predicted vs. observed specific low flow discharges q95 (l·s-1·km-2) in the cross-validation mode. Each panel corresponds to a regional regression model and each point corresponds to a catchment. Point markers L indicate leverage points. Legend for panels of 2 regions and 3 regions: squares - summer low flow dominated catchments, diamonds - winter low flow dominated catchments, crosses - catchments exhibiting mixed regime.

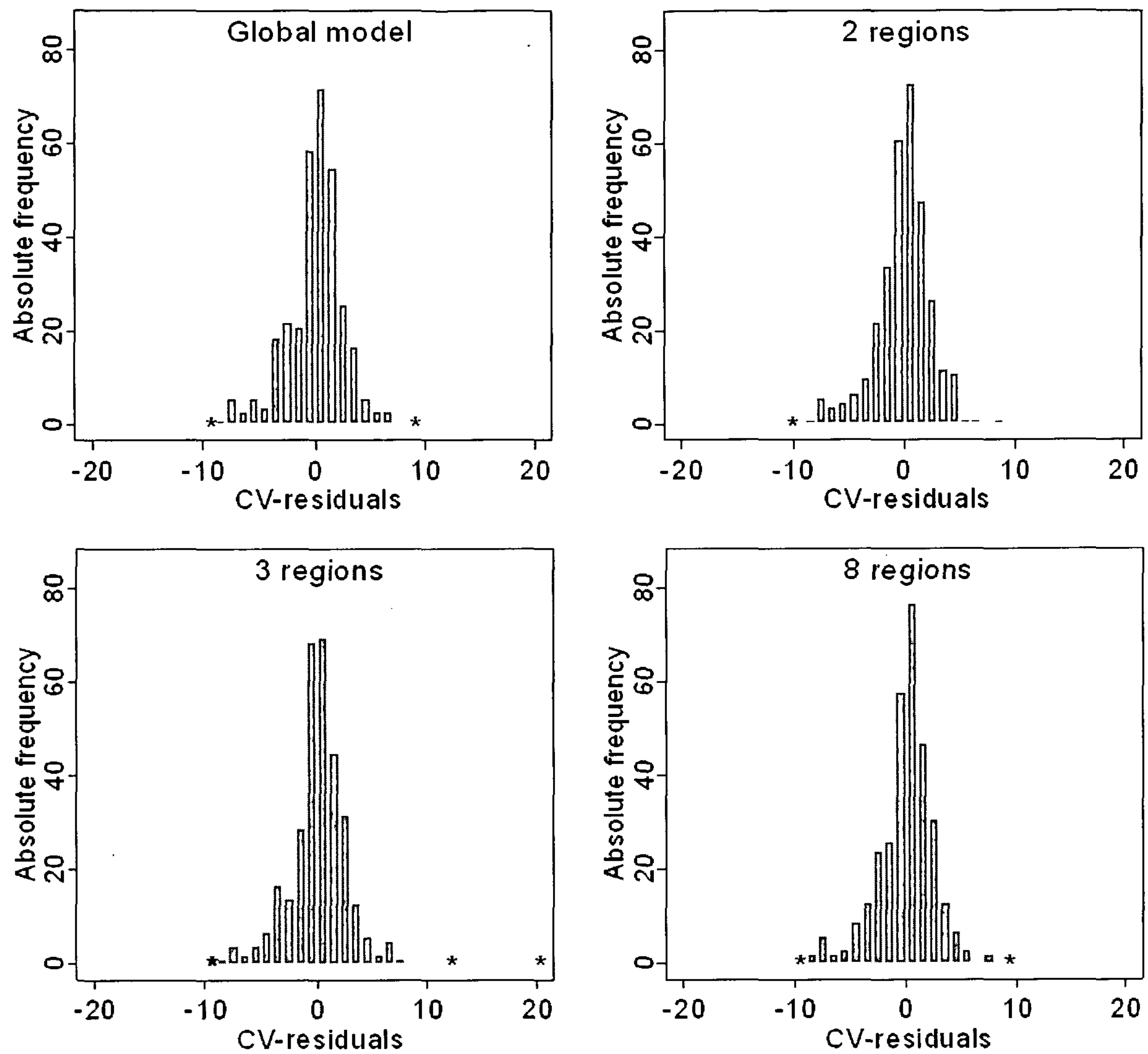

Fig. 2.11: Histograms of cross-validated residuals ($l.s^{-1}.km^{-2}$) of regionalisation. Each panel corresponds to one regionalisation model. Asterisks indicate single extreme values.

2.6 Discussion

2.6.1 Performance of regionalisation 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) and 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^2_{CV} = 57%, and explains 60% of the variance in q_{95} (goodness-of-fit R^2). It is interesting to compare this result to studies in the literature that used a similar number of catchments as in this section (325 catchments) and examined q95 specific discharges as in this section, rather than discharges. I 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 is 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} standardised 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 south-west 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 sub-catchments rather than complete catchments which may make the catchment characteristics more relevant to low flow regionalisation.

Most other studies in the literature used discharge rather than specific discharge and so are not directly comparable to the results in this section. As catchment size usually explains around 80%-90% of the variability of low flow discharges (see, e.g., Dingman and Lawlor, 1995; 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 (1992), for example, reported R^2 values of more than 90%.

2.6.2 Relative performance of regionalisation methods

In a first step I compared the overall performance of the models. The coefficients of determination in the cross-validation mode and scatter-plots of observed vs. predicted low flows q95 indicate only slight differences among the models (section 2.5.1 and 2.5.3). Grouping into three regions and separate regressions in each region performed best, followed by grouping into two regions and separate regressions in each region. Global regression performed worse, and 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 I compared the performance of sub-models for catchments that represent summer- and winter low flow dominated regimes. As can be seen from the cross-validated coefficients of determination of the sub-models, models for summer-dominated catchments exhibit a better performance than the models for the winter-dominated catchments (Table 2.2), and the models for the summer period low flows perform better than the models for the winter period low flows (Table 2.4). Summer-dominated catchments correspond to lowlands where the hydrologic situation is relatively simple and winter-dominated catchments correspond to alpine catchments where the hydrologic 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. The relatively poor performance of the model for the winter period low flows hence indicates that winter processes are not represented as well as the summer processes. This might point to both a higher complexity of winter processes as well as a poorer representation of winter processes by the available catchment characteristics.

Overall, models for summer-dominated catchments exhibit a better performance than models for winter-dominated 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 regionalisation. As this effect was not apparent in the performance measures discussed above, it appears that a more detailed assessment of residuals is needed.

In a third step, I therefore compared maps of residuals between predicted and observed values of specific low flow discharge q95 ($l.s^{-1}.km^{-2}$) for each sub-catchment (Fig. 2.12). In this figure, each panel corresponds to one model. Positive residuals indicate overestimation by the model, 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 bench mark 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 section 2.5.1. Large residuals occur in the southern part of Upper Austria which indicates a specific hydrological situation that would require a separate regionalisation model.

Compared to the global model, both regional regression models that fit separate models for 2 and 3 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 south-east 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 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. Regionalisation is clearly improved for a large area from the east (Wechsel region) to the south-west 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 regionalisation model between well represented and poorly represented situations, indicating the best performance for predicting low flows among all considered models.

The global regression using different Z parameters in each of the eight regions exhibits a pattern of cross-validated 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 2.3) and seasonality has no significant influence on the quantity of low flows. An alternative consideration of these eight regions in a complex regional regression model which fits separate models for each region similar to the more successful models for two and three zones may therefore be a promising approach and will be examined in section 3.

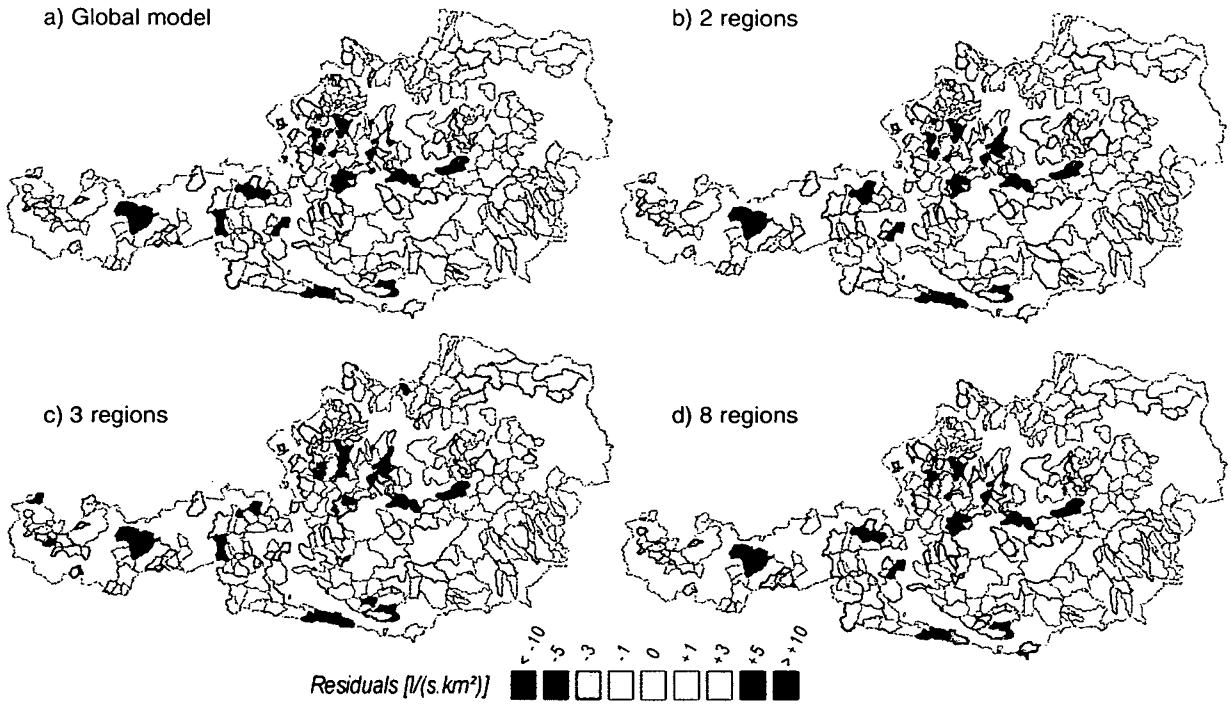

Fig. 2.12: Residuals of predicted minus observed specific low flows q95 ($l.s^{-1}.km^{-2}$) in cross validation mode. Each panel corresponds to one regionalisation model.

Overall, the assessment of models by maps of cross-validated residuals resulted in a somewhat different assessment from that of the coefficient of determination of cross-validated residuals. I 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 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 (section 2.3.2.2).

2.6.3 Relative performance of approaches to seasonality analysis

I used different approaches to seasonality analysis in this section. I first compared three seasonality measures, the seasonality ratio (SR), the seasonality index (SI) and the seasonality histogram (SH), which use a different number of parameters for representing the seasonal distribution of low flows. SR consists of one parameter, which is the ratio of summer period low flow characteristic q95s and winter period low flow characteristic q95w. SI consists of two parameters, the mean and the variability of days when discharges fall below Q95. SH consists of 12 parameters which represent the monthly frequency of days when discharges fall below Q95. SH therefore provides the most detailed information about low flow seasonality, SI contains less information, and 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. While general patterns are best visible for SR, the determination of patterns by the SI requires closer inspection. There is too much information in the SH for a synoptic interpretation at the regional scale, so some classification technique is needed to exploit the higher information content.

I hence 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 which also provides information about the optimal number of clusters, led to a classification of Austrian catchments into two contiguous regions which 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 sub-models for the regions and the residual pattern maps showed that the classification increases the performance of regionalisation and there is some value in the seasonality grouping. The second technique is the visual grouping of SH, which uses the large scale pattern of the SI to obtain a preliminary classification which 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 correspond generally 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 I believe there is some potential in this classification provided more sophisticated regionalisation models are used.

2.6.4 Most important controls in comparison with controls in the literature

2.6.4.1 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 the Alpine region where low flows are dominated by winter processes and in 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 (Flyschzone along the foothills of the Alps, Eastern Styria and Lower Carinthia). Smaller differences of mean seasonality between sub-regions 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 q95. This suggests that seasonality is an indicator of processes rather than an indicator of the magnitude of specific low flows.

2.6.4.2 Controls of specific low flow discharge

Further insights in important controls of low flows were gained from a comparison of the different regional regression models (Table 2.2). Although a large number of catchment characteristics were supplied to the stepwise variable selection algorithm, only 14 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. I therefore believe that the interpretation of the regression coefficients provides useful insights into the important process controls of low flows in Austria.

Catchment relief is represented in all equations, generally by one altitude parameter (H_M, H+ or H_R) and by one slope parameter (mainly S_{ST}, also S_M). 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. Also, the fact that one of the topographic characteristics is frequently selected first by the stepwise regression algorithm supports the finding that catchment topography (altitude and slope) is one of the most important 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 and is also frequently selected first by the stepwise regression algorithm. Hence, precipitation is an equally important control of low flows in Austria as catchment topography. 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.

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. Seven 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 and appears not to be an important control. There are only two land use characteristics that appear frequently (L_R, L_{GL}), but they are associated with high mountain conditions rather than independent indicators of landuse.

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

In the final step, I 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 in Demuth (2004). 73% of all catchment characteristics used are physiographic descriptors, 22% are climatic descriptors and 5% are hydrologic descriptors. Among the physiographic descriptors, morphometric descriptors make up the highest proportion (46%), followed by surface cover (17%), geology and soil (10% each). However, I believe that the frequency of the catchment characteristics used largely depends 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 regionalisation study of Switzerland presented by Aschwanden and Kan (1999) is similar to this study in both respects. The regionalisation of q95 specific low flows in Switzerland resulted in seven regional regression models for different regions in Switzerland. From all regression equations, catchment topography and precipitation 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). Landuse played an important role but the characteristics were proportion of horticulture and vineyards, and 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 are therefore in line with the regionalisation study of Switzerland. Relief, precipitation 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.

2.7 Conclusion

The objective of this study was to examine the value of different seasonality indices for low flow regionalisation. In a first step, three seasonality indices were compared. The main difference between the indices is the information content of low flow seasonality. Seasonality histograms (SH) are the most detailed indices, but classification techniques are needed to compare seasonality among a large number of catchments. The cyclic seasonality index (SI) is a more compact index and the spatial patterns can be delineated by visual inspection of a vector map of SI. The seasonality ratio (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, low lands and hilly landscapes. In a second step, three catchment classification methods that are based on seasonality have been examined. Cluster analyses of seasonality histograms resulted in a first 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 regionalisation was examined by comparing three multiple regression approaches which include the seasonality classifications in different ways, to 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 three regions (summer, winter and mixed seasonality) performs best (R^2_{CV} = 60%), followed by separate models fitted to two regions (R^2_{CV} = 59%). Including different calibration coefficients in each of the eight seasonality regions resulted in R^2_{CV} = 58% and hence performs only slightly better than the global regression model (R^2_{CV} = 57%). The models for the summer regions (R^2_{CV} = 66% and 60%), however, clearly perform better than the models for the winter regions (R^2_{CV} = 51%). The model for the catchments of the mixed seasonality type (R^2_{CV} = 35%) does not nearly perform as well. The residual maps of predicted minus observes q95 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. This will be examined in section 3.

3 A comparison of low flow regionalisation methods – catchment grouping

3.1 Introduction

Accurate estimates of low flow characteristics are needed for a range of purposes in water resources management and engineering including environmental flow requirements, water uses and discharges into streams, and hydropower operation (Smakhtin, 2001; Gustard et al., 2004). Low flow characteristics are best estimated from observed stream flow data but for sites where these data are unavailable hydrological regionalisation techniques can be used to infer them from other catchments where stream flow data have been collected.

The regionalisation of low flow characteristics is usually based on some sort of regression between the low flow characteristic of interest and catchment characteristics that are available for ungauged sites (e.g., Vogel and Kroll, 1992; Gustard et al., 1992; Schreiber and Demuth, 1997; Skop and Loaiciga, 1998). If the study domain is large or very heterogeneous in terms of the low flow processes a number of authors have suggested to split the domain into regions and apply a regression relationship to each of the regions independently. This is termed the regional regression approach. Gustard and Irving (1994), for example, tested a number of regression models between standardised Q_{95} low flows (the discharge that is exceeded on 95% of all days) and different soil group indices for 1530 catchments in Europe. Their global regression model of nine soil classes explained 29% of the spatial low flow variance while a regional regression model explained between 17% and 47% of the variance, depending on the region. In their study, the entire domain was subdivided into seven geographic regions. In a smaller scale study of 44 catchments in New Zealand, Clausen and Pearson (1995) showed that the variance of a drought index explained by catchment characteristics increased from 64% to between 68% and 91% if the domain is split into three geographically defined regions.

In some instances it is clear how to group a domain into regions of approximately uniform low flow behaviour but, more often, the choice is far from obvious. A number of methods of identifying homogeneous regions have therefore been proposed in the literature in the context of low flow regionalisation. All of these methods use low flow data and most of them use catchment characteristics as well. In the first technique, termed residual pattern approach, residuals from an initial, global regression model between flow characteristics and catchment characteristics are plotted, from which geographically contiguous regions are obtained by manual generalisation on a map (e.g. Hayes, 1992; Aschwanden and Kan, 1999).

This is an obvious method of improving on a global regression model. A drawback of the residual pattern approach, however, is that the initial model may be far from correct as it extends over the entire domain of interest. The shapes of the regions so obtained may then be artefacts of an inadequate model and the regional regression model will have little physical significance. Once the regions have been identified, the ungauged site of interest needs to be allocated to one of the regions. As the regions in this approach are spatially contiguous, the ungauged site can be allocated by its geographical location. As a final step, the low flow value for the site of interest is estimated from multiple regressions between observed low flows and catchment characteristics fitted to each of these regions independently.

In the second technique, multivariate statistics such as cluster analyses are used to delineate regions. In the multivariate analyses, both low flow data and catchment characteristics are used. They are usually standardised and/or weighted to enhance the discriminatory power of the methods. The use of multivariate statistics in the context of low flow regionalisation has been explored in detail by Nathan and McMahon (1990). They tested a number of approaches based on a combination of different techniques of cluster analysis, multiple regression and principal component analysis. They used Andrews curves (Andrews, 1972) for visualising similarity in catchment characteristics which allowed them to fine-tune the catchment grouping. Based on data from 184 catchments in South-east Australia, Nathan and McMahon (1990) found that the relative performance of the methods depended on the low flow characteristic examined. Their overall assessment suggested that the weighted cluster analysis (Ward's method based on a Euclidean distance measure) using weights according to the coefficients of an initial stepwise regression model performed best. Since regions obtained by the cluster analysis approach are generally discontiguous in space, the allocation of ungauged sites to the most similar group requires decision criteria which are usually based on catchment characteristics. Nathan and McMahon (1990) assumed in their analysis that the catchment allocation was known and proposed to use Andrews curves for assigning ungauged catchments. Possible alternative methods are discriminant analyses and classification trees (Haines et al., 1988). As a final step, again, the low flow value for the site of interest is estimated from multiple regressions between observed low flows and catchment characteristics fitted to each of the regions independently.

A third technique are Classification And Regression Tree (CART) models (Breiman et al., 1984) which, to my knowledge, have not yet been used in low flow regionalisation. However, there do exist a number of interesting applications in hydrology, including the classification of satellite images of snow cover and the interpolation of ground snow measurement (e.g. Rosenthal & Dozier, 1996; Elder, 1995) and a first attempt of using the regression trees for regionalizing low flows is given in Laaha (2002). In the context of low flow regionalisation, the independent variables in the regressions trees are the catchment characteristics and the dependent variables are the low flows. Regression trees then divide a heterogeneous domain into a number of more homogeneous regions by maximising the homogeneity of low flows and catchment characteristics within each group simultaneously. Regression trees have a number of advantages over other models. Their structure is non-parametric, small trees are readily interpretable, there is no global sensitivity to outliers and they are able to handle non-linear relationships well. However, big trees are difficult to interpret, there is a lack of smoothness and there are potential problems with overfitting the data, so some method for pruning the tree is needed (Breiman et al., 1984). Once the regression tree is fitted to the data, it can be used to allocate ungauged sites to the regions obtained by the regression tree. Alternatively, classification trees can be used to allocate group names to catchment characteristics. Classification trees operate on categorical variables while regression trees operate on continuous variables. The final step of estimating low flows for the ungauged site of interest is a regional regression as in the other grouping methods.

In a fourth technique, the seasonality of low flows is used to delineate homogeneous regions. The rationale of this approach is that differences in the occurrence of low flows within a year are a reflection of differences in the hydrologic processes and are hence likely to be useful for finding homogeneous regions. 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 regionalisation 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 visualise the spatial patterns of the seasonalities of floods and other hydrologic variables. In contrast, an application of a low flow seasonality index in the UK (Young et al., 2000) suggested there is little discriminatory power in this index because the spatial variability of low flow seasonality was rather weak. 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 in Upper Austria and found that both measures were capable of classifying catchments into summer and winter low flow dominated sub-regions. An extension of this work is presented in section 2, where homogenous regions with respect to low flow seasonality were visually delineated from a number of seasonality measures. The results indicated that, in a humid, mountainous country such as Austria, the spatial variations in low flow seasonality are indeed enormous. There is likely some potential in this approach.

If the regions are spatially contiguous such as those presented in section 2, the ungauged site can be allocated by its geographical location. The final step of estimating low flows for the ungauged site of interest is an analogous regional regression to the other grouping methods.

While much work has been done in the literature on catchment grouping in the context of low flow regionalisation I am unaware of any comprehensive comparison of the grouping methods for the same data set to assess their relative merits. The aim of this section therefore is to examine the relative performance of different grouping techniques to investigate what is the optimum grouping method for regionalising low flows. The comparison will be made on the same data set as it was used in section 2, i.e., 325 sub-catchments and catchments without an upstream gauge, respectively, and the low flow characteristic chosen is, again, the q_{95} specific low flow quantile, i.e. the specific discharge that is exceeded on 95% of all days.

3.2 Method

3.2.1 Classification of catchments

3.2.1.1 Residual pattern approach

The residual pattern approach to catchment grouping consisted of three steps:

(1) Perform stepwise regression to obtain a global regression model;

(2) Plot the residuals from the global regression model in geographic space;

(3) If residual patterns are apparent, delineate contiguous regions of similar sign and magnitude of residuals.

Stepwise regression may lead to over-fitted models where omission of a single catchment characteristic only slightly reduces the global model quality. When choosing the number of catchment characteristics in the global regression I therefore tended to use the more parsimonious model as it produced clearer residual patterns.

3.2.1.2 Weighted cluster analysis

Weighted cluster analysis has been recommended by Nathan and McMahon (1990) as the optimal technique to identify homogeneous regions and I used their method consisting of the following steps:

(1) Identify the catchment characteristics most relevant to the problem at hand by performing an overall stepwise regression analysis;

(2) Weight the selected catchment characteristics according to their relative importance, as determined by the magnitude of their ß-coefficients which are the coefficients of the stepwise regression model based on standardised catchment characteristics;

(3) Perform a number of cluster analyses on the weighted catchment characteristics using different measures of similarity and linkage methods;

(4) Prepare plots of Andrews curves for each of the groupings derived in (3), and identify the set of clusters exhibiting the least within-group variation. This will give the optimal classification of catchments into homogeneous groups;

(5) Remove outliers in the optimum grouping based on the Andrews plots, if needed. Derive a set of mean catchment characteristics for each homogeneous group;

(6) Refine the optimum grouping derived by the cluster analysis by comparing the catchment characteristics of each catchment with the group mean and reclassify the catchment in case the catchment characteristics are too different.

In the spirit of Nathan and McMahon (1990), I compared several cluster analysis techniques of the S-Plus statistics package. These were two hierarchical cluster analysis methods, *hclust* (Hartigan, 1975) and *agnes* (Kaufman and Rousseeuw, 1990), which are similar to the algorithm used by Nathan and McMahon, as well as the *pam* partitioning method (Kaufman and Rousseeuw, 1990). Several combinations of linkage methods (single linkage, average linkage and complete linkage) and distance measures (Euclidean distance and Manhattan distance) were evaluated for different numbers of clusters. The most appropriate method was selected by a visual assessment of Andrews plots. In Andrews plots, a point in multi-dimensional space $x = [x_1, x_2, ..., x_n]$ is represented by a function of the form:

$$F(t) = x_1 / \sqrt{2} + x_2 \sin(t) + x_3 \cos(t) + x_4 \sin(2t) + x_5 \cos(2t) + \cdots \qquad (1)$$

plotted over the range of $-\pi \leq t \leq +\pi$. A set of multivariate observations can be displayed as a collection of these plots and those functions that remain close together for all values of t correspond to observations that are close to one another in terms of their Euclidean distance. This property implies that these plots can be used to both detect groups of similar observations and identify outliers in multivariate data.

Since the regions obtained by weighted cluster analysis, generally, are not contiguous, the prediction of low flow characteristics at ungauged sites requires a decision rule based on catchment characteristics in order to allocate the site of interest to the most appropriate region. Nathan and McMahon proposed a procedure similar to step (6), i.e. comparing the Andrews curve of an ungauged catchment with the mean curve of each cluster. Because of the subjectivity of a visual assessment, this method is not suitable for automatic cross-validation of the regional regression model. I therefore adopted an alternative approach and used classification trees for automatically allocating ungauged catchments to the most appropriate cluster. Similarly to the regression trees (see below), the classification tree was fitted based on 10-fold cross-validation to determine the optimum tree size for prediction.

3.2.1.3 Regression tree

In this section, I propose regression trees for obtaining homogeneous regions to be used in a regional regression approach. Regression trees are an exploratory technique for finding homogeneous regions among predictor variables (i.e. catchment characteristics) with respect to a target variable (i.e. q_{95} low flow). The regression tree is constructed by an algorithm known as *binary recursive partitioning* (Clark and Pregibon, 1991). By this algorithm, groups of catchments are subsequently subdivided by binary conditions (e.g. IF P_s<534mm THEN sub-group x ELSE sub-group y), starting from the most important catchment characteristics and proceeding to the less important ones. Each condition yields the optimal subdivision of a group which minimises the sum of squared differences between observed values of q_{95} and the group mean, a measure that is termed the *deviance* of the node. The algorithm identifies the most important catchment characteristics, and potential interactions between catchment characteristics are handled implicitly (Venables and Ripley, 1999).

Tree construction can be carried out until each terminal node consists of one single catchment but this leads to a model with little significance for prediction or classification problems. To avoid such over-fitting, trees need to be pruned back, and the optimal number of nodes is best determined by an independent validation data set. If no such validation data set is available, I can split the data set into 10 (roughly) equally sized parts, subsequently use nine parts for calibration and one part for validation, and calculate the average prediction error (total deviance of a tree) for several tree sizes. This procedure, termed 10-fold cross-validation, is part of the S-Plus *tree*-package and was used in this study.

Regression trees have the convenient property of invariance against monotone transformation of predictor variables (i.e. catchment characteristics). However, the dependent variable (i.e. q_{95}) needs to be normally distributed for optimal tree construction. I examined the distribution of q_{95} in the data set of this section and found that a square-root transformation of q_{95} yields a distribution that is close to normal. Since the regression tree is used for classification but not for prediction, no retransformation is needed which may be non-unique if the transformed variable changes sign.

The regression tree approach to catchment grouping consisted of the following steps:

(1) Perform transformation to normality;

(2) Fit an initial regression tree to the data;

(3) Determine the optimal tree size by 10-fold cross-validation;

(4) Prune the initial tree back to the tree size derived in (3).

While regression trees are suitable for allocating unobserved catchments to the most appropriate clusters, they are not suitable for cross-validation of the resulting regional regression model as the names of the clusters may change when models are refitted for subsets of the data. I therefore fitted a classification tree to the group names of the regression tree as (categorical) dependent variable which exhibited an identical structure to the regression tree, but had the advantage of producing the same group names for various data subsets. This allowed me to assign each ungauged catchment to one of the clusters of the regression tree in the cross-validation.

3.2.1.4 Regions of similar low flow seasonality

Regions of similar low flow seasonality as defined by Laaha and Blöschl (2003) were used as the final scheme for catchment grouping. Laaha and Blöschl (2003) classified Austria into eight contiguous regions based on a visual assessment of two seasonality measures. The first seasonality measure was a seasonality index based on circular statistics (Young et al., 2000) that represented the mean and the variance of the days of low flow occurrence. The second seasonality measure were seasonality histograms (Laaha, 2002) which were used to refine the information from the seasonal statistics. Catchment elevation was used to assist in the delineation of the regions as, in Austria, topographic elevation is one of the main controls of hydrologic regimes. The method consisted of the following steps:

(1) Determine the Julian dates (i.e. days from 1 to 365) of days of low flow occurrence for each sub-catchment by selecting all days when daily discharge was below Q_{95};

(2) Calculate the seasonality index from the dates for each sub-catchment and plot the seasonality indices as a vector-map in geographical space;

(3) Delineate preliminary regions on the vector map;

(4) Plot monthly histograms of low flow occurrence for each sub-catchment and use them to refine the preliminary classification;

(5) Use topographic elevation to refine the exact position of the region boundaries.

3.2.2 Regional regression approach

For each group identified by the classification methods, a multiple regression model was fitted independently with specific low flow discharge q_{95} as the dependent variable and a set of catchment characteristics as the independent variables. Catchment characteristics are often subject to inter-correlations and multicollinearity. Rather than performing a selection of the most important variables prior to regionalisation, I used a stepwise regression approach. The stepwise regression procedure used *Mallow's* C_p (Weisberg, 1985, p. 216) as the criterion of optimality, which was calculated as:

$$C_p = \frac{RSS_p}{\hat{\sigma}^2} + 2p - n \qquad (2)$$

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 penalised by the second term which adds the number of coefficients p minus the number of catchments n. C_p is therefore a penalised selection criterion which 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 minimise the C_p criterion. After each step it is tested if replacing one of the variables by any remaining catchment characteristic will further decrease the criterion. 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 in hydrology is often complicated by single extreme values or outliers. Eliminating outliers may improve the goodness-of-fit but this does not necessarily entail an increase in the predictive power of the model. On the other hand, extreme values may act as leverage points and force the fitted model close to them, particularly if the regression model is fitted by the least squares method, which increases the magnitude of the residuals of the remaining points. I therefore adopted an iterative robustified regression technique in this section. Initial models were fitted by stepwise regression and then checked for leverage points using Cook's distance (e.g. Weisberg, 1985). These leverage points were removed from the sample and the regression model was refitted iteratively until no leverage points remained. The final model quality was assessed for all data including leverage points. q_{95} was used in all regional regressions without transformation, as exploratory analyses of the data suggested that transformations did not increase the predictive performance.

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 ten to give similar magnitudes in terms of their ranges (see Table 2.1). Since linear regression is scale invariant (Weisberg, 1985, p.185) the regression models, including their residual statistics, remain unaffected by the rescaling but the numerical stability is improved.

3.2.3 Analysis of predictive performance

3.2.3.1 Analysis of variance

In a first step, I was interested in how well the classification into homogeneous regions may explain the spatial variability of specific low flow discharges, q_{95}. A widely used measure of the explanatory power of groupings is the one-factorial analysis of variance (ANOVA) which was used here with q_{95} as the dependent variable and the classification number as the independent variable. The ANOVA may be interpreted as an assessment of a simple regionalisation model, where predicted q_{95} is simply the average low flow discharge in each group of a classification. The coefficient of determination (R^2) of this model, i.e. the ratio of the variance explained by the classification and the total variance of low flows, is a measure of the goodness-of-fit of this simple model. R^2 values close to 100% indicate an excellent goodness-of-fit while smaller values indicate a poorer goodness-of-fit.

3.2.3.2 Goodness-of-fit of component regressions

In a second step I examined how well the regression models in each of the regions fitted the data. I assessed the goodness-of-fit by the coefficient of determination of the regressions separately in each of the regions.

3.2.3.3 Cross-validation of regional regression

The value of the classification methods for the ultimate purpose of estimating low flow characteristics at ungauged sites can not be fully assessed by goodness-of-fit statistics. A more appropriate measure of the prediction errors are the error statistics from leave-one-out cross-validation. In this section, the cross-validation procedure consisted of the following steps:

(1) Remove catchment i from the data set;

(2) Update the catchment classification (if appropriate) for the remaining n-1 catchments;

(3) Assign catchment i to one of the regions obtained in (2);

(4) Estimate the coefficients of the regression equation for this region using all catchments in this region apart from catchment i;

(5) Apply the regression equation obtained in (4) to predict the low flow characteristic q_{95} at site i;

(6) Repeat steps (1) to (5) for all n catchments;

(7) Calculate the predictive error for each catchment i as q_{95} estimated in (5) minus observed q_{95} and analyse the error statistics.

In some of the classification methods the catchment classification was updated during the cross-validation procedure while in others it was not. In the weighted cluster analysis and the regression tree approaches the regions are discontiguous, and will hence significantly change if a single catchment is added. In these methods the classification was updated. In the residual pattern and the seasonality region approaches, however, the regions are contiguous and will therefore not change much if a single catchment is added. In these methods the classification was not updated.

From this prediction vector, the cross-validation prediction error V_{cv} was then estimated by:

$$V_{cv} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{q}_{95_i}^{(-i)} - q_{95_i}\right)^2 \qquad (3)$$

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

$$rmse = \sqrt{V_{cv}} \qquad (4)$$

and the coefficient of determination based on cross-validation is:

$$R_{cv}^2 = \frac{V_q - V_{cv}}{V_q} \qquad (5)$$

where V_q is the spatial variance of the observed specific low flow discharges q_{95}.

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

3.3 Results

3.3.1 Residual pattern approach

A preliminary global regression model was fitted to the data by stepwise regression. Since the primary purpose of the global model was to calculate a meaningful residual pattern, the residuals were carefully checked for the general assumptions underlying multiple regression, unbiasedness ($E[res_i]=0$) and homoscedasticity ($Var[res_i]=constant$), where res_i is the residual of catchment i. The analysis indicated slight heteroscedasticity which appeared to be a consequence of a significant skew of the distribution of q_{95}. I therefore transformed q_{95} by a square-root transformation which resulted in approximate normality. The global regression model was then fitted to the transformed data. The retransformation is non-unique if the variable changes sign but since all predictions were positive this was not a problem.

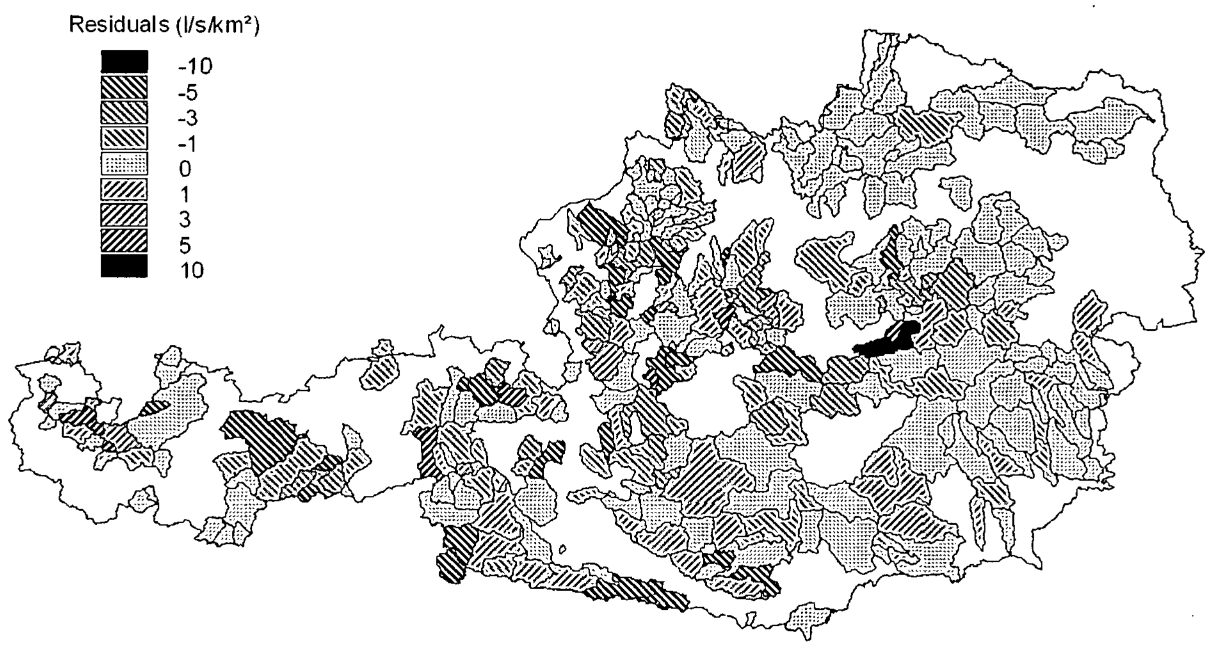

Fig. 3.1. Residual pattern of the global regression model (goodness-of-fit residuals). Positive residuals indicate an overestimation by the model.

Stepwise regression resulted in seven catchment characteristics used as predictors. This equation was manually revised and three of the predictors were removed to avoid overfitting. There was only a slightly loss in the goodness-of-fit when removing these predictors (R^2 decreased from 66% to 62%). The remaining predictors were H_R (range of altitude), L_R (fraction of wasteland or rocks), G_F (fraction of Flysch) and P_W (average winter precipitation). The more parsimonious model indicated a clearer residual pattern than the full model based on seven predictors and hence seemed to be more suitable for detecting homogeneous regions. The residual map is presented in Fig. 3.1. The residual pattern suggests that Austria can be classified into two main units. The first unit consists of the Bohemian massif in the North, lowlands and the foothills of the Alps in the East and South, and some of the Alpine catchments of West- and East-Tyrol. In this unit, the magnitude of the residuals is generally low ($< 1\ l \cdot s^{-1} \cdot km^{-2}$ for most catchments, except for East-Tyrol) and the pattern of the residuals is random, so the global model seems to work well in this unit. The second unit consists of the Alpine catchments and the Molassezone in the North. In this unit, the magnitude of the residuals is larger although there are no clear patterns. I chose to subdivide the second unit into four regions based on the geology. This gave me a total of five regions as shown in Fig. 3.2. These are (0) the Bohemian massif, lowlands and the foothills of the Alps, (1) Central-Alps and Pre-Alps, (2) parts of the Northern Calcerous Alps, (3) parts of Carinthia and (4) the

Bregenzerwald (Vorarlberg). Region 0 relates to small residuals, region 1 relates to negative residuals, and the remaining regions 2 - 4 relate to positive residuals.

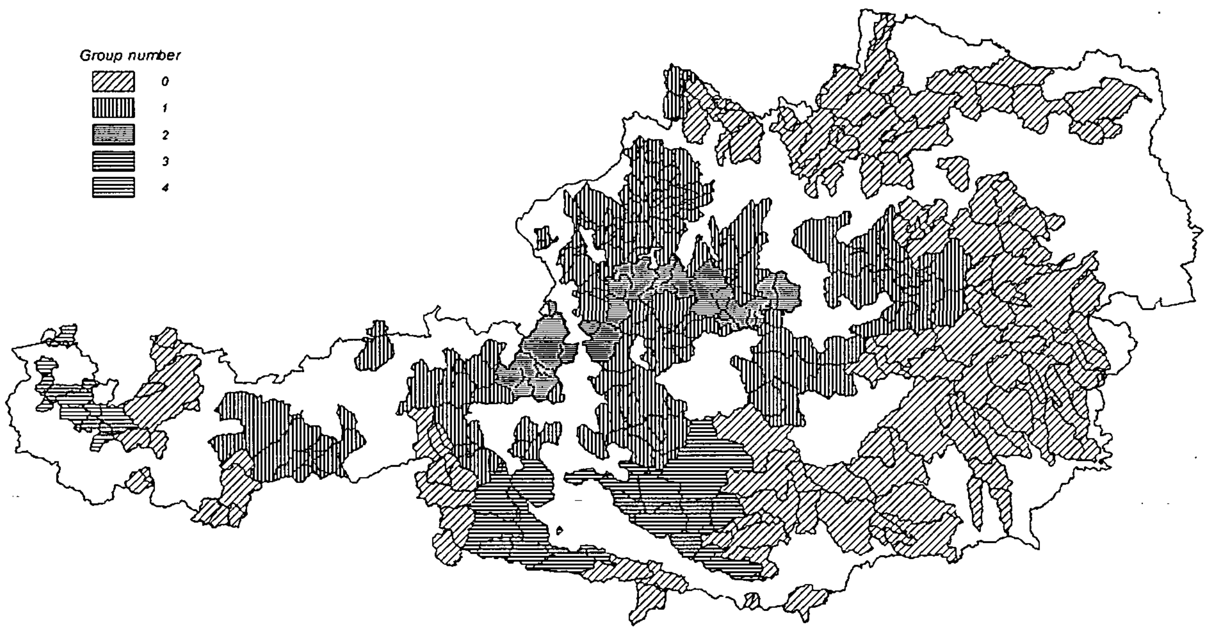

Fig. 3.2. Classification of catchments based on the residual pattern of Fig. 3.1.

The coefficient of determination of this classification calculated by one-way ANOVA was R^2=25% which means that this classification explains 25% of the total spatial variance of the specific low flow discharges q_{95}. Although this is not much, the delineated regions were used as a basis for a regional regression model. The model consisted of five independent regionally restricted models. A statistical summary of these component models is presented in Table 3.1. Three out of the five regions are well represented by the regional models (regions 0, 1, and 3). However, the regression models for region 2 (Northern Calcerous Alps) and region 4 (Bregenzerwald) indicate very poor model performance which suggests that there may be significant heterogeneity of low processes within these regions. Note that R^2 represents the model goodness-of-fit coefficient of determination and hence does not fully capture the predictive performance for ungauged sites.

Table 3.1. Components of the regional regression model based on the residual pattern approach. R^2 denotes the goodness-of-fit coefficient of determination. Group numbers are as of Fig. 3.2. $\hat{q}_{95}$ are the estimated specific low discharges in $l \cdot s^{-1} \cdot km^{-2}$ and the units of the catchment characteristics are given in Table 2.1.

Group	Region	R^2	Model
0	N,E,SE of Austria, E-Tyrol, W-Tyrol	87%	$\hat{q}_{95}$ = -3.46 + 0.67*P - 0.19*L_{GL} - 0.03*G_F + 0.10*S_M
1	Central-Alps and Pre-Alps	60%	$\hat{q}_{95}$ = -0.81 + 0.69*P + 0.41*H_R - 0.52*H_M + 0.08*S_M
2	Part of the Northern Calcerous Alps	15%	$\hat{q}_{95}$ = 7.66 + 0.12*G_Q
3	Carinthia	82%	$\hat{q}_{95}$ = 1.51 + 1.02*D - 0.08*S_{MO}
4	Bregenzerwald (Vorarlberg)	32%	$\hat{q}_{95}$ =14.49 - 0.12*S_{MO}

The predictive performance of the complete regional regression model was finally checked by cross-validation. Ungauged catchments were assigned based on the regions in Fig. 3.2. The overall predictive performance was found as R^2_{cv}=63%. This is significantly better than the coefficient of determination of the classification (goodness-of-fit R^2=25%). This improvement is also apparent when comparing the residual pattern of the global regression model (Fig. 3.1) with that of the regional regression model (Fig. 3.3). The latter pattern is more random and the magnitudes of the residuals are significantly smaller. This means that there is a lot of value in using regionally restricted regression models over one single, global regression model.

Fig. 3.3. Residual pattern (cross-validation residuals) of the regional regression model based on the classification presented in Fig. 3.2. Positive residuals indicate an overestimation by the model.

3.3.2 Weighted cluster analysis

For the weighted cluster analysis, all catchment characteristics were standardised to zero mean and unit variance. A stepwise regression was then conducted between q_{95} and the standardised catchment characteristics in order to identify the most relevant catchment characteristics. The catchment characteristics so obtained and the respective ß-coefficients of the regression are presented in Table 3.2. From this analysis, winter precipitation P_w appeared as the most important characteristic. The positive ß-coefficient indicates that the low flows increase with winter precipitation which is plausible. Low flows also increase with the mean topographic slope, S_M, but they decrease with the portion of rock, L_R. L_R is highly correlated with altitude and it is likely that it is an indicator of catchment altitude rather than a physical consequence of rock cover *per se*. These ß-coefficients were used as weights in the weighted cluster analysis.

Table 3.2. Catchment characteristics and associated weights obtained by a preliminary stepwise regression. For symbols see Table 2.1.

Catchment characteristic	H_R	L_R	G_F	P_W	G_{GD}	S_M	G_Q
Weight (ß-coefficient)	0.22	-0.27	-0.12	0.42	0.13	0.33	0.11

A number of cluster analyses were carried out, combining different distance measures and linkage methods for a range of numbers of clusters. In each case, the homogeneity of the groups was assessed by a visual examination of Andrews plots. This comparison suggested that the hierarchical cluster analysis (*agnes*) that combines Ward's method and a Euclidean distance metric (using 10 clusters) was preferable to other methods and slightly preferable to the *pam* partitioning method (10 clusters, Euclidean metric). Fig. 3.4 shows the Andrews curves for the optimum classification method (*agnes*, 10 clusters). Each panel represents a cluster and each line corresponds to one catchment. The x_i of Eq. 1 are the catchment characteristics in Table 3.2 from left to right, standardised to zero mean and unit variance, and weighted by the ß-coefficients. I now examined the Andrews curves for homogeneity. Overall, the between-group variability is much larger than the within-group variability, although in groups 4 and 5 individual catchments appear to be different from the rest. However, given that I used a robustified regression technique which gives little weight to single outliers, I deemed the groups sufficiently homogeneous for the further analysis. I was hence able to avoid any subjective steps of manual re-classsification of outliers. The coefficient of determination of the classification by the weighted cluster analysis alone (i.e. without regional regressions) was R^2=56% which means that this classification explains 56% of the total spatial variance of the specific low flow discharges q_{95}. This is significantly more than that of the residual patterns approach.

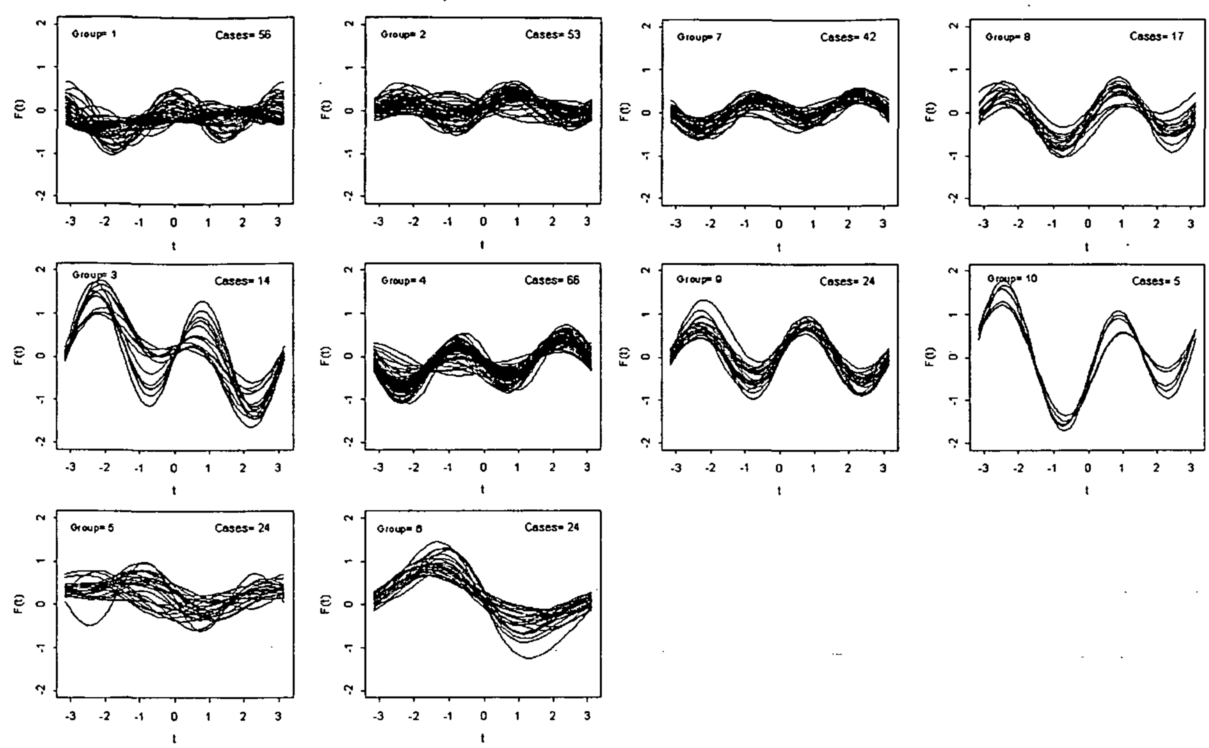

Fig. 3.4. Graphical representation of weighted catchment characteristics by Andrews curves. Each panel represents a cluster and each line corresponds to one sub-catchment.

In a next step, the clusters were plotted on a map (Fig. 3.5). Even though the cluster analysis did not use any information on the geographical location of catchments, most of the clusters are contiguous and there are only some of the Alpine catchments that are scattered in terms of their location. This result gives additional credence to the weighted cluster analysis approach. The spatial contiguity of the regions is apparently related to the spatial dependence of the weighted catchment characteristics.

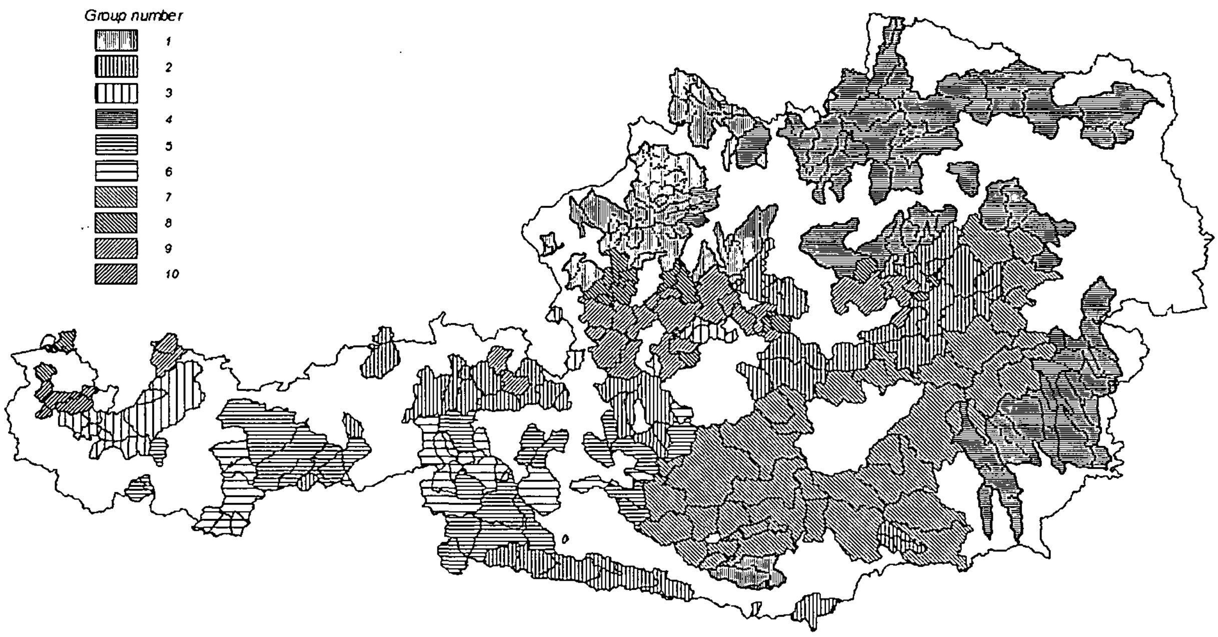

Fig. 3.5. Classification of catchments based on the weighted cluster analysis. Group numbers correspond to Fig. 3.4.

Regression models were then fitted to each of the regions independently. They are shown in Table 3.3. Most regions are represented rather poorly by the respective multiple regression models. For some regions (regions 8, 6, 4, 3), however, the model performance is very good. These differences may be related to the weights of the catchment characteristics. Constant weights have been used across the entire study domain which may be more appropriate in some parts of the domain than in others, as local deviations from the average behaviour may exist. The catchment characteristics used in the context of a weighted cluster analysis are hence not able to fully represent regional anomalies in the low flow patterns.

Table 3.3. Components of the regional regression model based on the weighted cluster analysis. R^2 denotes the goodness-of-fit coefficient of determination. Group numbers are as of Fig. 3.5. $\hat{q}_{95}$ are the estimated specific low discharges in $l \cdot s^{-1} \cdot km^{-2}$ and the units of the catchment characteristics are given in Table 2.1.

Group	Region	R^2	Model
1	Upper Austria	35%	$\hat{q}_{95} = 8.30 + 5.45*H_0 + 2.01*A - 1.08*L_F + 1.37*P_S$
2	Central Alps	32%	$\hat{q}_{95} = 8.20 + 2.07*G_Q + 3.62*P_W + 0.91*A$
3	Northern Calcerous Alps I	66%	$\hat{q}_{95} = 9.36 - 2.10*S_{MO} + 2.60*G_F$
4	Flatland and hilly terrain (N,E of Austria)	67%	$\hat{q}_{95} = 4.66 + 2.45*P - 0.30*G_F$
5	High Alps I (Tyrol, Carinthia)	44%	$\hat{q}_{95} = 7.75 + 3.26*P_S$
6	High Alps II (Tyrol, Carinthia)	70%	$\hat{q}_{95} = -1.67 + 4.24*S_M$
7	Low Alps (Styria and Carinthia)	41%	$\hat{q}_{95} = 5.89 + 1.69*H_+ - 0.87*S_{MO}$
8	Flyschzone (Upper- and Lower Austria)	75%	$\hat{q}_{95} = 17.35 - 1.98*G_F + 11.04*A$
9	Northern Calcerous Alps II	32%	$\hat{q}_{95} = 10.65 - 1.87*D + 3.55*G_Q$
10	Pre-alps (Bregenzerwald)	0%	$\hat{q}_{95} = 8.45$

Even though most of the clusters in Fig. 3.5 were coherent I did not judge them to be sufficiently contiguous for allocating ungauged catchments to regions in a unique way. I therefore approximated the grouping of Fig. 3.5 by a classification tree. The classification tree is shown in Fig. 3.6. The quality of approximation was assessed by the misclassification error which is the ratio of misclassified catchments and all classified catchments. The overall misclassification error is 21 out of 325 catchments (i.e. 21/325 = 0.06) which represents an excellent approximation to the grouping from the weighted cluster analysis. Fig. 3.6 shows in detail what catchment characteristics are most significant in representing the clusters. Precipitation (P_W) and topography (S_M, H_+, H_0, S_{ST}) are the most important catchment characteristics. This result is similar to the weights found by the regressions using standardised catchment characteristics in Table 3.2. Note that region 10 does not appear in the classification tree as the number of catchments is very small in this region. Also note that some of the catchment groups appear in two nodes (e.g. group 4) which means that this group consists of both terminal nodes in the classification tree.

The predictive performance of the complete regional regression model was finally examined by cross-validation, using the classification tree of Fig. 3.6 for assigning ungauged catchments to the regions. The cross-validation gave a predictive performance of R^2_{cv} =59%. Although the variance explained by the grouping alone was relatively large, the weighted cluster analysis does not appear to be as useful for delineating regions for the regional regressions.

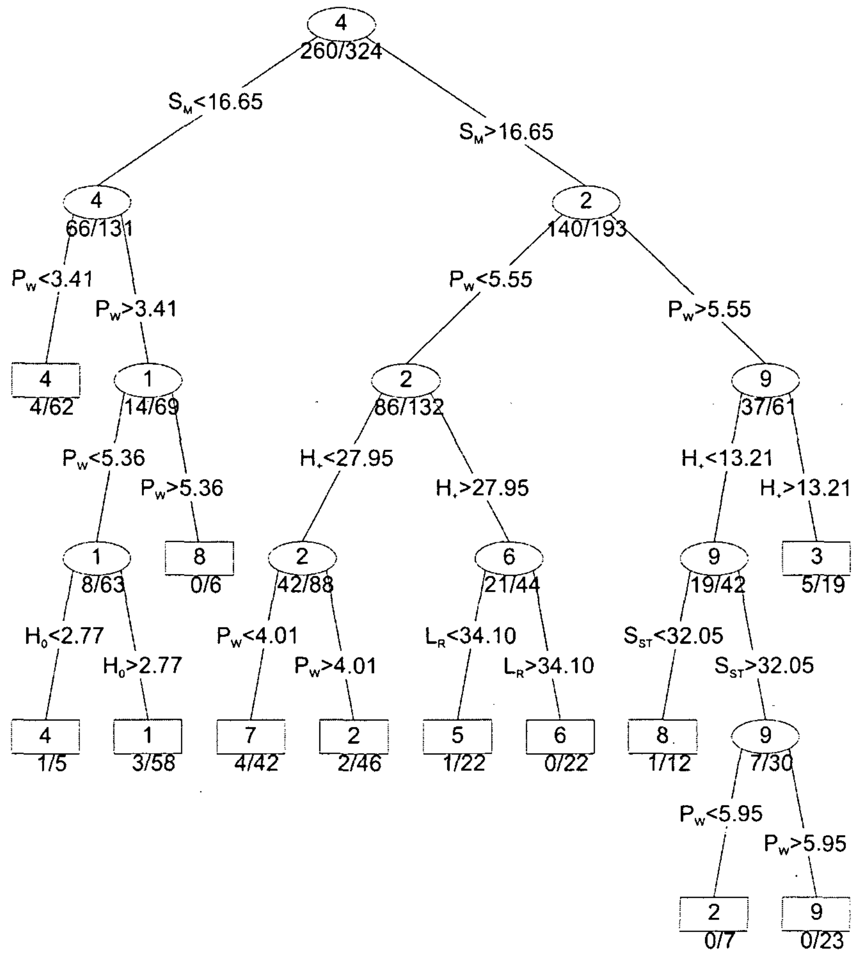

Fig. 3.6. Approximation of the classification based on the weighted cluster analysis by the classification tree. Ellipses indicate interior nodes, rectangles indicate terminal nodes (groups of catchments). Numbers within nodes represent group number (see Table 3.3), numbers below nodes represent misclassification error rate (misclassified catchments / classified catchments).

3.3.3 Regression tree

In the regression tree approach, the target variable was the specific low flow discharge q_{95} transformed by a square-root transformation. As descriptive variables, the complete set of non-standardised catchment characteristics was used. From an initial regression tree that was completely fitted to data, the optimal tree size was determined by 10-fold cross-validation.

Fig. 3.7 shows the cross-validated total deviance of trees of different sizes. Since the cross-validated deviance is a measure of the prediction error of the model, the optimum size of the regression tree is where the prediction error is at a minimum. Fig. 3.7 indicates that the optimum size is seven nodes. The initial regression tree was then pruned back to seven nodes using cost-complexity pruning (Clark and Pregibon, 1991).

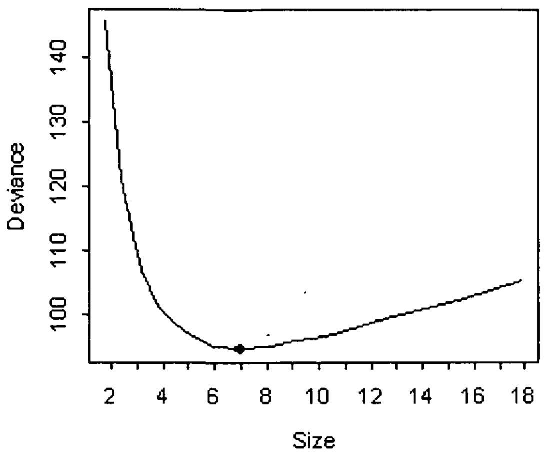

Fig. 3.7. Cross-validated deviance of regression trees as a function of number of splits. The minimum prediction error (cross-validated deviance) is obtained by a tree size of seven terminal nodes.

The regression tree so obtained is shown in Fig. 3.8 and divides Austria into seven regions. The structure of the regression tree indicates that the precipitation characteristics (P_S and P_W) are the most important factors for explaining the spatial variability of low flows. The second most important factors are three topographical indices, the range of altitude (H_R), mean altitude (H_M) and the fraction of catchment area exhibiting moderate slopes (S_{MO}). This result is similar to the weights of catchment characteristics (Table 3.2) used in the weighted cluster analysis, although geological characteristics do not appear in the regression tree of Fig. 3.8. The classification obtained from the regression tree hence partitions the landscape into regions of similar relief and similar seasonal precipitation. The variance explained by the grouping, calculated by one-way ANOVA, is 62%. This is the largest value of all classification approaches. This means that the regression tree is an excellent classification method if one is interested in finding groups that are most distinct in terms of both catchment characteristics and catchment response.

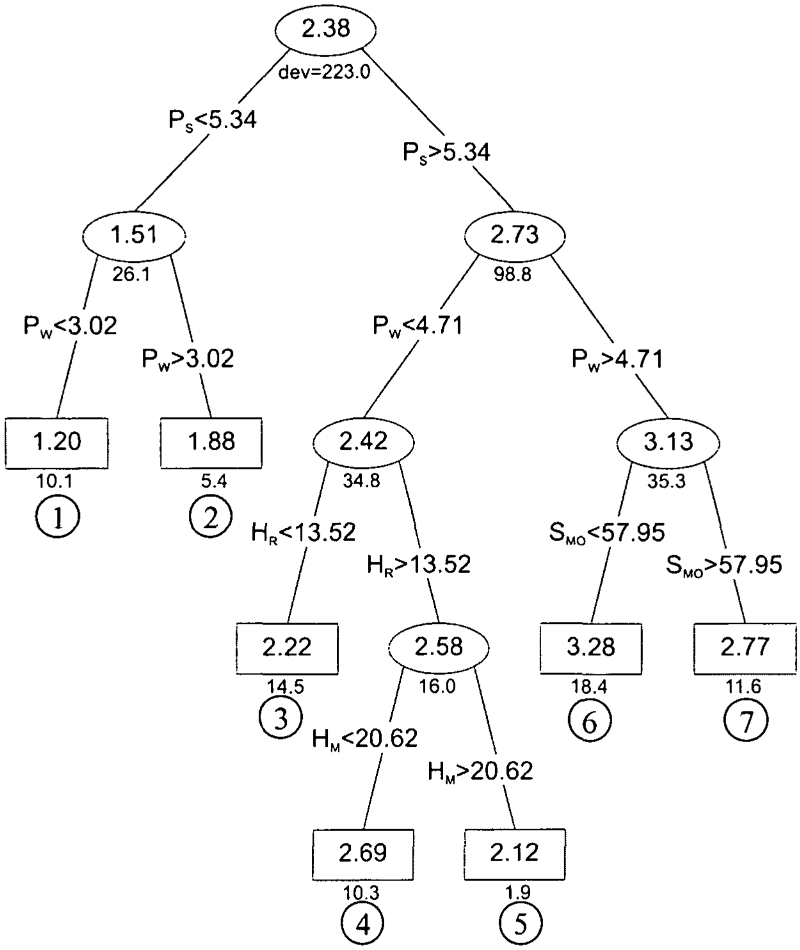

Fig. 3.8. Regression tree model. Ellipses indicate interior nodes, rectangles indicate terminal nodes (groups of catchments), circles represent group numbers. Numbers within nodes represent node means of square root-transformed specific discharge q_{95}, numbers below nodes represent node deviances in terms of square root-transformed specific discharge.

Fig. 3.9 shows the catchment groups plotted on a map. Overall, the regions are consistent with both the geographical classification of Austria and the main geological units. Some of the regions are contiguous while others are not. For instance, a mountain range consisting of limestone formations (Calcerous Alps) is divided into two regions (regions 6 and 7) by the strength of the relief and these are scattered in space. Similarly, region 3 (foothills of the Alps) consists of a thin band around the main Alps which is broken up into pieces in the south of Austria.

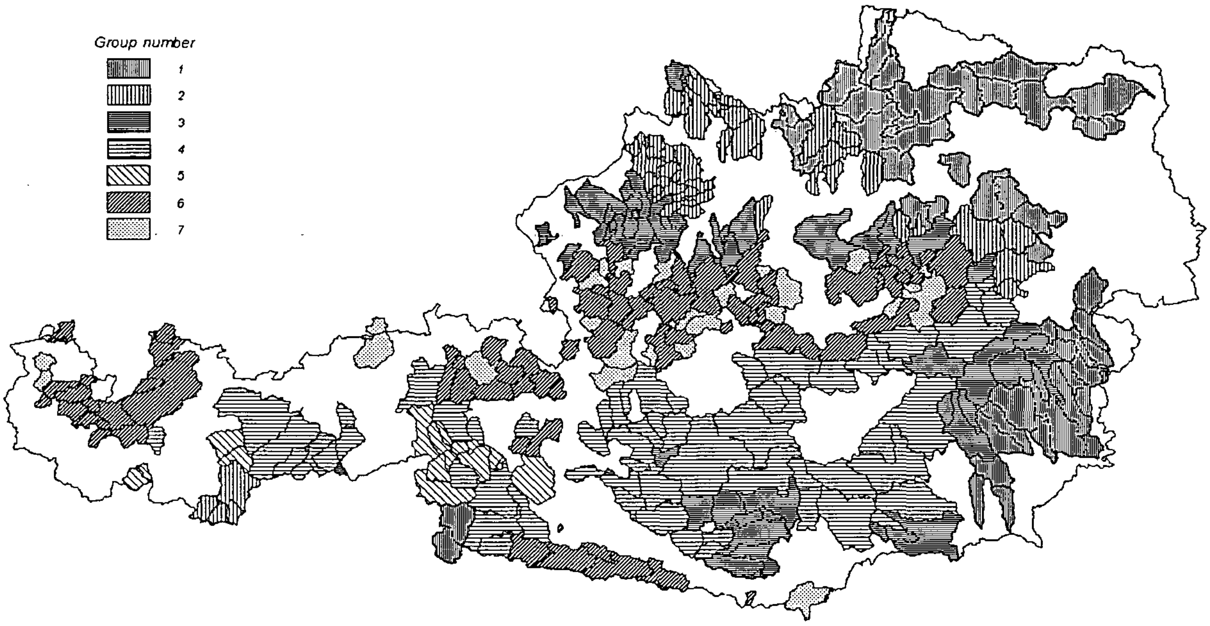

Fig. 3.9. Classification of catchments based on the regression tree approach. Group numbers correspond to Fig. 3.8 and Table 3.4.

Regression equations were now fitted to each region independently (Table 3.4). Two regions (regions 1 and 5) are well represented by the regression models, three regions (regions 2, 4, 7) exhibit a moderate model fit, and two regions (regions 2, 6) are poorly represented by the models. In the main, the goodness-of-fit of the regional regression model is similar to that of the weighted cluster analysis (Table 3.3).

Table 3.4. Components of the regional regression model based on the regression tree. R^2 denotes the goodness-of-fit coefficient of determination. Group numbers are as of Fig. 3.9. $\hat{q}_{95}$ are the estimated specific low discharges in $l \cdot s^{-1} \cdot km^{-2}$ and the units of the catchment characteristics are given in Table 2.1.

Group	Region	R^2	Model
1	Flatland and hilly terrain (N,E of Austria)	70%	$\hat{q}_{95}$ = -2.28 + 0.33*P + 0.04*G_{GS} + 0.25*H_M + 0.40*S_{ST}
2	Mühlviertel and Pre-alps (Lower Austria)	51%	$\hat{q}_{95}$ = 2.25 - 0.60*D - 0.08*L_{GL} + 1.91*P_W
3	Foothills of Alps	25%	$\hat{q}_{95}$ =-0.19 + 0.57*D + 0.03*G_{GD} - 0.10*G_{GS}
4	Central Alps	54%	$\hat{q}_{95}$ = -1.99 + 0.90*P - 0.20*G_T + 0.11*G_Q
5	High Alps (Tyrol)	67%	$\hat{q}_{95}$ = -9.57 + 0.30*S_M
6	Calcerous Alps I (S_{MG} < 57.95%)	13%	$\hat{q}_{95}$ = 14.68 + 0.19*L_A - 0.56*D
7	Calcerous Alps II (S_{MG} < 57.95%)	47%	$\hat{q}_{95}$ = 10.51 + 0.05*G_L - 1.47*P_W + 0.15*L_G

As the regions are not sufficiently contiguous to permit a unique allocation of ungauged catchments (Fig. 3.9) I allocated them by a classification tree. The cross-validation of regional regression estimates based on the regression tree approach was found as R^2_{cv} = 64%. This is significantly better than the estimates from the weighted cluster analysis where the performance was only R^2_{cv} = 59%. The main difference in terms of predictive performance of the two methods seems to be related to the allocation of ungauged catchments. The classification tree for the grouping in the weighted cluster analysis method exhibited a significantly larger misclassification rate than the classifications in the regression tree approach. It appears that one advantage of the regression tree method is a very efficient classification and allocation of ungauged catchments.

3.3.4 Regions of similar low flow seasonality

The last approach to catchment grouping considered in this study is based on types of low flow seasonality as defined by Laaha and Blöschl (2003). Most regions of the grouping of Laaha and Blöschl (2003) are contiguous with the exception of three sub-types of winter low flows (types A, B, C), which are scattered within the winter low flow dominated Alpine region. Since I focused in this approach on contiguous regions, these three types were merged into one single type of winter low flows. The resulting classification, consisting of eight regions of approximately homogeneous seasonality, is presented in Fig. 3.10. Since all regions are contiguous, the allocation of ungauged sites is well defined by their location and no re-classification was needed in the cross-validation procedure. Examples of the seasonal distribution of low flows for each of the regions are given in Fig. 3.11. From Fig. 3.11 it is quite clear that the seasonality of low flows shows major differences in the study domain, so one would expect seasonality to possess significant predictive power for delineating regions of similar low flow processes.

Regional regressions were now fitted independently to each of the regions. The results are summarised in Table 3.5. In most regions, the models fit well, with coefficients of determination ranging from 60% to 70%. The regression models for the Pre-Alps of Styria and Lower Carinthia (regions 3 and 4) exhibit even better coefficients of determinations of 89% and 83%, respectively. The exception is the Alpine, winter low flow dominated region (A-C), where the goodness-of-fit is only R^2=51%. This low coefficient is not surprising as three types of seasonality have been lumped into a single region.

In a final step, the predictive performance for the case of ungauged catchments was assessed by cross-validation. The cross-validated coefficient of determination for the approach based on seasonality regions was R_{cv}^2=70%. This is a better predictive performance than the other grouping methods. It appears that the stream flow characteristics as illustrated in Fig. 3.11 contain a lot of information highly relevant to low flow regionalisation.

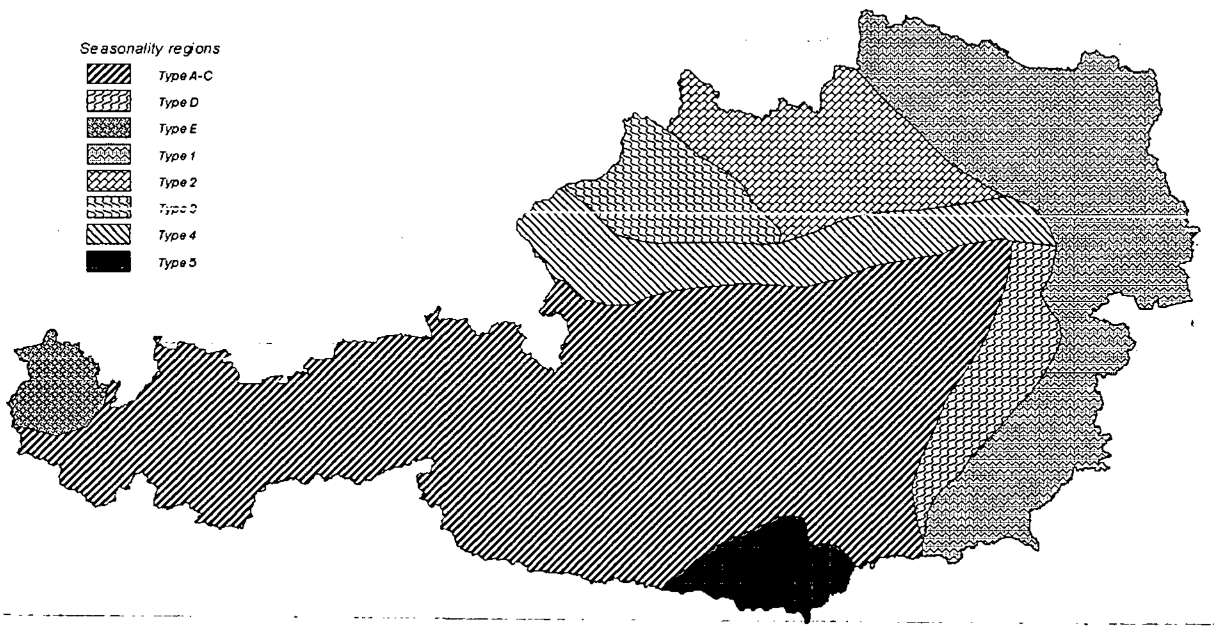

Fig. 3.10. Regions of similar seasonality in Austria. The labels of the regions correspond to the seasonality types in Fig. 3.11. Letters relate to winter low flows, numbers relate to summer low flows.

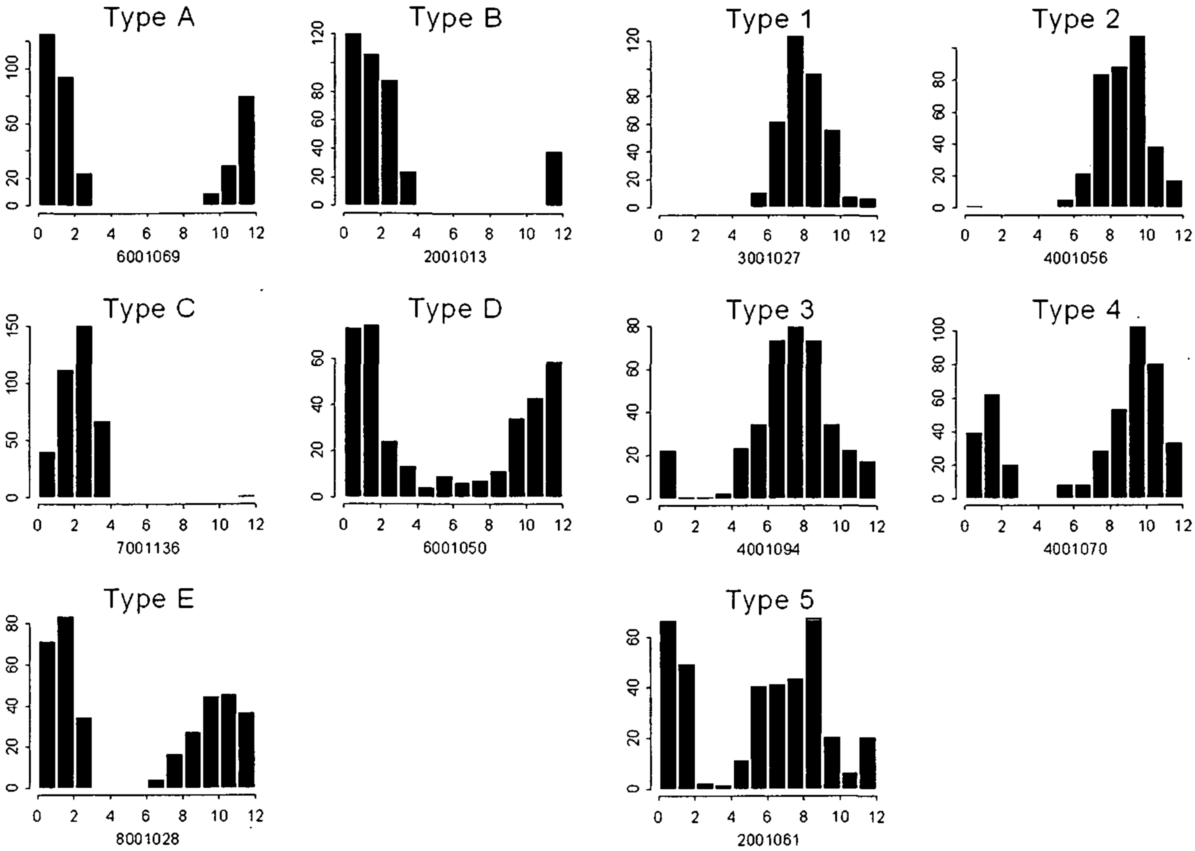

Fig. 3.11. Seasonality types of low flows in Austria illustrated by the non-exceedance frequencies of Q_{95} for each month for a typical catchment in each region. Letters relate to winter low flows, numbers relate to summer low flows (see Fig. 3.10).

Table 3.5. Components of the regional regression model based on regions of similar low flow seasonality. Group numbers are as of Fig. 3.10. $\hat{q}_{95}$ are the estimated specific low discharges in $l \cdot s^{-1} \cdot km^{-2}$ and the units of the catchment characteristics are given in Table 2.1.

Group	Region	R^2	Model
A-C	Alps	51%	$\hat{q}_{95} = 0.67 + 0.40*P + 0.17*G_Q - 0.01*G_C + 6.43*L_{WE} + 0.14*S_M - 0.04*S_R - 0.20*H_0$
1	Flatland & hilly terrain (N,E of Austria)	71%	$\hat{q}_{95} = -0.12 + 0.11*S_M + 0.05*G_{GS} + 0.02*G_C$
2	Bohemian Massif	64%	$\hat{q}_{95} = -3.31 + 1.96*P_W$
3	Foothills of Alps (Upper Austria)	68%	$\hat{q}_{95} = -10.04 - 0.76*D + 3.27*P - 2.22*H_0$
4	Flyschzone	63%	$\hat{q}_{95} = -6.17 + 0.06*G_L + 2.07*P_S - 0.06*G_W$
5	Lower Carinthia	83%	$\hat{q}_{95} = -17.48 + 3.56*D + 20.06*L_{WE}$
D	Pre-Alps (Styria)	89%	$\hat{q}_{95} = -7.99 + 1.08*P + 0.04*L_F$
E	Pre-Alps (Vorarlberg)	60%	$\hat{q}_{95} = 18.20 - 0.18*S_{MO}$

3.4 Discussion

3.4.1 Variance explained by grouping alone (ANOVA)

In a first step of comparing the methods of catchment grouping I examined the part of the variance (R^2) of specific low flows q_{95} that can be explained by the grouping alone without using regressions. The R^2 values are large if the variability between the estimated group means of q_{95} (SS_G) are large relative to the variability of the residuals (observed q_{95} minus group mean) within each group (SS_R). R^2 is a goodness-of-fit measure.

The regression tree approach performs best. Out of the total sum of squared specific low flow discharges of 5246 $l^2 \cdot s^{-2} \cdot km^{-4}$ the regressions tree explains 3244 $l^2 \cdot s^{-2} \cdot km^{-4}$, i.e., the variance explained by the grouping, calculated by one-way ANOVA, is 62%. This means that the regression tree is an excellent classification method if one is interested in finding groups that are most distinct in terms of both catchment characteristics and low flow catchment response. I believe that the reason for the good performance is that the splitting algorithm simultaneously maximises group homogeneity in terms of catchment characteristics and low flows. The regression tree is flexible in that it can choose the locally most relevant catchment characteristics, as each group can be subdivided by different decision criteria. This means that there is no need to select global similarity measures. This is an advantage for low flow regionalisation where global similarity measures may not exist. Application of the regression tree is straightforward and it provides an objective and robust classification. The most relevant catchment characteristics are apparent in the structure of the fitted regression tree. In contrast to the weighted cluster analysis, the regression tree is suitable for non-linear relationships between low flows and catchment characteristics which is an additional advantage. Using regression trees prior to linear regressions is therefore an attractive approach of combining the merits of non-linear and linear models.

The weighted cluster analysis approach performs second best and explains 56% of the variance of q_{95}. The weighting of the catchment characteristics by the coefficients of a regression model between q_{95} and catchment characteristics transfers information on low flow discharges to the distance measures used in the cluster analysis which seems to be a rather efficient approach. However, it should be noted that the weighted cluster analysis consists of 10 groups so one would expect a better goodness-of-fit than for the other methods. The seasonality regions and residual pattern approaches yield low R^2 values of 34% and 25%, respectively. It is clear that these two methods give little weight to finding regions that are most homogeneous in terms of low flows. It is also interesting that even though there are large differences in the goodness-of-fit between the groupings, they are all significant at the 95% level (Table 3.6).

Table 3.6. Variance explained by the groupings alone without using regressions. SS_G is the sum of squares of the mean group specific low flows q_{95}, SS_R is the sum of squares of the residuals of group mean minus observed q_{95} and SS_T is the total sum of squares of the observed q_{95}. Units of SS are $l^2 \cdot s^{-2} \cdot km^{-4}$. R^2 is the coefficient of determination of the group mean and the p-values are the empirical significance levels of F-tests of the group means.

Classification method	# of Groups	SS_G	SS_R	SS_T	R^2	p-value
Residual pattern approach	5	1319.1	3927.1	5246.3	25%	<0.001
Weighted cluster analysis	10	2911.8	2334.4	5246.2	56%	<0.001
Regression tree	7	3244.4	2001.9	5246.3	62%	<0.001
Seasonality regions	8	1787.0	3459.3	5246.2	34%	<0.001

3.4.2 Goodness-of-fit of regression models

In a second step I compared the goodness-of-fit of the regressions models for each of the groups identified by the various grouping methods. I also compared these goodness-of-fit values to the global regression model.

The global regression model uses four catchment characteristics as predictors. These are H_R (range of altitude), L_R (fraction of wasteland or rocks), G_F (fraction of Flysch) and P_W (average winter precipitation). The global model explains 62% of the variance in q_{95}. This is the same value as the best grouping method without regressions. It is interesting to compare this result to studies in the literature that used a similar number of catchments as in this section (325 catchments) and examined specific discharges as in this section, rather than discharges. Gustard et al. (1992) obtained R^2= 57% between Q_{95} standardised 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 south-west 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 obtained in this study are hence somewhat 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 sub-catchments rather than complete catchments which may make the catchment characteristics more relevant to low flow regionalisation.

The R^2 values of the component models vary vastly depending on the grouping method and the region (Tables 3.1, 3.3, 3.4, and 3.5). For the residual pattern approach, the R^2 values vary from 15-87%, for the weighted cluster analysis they vary from 0-75%, for the regression tree they vary from 13-70% and for the seasonality regions they vary from 51-89%. Overall the seasonality regions provide the best goodness-of-fit of the component regression models.

Aschwanden and Kan (1999) obtained R^2 values between 59% and 84% using the residual pattern approach and regional regressions of q_{95} in a very similar analysis to this section. This R^2 range is a similar order of magnitude found for the residual patterns approach in this study. The low goodness-of-fit for one of the regions of 15% in this study (region 2, see Table 3.1) may be related to karstic effects as this is a limestone area of the Pre-alps. It is possible that the specific discharges derived from the observations are inaccurate as the hydrologic catchment areas in these regions may differ from the topographic catchment areas but are not well known. Most other studies in the literature used discharge rather than specific discharge and so are not directly comparable to the results in this section. As catchment size usually explains around 80-90% of the variability of low flow discharges (see, e.g., Dingman and Lawlor, 1995; 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 (1992), for example, reported R^2 values of more than 90%.

3.4.3 Predictive performance of regional regressions for various grouping methods

The global regression model, i.e., without using any grouping, gives an R^2_{cv} = 57% in the cross-validation mode (Table 3.7). This is a significantly lower value than the goodness-of-fit R^2 of the global model (R^2 = 62%). Part of the difference may be related to an overfitting of the global regression model although this is unlikely to explain the full difference as only four catchment characteristics have been used as predictors. A more important reason for the difference may be heteroscedasticity of the sample and the existence of outliers which contribute significantly to the estimation error. This issue is discussed later in this section.

In the regional regression models, the grouping based on seasonality regions performs best (Table 3.7). The explained variance, in a cross-validation mode, is R^2_{cv} = 70%. This is significantly more than for the global model (R^2_{cv} = 57%). It appears that delineating regions based on the seasonality of low flows provides information on the hydrological regimes not captured by the catchment characteristics and the low flow discharges. Note that all four grouping methods use information on the low flow discharge q_{95}, albeit in different ways, and all grouping methods, with the exception of the seasonality regions approach, use catchment characteristics as well.

It is interesting that this performance is significantly better than that of an alternative model proposed in section 2 which gave R^2_{cv} =58% for the same data set. The model of section 2 is a global regression model that uses a region index as a predictor variable in addition to the catchment characteristics. This index value differs by the region and has been calibrated. It appears that the seasonality types are not mainly related to the magnitude of the low flows, so they are not very efficient as a predictor variable. However, the relationship between catchment characteristics and low flows appears to be significantly different for different seasonality regions. Various processes may combine in different ways in different seasonality regions, as a result of differences in the hydrologic and climatic regime. The seasonality grouping is hence very efficient in the context of the regional regression approach of using separate regressions in each of the groups.

The favourable performance of the grouping method based on seasonality regions may be related to the striking differences in low flow seasonalities in the study domain (Fig. 3.11). These differences are clearly related to different processes. Winter low flows are a result of the retention of solid precipitation in the seasonal snow pack of the catchment and of freezing processes in the soils. In contrast, summer low flows are related to the relatively large moisture deficits in the lowland regions of Austria during summer. It appears that grouping the domain according to low flow seasonalities does capture some of the effects of these processes.

The regression tree grouping performs second best (R_{cv}^2=64%) and the performance of the residual pattern approach is similar (R_{cv}^2=63%). As compared to the global regression model (R_{cv}^2=57%) there is some improvement in the performance although it is not large. The weighted cluster analysis, only yields a minor improvement (R_{cv}^2=59%) over the global model. The improvement of the regional regression models (including grouping) over the global model (without grouping) is related to the degree of non-linearity that can be captured by the grouping method. In the weighted cluster analysis method, the performance is similar to the fully linear global model, so does poorly in representing any non-linearity. The other two methods do capture some of the non-linearity.

It is interesting that the relative performance of the grouping methods combined with regional regressions differs from the relative goodness-of-fit of the grouping methods alone. While for the grouping methods alone the regression tree approach performed best, it is the grouping based on seasonality regions that performs best when combing the grouping with regional regressions. It is clear that in the latter case, the important feature the catchment groupings need to capture is the way the catchment characteristics are related to low flows rather than the low flows themselves. Within group homogeneity and between group heterogeneity in terms of low flow discharges are hence not a good indicator for the predictive performance of low flow regional regressions. Cross-validation of the regression estimates is certainly a preferable way of measuring the performance of regionalisation methods.

It should be noted that in the residual pattern and the seasonality region approaches the regions were not updated in the cross-validation procedure. This was because the regions were deemed sufficiently contiguous not to change much if a single catchment is added. It is possible that the cross-validation performance of these two methods may very slightly decrease if the regions were updated but given the relative magnitude of the cross-validated coefficients of determinations it is unlikely that this will change the ranking of the predictive performance of the methods.

Table 3.7. Predictive performance of regional regression models based on different grouping methods. R^2_{cv} is the coefficient of determination of cross-validated estimates.

Catchment grouping	**Allocation of ungauged site via**	R^2_{cv}
Residual pattern approach	Geographic location	63%
Weighted cluster analysis	Classification tree	59%
Regression tree	Classification tree	64%
Seasonality regions	Geographic location	70%
No grouping	-	57%

3.4.4 Heteroscedasticity, outliers and bias

As a final step of assessing the methods of catchment grouping I examined scatter plots of predicted vs. observed specific low flow discharges q_{95} (Fig. 3.12). 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. Overall the relative scatter of the methods (Fig. 3.12) corresponds well with the cross-validated coefficients of determination in Table 3.7 and it is clear that the seasonality regions approach performs best and the weighted cluster analysis approach performs poorest. The weighted cluster analysis approach overestimates low flows significantly for three catchments and the magnitude of the estimation error is relatively large for a number of catchments. The outliers tend to increase with q_{95}, which suggests that the predictions are heteroscedastic. One would usually apply a variance-stabilising transformation in this case, such as taking the logarithms of q_{95}, but preliminary analyses showed that this transformation improved the heteroscedasticity of the transformed data but did not improve the heteroscedasticity of the residuals of the back-transformed predictions. The residual pattern approach generally performs quite well although it gives negative predictions of q_{95} for two catchments and a few outliers. The regression tree approach performs equally well for the bulk of the catchments, but appears slightly superior to the residual pattern approach as far as outliers are concerned. The approach based on seasonality regions performs best. The points are scattered around the

1:1 line indicating low prediction errors for a broad range of discharges. The scatter is almost homoscedastic and there are only a few minor outliers.

One apparent deficiency of all models is the large scatter and clear bias for very wet catchments. In catchments where observed specific low flow discharges are more than about 12 $l{\cdot}s^{-1}{\cdot}km^{-2}$ the low flows are consistently underestimated, 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 $l{\cdot}s^{-1}{\cdot}km^{-2}$ corresponds to 378 mm of low flow depth per year which is a relatively large value for Austrian conditions. In all catchments in the q_{95}>12 $l{\cdot}s^{-1}{\cdot}km^{-2}$ range, with the exception of two catchments, limestone is the main geologic formation (75% of the catchment area on average) so karst effects are likely to occur. It is possible that the specific discharges derived from the observations are inaccurate as the hydrologic catchment areas in these regions may differ from the topographic catchment areas. A more detailed analysis is needed to ascertain the extent to which the low flow observations in these catchments are actually biased. It should also be noted that it is not uncommon for regionalisation models to have a tendency for underestimating large values. For example, the flood regionalisation analysis of Merz and Blöschl (2004a) showed that flood quantiles in the same study area were consistently underestimated by their method for catchments with above-average specific flood discharges.

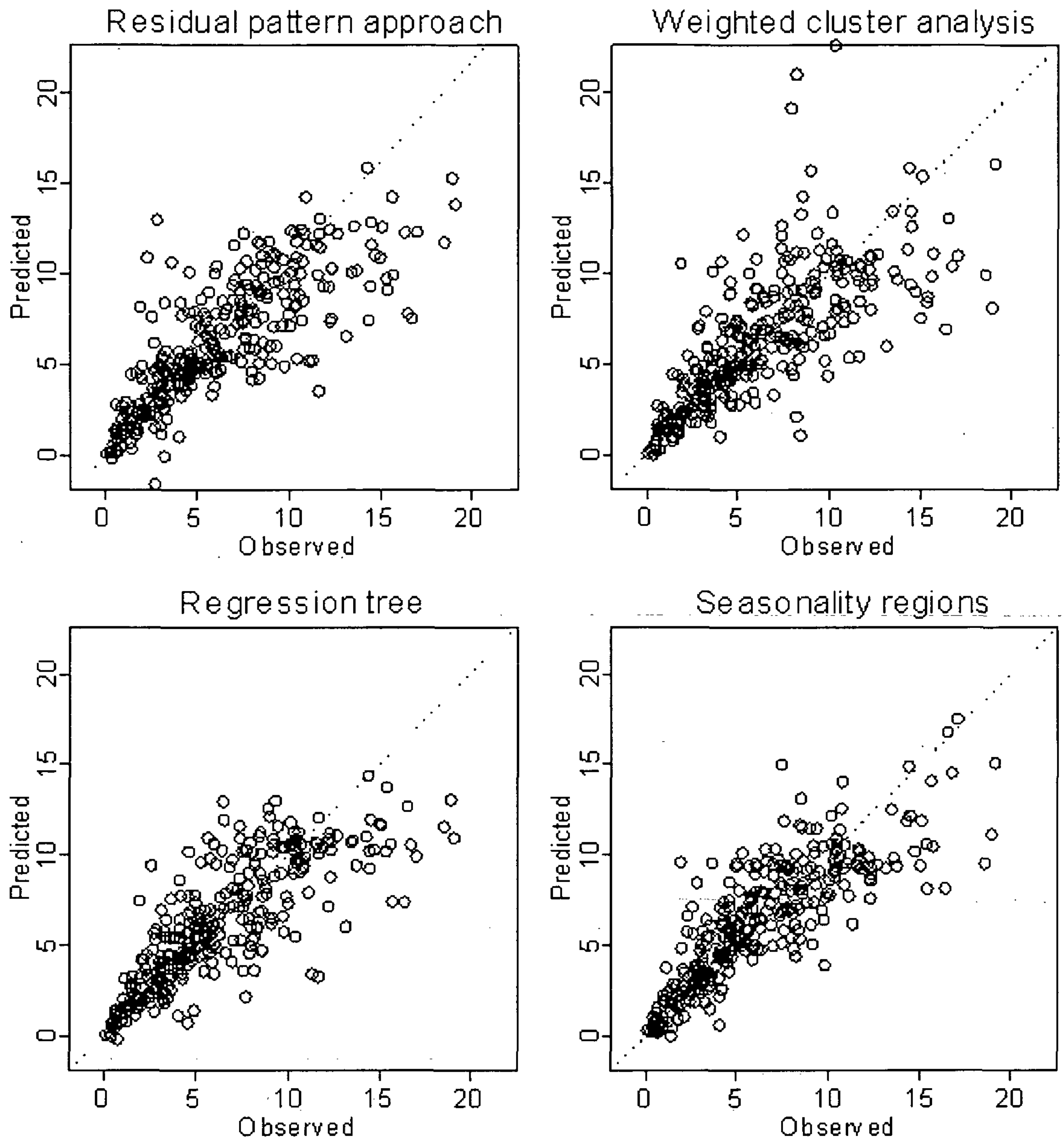

Fig. 3.12. Scatter plots of predicted vs. observed specific low flow discharges q_{95} ($l \cdot s^{-1} \cdot km^{-2}$) in the cross-validation mode. Each panel corresponds to one regional regression model and each point corresponds to one catchment.

3.5 Conclusion

I compared four catchment grouping methods in terms of their performance in predicting specific low flow discharges q_{95}. These methods are the residual pattern approach, weighted cluster analysis, regression trees and an approach based on seasonality regions. In a first step I examined the part of the variance (R^2) of specific low flows q_{95} that can be explained by the grouping alone without using regressions. In this comparison, the regression tree approach performs best and explains 62% of the spatial variance. This means that the regression tree is an excellent classification method if one is interested in finding groups that are most distinct in terms of both catchment characteristics and low flow catchment response. In a second step I compared the goodness-of-fit of the regressions between catchment characteristics and q_{95} for each of the groups identified by the various grouping methods. Here, the seasonality regions approach provides the best goodness-of-fit of the component regression models and explains between 51 and 89% of the spatial variance of q_{95}, depending on the region. A global regression model explains 57% of the variance in q_{95}. It uses range of altitude, fraction of rock, fraction of Flysch, and average winter precipitation as the predictor variables. In a third step I examined the predictive power of the regional regressions based on each of the grouping methods using leave-one-out cross-validation. The cross-validation represents the regionalisation error of the low flows one has to expect for the case of ungauged sites. Among the grouping methods tested here, the grouping based on seasonality regions performs best and explains 70% of the variance in a cross-validation mode. The favourable performance of this grouping method is likely related to the striking differences in seasonal low flow processes in the study domain. Winter low flows are a result of the retention of solid precipitation in the seasonal snow pack of the catchments and of freezing processes in the soils while summer low flows are related to the relatively large moisture deficits in the lowland regions of Austria during summer. The regression tree grouping performs second best (explained variance of 64%) and the performance of the residual pattern approach is similar (explained variance of 63%). The weighted cluster analysis only explains 59% of the spatial variance of q_{95} which is only a minor improvement over the global regression model, i.e. without using any grouping, in a cross-validation mode (explained variance of 57%). An analysis of the sample characteristics of all methods suggests that, again, the grouping method based on the seasonality regions has the most favourable characteristics although all methods tend to underestimate specific low flow discharges in the very wet catchments.

This study has examined a single low flow characteristic (q_{95}) and it would be interesting to see whether the relative performance of the grouping methods remains the same if different characteristics are examined. There is also some potential in using short discharge series in the low flow regionalisation. Short series and, perhaps, snapshot discharge measurements may be available in a much larger number of catchments. The value of short time series for low flow regionalisation will be examined in section 4.

4 The value of short stream flow records in regional low flow estimation

4.1 Introduction

Characteristic values of low flow discharge are needed for a number of purposes in water resources management and engineering including environmental flow requirements, water uses and discharges into streams, and hydropower operation (Smakhtin, 2001). The interest usually resides in characteristic low flow values that represent the long-term average behaviour of low flows, commensurate with the life time of a structure or the design period of a management measure. Due to climatic variability and other sources of variability that occur over short time scales, low flow characteristics estimated from a few years of stream flow data deviate from the long-term average. Because of this, it is usually recommended to use stream flow records of 20 years or more for low flow estimation (Tallaksen and van Lanen, 2004). However, in many countries, for a significant part of the gauged catchments the records are shorter than the recommended period. While these short records are unlikely to provide the full information of long records it is clear that they do provide some information which may be used in estimating the long term low flow characteristics for these stream gauge locations.

A number of methods exist for inferring the long-term low flow characteristics from short records. These methods all account for climatic variability, in some way, and are therefore referred to as climate adjustment methods. They are used to estimate the low flow characteristics for the site of interest (which I term the subject site) where a short stream flow record is available, based on stream flow data from other catchments (which I term donor sites) where long records are available. The climate adjustment is usually limited to random effects (e.g. random climate variability and measurement errors) and cyclic effects (e.g. climatic variation), while systematic effects such as trends caused by climatic change or changes of the catchment response characteristics as a result of human activities are often treated in an explicit way rather than by climate adjustment procedures (e.g. Kundzewicz and Robson, 2000).

Climate adjustment methods consist of three main steps, (a) selecting donors, (b) calculating adjusted low flow characteristics at the subject site for each donor by record augmentation techniques, and (c) combining the adjusted values associated with each donor to obtain an estimate of the long-term low flow characteristic at the subject site (Robson, 1999).

Donor-sites are often selected by expert judgement based on the hydrogeology and climate in the study region. More formal procedures of selecting donor sites make use of spatially contiguous regions, spatial distance, catchment characteristics, or a combination thereof. In a number of countries, mapped regions exist that are spatially homogeneous with respect to low flows or other flow characteristics (e.g. NERC, 1975; section 3) and one option is to select the donor from the region where the subject site is located. Spatial patterns of the seasonal occurrence of low flows can be used to assist in the identifications of homogeneous regions (section 2; Merz et al., 1999). Spatial proximity, i.e. using the nearest stream gauge is also a widely used method of donor selection (Stedinger et al., 1992) which is particularly useful if the donor site is downstream of the subject site and the catchment area is not much larger. An alternative is the use of catchment characteristics such as geology and mean annual precipitation. Catchment characteristics play an important role in a range of hydrologic regionalisation methods (e.g. Nathan and McMahon, 1990; Holmes et al., 2002; section 2 and 3). There are numerous ways of formulating similarity measures based on catchment characteristics. The most straightforward way is a Euclidean distance measure, i.e. a linear combination of the squared differences of the catchment characteristics of the subject and donor sites. The catchment characteristics can be scaled to unit variance and they can be weighted, and here again, there exist a range of possibilities (Nathan and McMahon, 1990). Methods for visualising similarity in catchment characteristics can assist in the expert assessment of choosing a suitable donor catchment (Andrews, 1972). If the flow record at the subject site is not too short, the donor selection can also be based on the correlation of annual low flows between the subject and donor sites. The catchment that exhibits the largest correlation with the subject site can then be used as a donor. An example in the context of climate adjustment of flood records is Robson (1999) who used rank correlation coefficients between annual values of subject and donor sites. More details of various measures for assessing the similarity of catchments in the context of low flow regionalisation are given in section 3).

Once one or more donors have been identified, some sort of record augmentation technique is needed to take advantage of the climate variability signal in the longer record of the donor for estimating the flow characteristics for the subject site. Fiering (1963) and Matalas and Jacobs (1964) proposed a theoretical framework of minimum variance stream flow record augmentation procedures. The basic idea of these methods is to employ the cross-correlations between a long record and a short record to estimate the mean and the variance of flow at the (short record) subject site. Vogel and Stedinger (1985) improved on these estimators and assessed them by Monte-Carlo experiments for annual flood peaks and monthly stream flows. They found very significant gains of information provided the correlations were large and the record length of the donor was much larger than that of the subject site. However, they also stated that the estimates are likely to be poor if the stream flow record of the subject site is too short. Using this method, Vogel and Kroll (1991) examined the value of stream flow record augmentation procedures in low-flow and flood-flow frequency analysis for 23 catchments in Massachusetts. They defined an effective record length as the length of an unadjusted record that gives the same estimation error as a shorter record that is adjusted. They found that the record augmentation increased the effective record length but the presence of serial correlations in the flow data decreased the effective record length. The net effect of these two components was a gain in information for subject site records shorter than 30 years only. The value of the record augmentation procedure also depended on the flow characteristics examined and slightly increased with the return period of the low flow characteristics.

In case of multiple donors, low flow characteristics adjusted by each donor are usually combined by some statistical average to obtain the low flow estimate at the subject site. Robson (1999) combined adjusted values from multiple donors by a weighted geometric average. The weights were calculated from the distance between subject site and donor, the length of the overlap period and the additional years of data provided by the donor based on a rank correlation coefficient of annual values.

When a number of base flow measurements can be obtained at an otherwise ungauged site they can be correlated with concurrent stream flows at a nearby gauged site for which a long flow record is available. This is sometimes termed the base flow correlation procedure (Hayes, 1992; Stedinger et al., 1992). The base flow spot gaugings can be thought of as the limiting case as the record length approaches zero. In this method, parameters of a linear regression model estimated from concurrent stream flows are used to infer the low flow characteristic at the subject site from that of the donor site. This is typically done for $Q_{d,T}$ low flows (*d*-day low flow discharge for a return period of *T* years, Demuth et al., 2004) but the method can be subject to considerable error if only a few discharge measurements are used (Stedinger et al., 1992). If base flow measurements are only available for a single point in time one cannot estimate regression parameters but one can assume that the spot gauging is representative of the low flow characteristic of interest, provided the flow conditions of the streams in the region on the day of measurement are similar to the low flow characteristic of interest. Kroiß et al. (1996), for example, were interested in finding the low flow characteristic Q95 (i.e. the discharge that is exceeded on 95% of all days) for numerous sites in the Lainsitz region, Northern Austria, to assist in the siting of wastewater treatment plants. They conducted stream flow spot measurements during a few days of a low flow autumn period and adjusted the discharge values so obtained by scaling them by Q95 observations from gauged catchments in the region. Although they did not test the estimates against longer records, they were able to interpret the regional patterns of Q95 based on the hydrological heterogeneity in the region.

The climate adjustment techniques in the literature for estimating stream flow characteristics from short records have, to my knowledge, never been compared in a comprehensive way for the case of low flows and it is so far unclear which of the methods performs best. The aim of this section therefore is to examine the relative performance of different climate adjustment techniques for estimating low flow characteristics from short stream flow records. I will address the following questions: (i) How accurate are low flow characteristics estimated from short records and what is the role of the record length? (ii) What is a suitable donor selection method? (iii) What are the relative merits of various methods of exploiting the information of a donor? (iv) What is the value of using short stream flow records at the subject site over using data from neighbouring sites only (i.e. regionalisation)? The analyses will be made for a comprehensive data set in Austria and the low flow characteristic chosen is the Q95 flow quantile which is the discharge that is exceeded on 95% of all days for one particular site. The value of each technique is assessed by using hypothetically shortened stream flow records and comparing the Q95 estimated from the shortened records with the Q95 estimated from the full record.

The section is organised as follows: Section 4.2 summarises the data used. Section 4.3 details the methods of climate adjustment examined in this section which consist of three donor selection techniques and two record augmentation techniques. The evaluation procedure based on hypothetically shortened records is presented in section 4.4. Results of the comparisons are presented in section 4.5 and discussed in section 4.6. Section 4.7 gives conclusions.

4.2 Data

4.2.1 Study area

The study has been carried out in Austria which is physiographically quite diverse. There are three main zones in terms of the geographical 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). Elevations range from 117 to 3798 m a.s.l.. Austria has a varied climate with mean annual precipitation ranging from 500 mm in the eastern lowlands 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 hydrologic regimes (Kresser, 1965) and low flows exhibit important regional differences in terms of their quantity and their seasonal occurrence (Laaha and Blöschl, 2003).

4.2.2 Discharge data and selection of gauges

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 (section 2). 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, I have chosen to give priority to a consistent observation period to make all records comparable in terms of climatic variability. I use all of these 325 catchments as possible donor sites.

For the subject sites, i.e. the sites where I test the value of short stream flow records, I have chosen to only use those catchments that do not have an upstream neighbouring gauged catchment. I did this for ease of comparison with regionalisation studies in the study area which were based on discharges of catchments without upstream gauges and on discharges of residual catchments between subsequent gauges (section 3). Also, this tends to be a set of relatively small catchments which are usually of most interest in estimating low flows from short records. One of the donor selection techniques requires the availability of downstream flow data and I therefore excluded those catchments that did not have a downstream neighbour. What remained was a set of 132 gauged catchments which I used as subject sites in this section. These are the sites for which I analyse the effects of record length and climate adjustment method on estimating low flow characteristics.

4.2.3 Low flow characteristics

The low flow characteristic chosen in this section is the flow quantile Q95, i.e. the discharge equalled or exceeded during 95% of the observation period (Pr(Q≥Q95=0.95)). Values of Q95 have been calculated for all 325 gauges from continuous daily discharge records between 1977 and 1996 and are assumed to represent the long-term averages of Q95. The statistical characteristics of the Q95 discharges of the 132 catchments used as subject sites are given in Table 2 along with those of the specific discharges q95 and the catchment areas.

Table 4.1. Characteristics of the 132 catchments used as subject sites. Q95 are low flow discharges, q95 are specific low flow discharges, area is the catchment area. The percent values are the quantiles.

	Minimum	25%	50%	75%	Maximum	Mean
Q95 (m^3/s)	0.013	0.194	0.449	0.927	3.890	0.692
q95 ($l/(s.km^2)$)	0.65	3.32	5.93	8.81	16.76	6.24
Area (km^2)	8.7	40.7	77.9	145.0	479.0	114.8

4.2.4 Catchment characteristics

One of the investigated donor selection techniques is based on hydrological similarity of catchments. To define the similarity measures, I used 31 catchment characteristics (Table 2). They relate to catchment area (A), topographic elevation (H), topographic slope (S), precipitation (P), geology (G), land use (L), and drainage density (D). All percent values with the except of mean slope (S_M) relate to the area covered by a class relative to the total catchment area. Some of the catchment characteristics had to be adapted from the original sources to make them more useful for regionalisation. For instance, the original classification of the metallurgic map used here distinguishes 670 geological classes from which I derived 9 hydrogeological classes I deemed relevant for low flow regionalisation. One of them is termed source region which is the percent area where the density of springs is large. In a similar vein, I condensed the original Corine Landcover classification (Aubrecht, 1998) into nine land-use classes. Three precipitation characteristics of average annual, summer and winter precipitation from 1977 to 1996 estimated by the regionalisation 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 catchment 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 2.

Table 4.2. Statistical summary of the characteristics of the 325 catchments used in this section.

Acronym	Variable description	Units	Min.	Mean	Max.
A	Catchment area	km²	7.00	313.31	7012.10
H_0	Altitude of streamgauge	m	159.00	591.38	2215.00
H_+	Maximum altitude	m	298.00	1862.29	3770.00
H_R	Range of altitude	m	82.00	1270.91	3324.00
H_M	Mean altitude	m	231.90	1103.56	2944.70
S_M	Mean slope	%	0.03	0.25	0.56
S_{SL}	Slight slope	%	0.00	25.99	100.00
S_{MO}	Moderate slope	%	0.00	47.30	93.00
S_{ST}	Steep slope	%	0.00	26.62	80.00
P	Average annual precipitation	mm	467.06	1082.31	2103.40
P_S	Average summer precipitation	mm	293.75	652.20	1208.10
P_W	Average winter precipitation	mm	155.33	430.09	895.30
G_B	Bohemian Massif	%	0.00	10.09	100.00
G_Q	Quaternary sediments	%	0.00	5.88	93.00
G_T	Tertiary sediments	%	0.00	15.05	100.00
G_F	Flysch	%	0.00	6.87	100.00
G_L	Limestone	%	0.00	26.04	100.00
G_C	Crystalline rock	%	0.00	26.97	100.00
G_{GS}	Shallow groundwater table	%	0.00	1.29	18.30
G_{GD}	Seep groundwater table	%	0.00	6.06	76.10
G_{SO}	Source region	%	0.00	1.35	35.20
L_U	Urban	%	0.00	0.53	7.79
L_A	Agriculture	%	0.00	19.62	97.30
L_C	Permanent crop	%	0.00	0.12	20.30
L_G	Grassland	%	0.00	20.60	71.70
L_F	Forest	%	0.00	47.45	100.00
L_R	Wasteland (rocks)	%	0.00	0.07	9.61
L_{WE}	Wetland	%	0.00	9.05	81.20
L_{WA}	Water surfaces	%	0.00	0.39	14.60
L_{GL}	Glacier	%	0.00	1.78	43.80
D	Stream network density	m/km²	160	790	1320

4.3 Climate adjustment techniques

4.3.1 General concept

My approach to climate adjustment consists of three steps: (a) selection of appropriate donors for each subject site, (b) calculation of adjusted low flow characteristics for the subject site from data of each donor (i.e. record augmentation), and (c) combination of adjusted values associated with each donor in the case of multiple donors. I examine three donor selection techniques plus the case of no donor (i.e. no adjustment), and two record augmentation methods. The techniques are presented below.

4.3.2 Donor selection

4.3.2.1 No donor

In the first technique, no donor is selected which corresponds to the case of calculating low flow characteristics from short records without any adjustment for climatic variability. The estimation error of this technique will be a benchmark against which the other methods are to be tested. Any of the other methods should improve on this benchmark case.

4.3.2.2 Downstream site

The second technique uses the nearest gauge at the same stream as the subject site. The rationale of this technique is that the nearest down stream gauge is usually close to the subject site and there will be some overlap in catchment area, so they should have similar hydrological and climatic catchment characteristics. One drawback of the downstream site technique is that only one gauge is considered as a donor. Because of this, the method is probably less robust than the methods that use more than one donor, particularly for catchments where land use changes have occurred and/or some constructions have taken place at the stream. The procedure consists of a single step:

(a) Select adjacent downstream gauge at the same stream as a donor;

4.3.2.3 Catchment similarity

In the third technique, donors are selected according to the similarity of physiographic catchment characteristics. The basic assumption of this method is that hydrological processes are related to catchment physiography, so discharges from physiographically similar catchments should experience similar effects of climatic variability. The difficulty with this approach is that information on catchment similarity is probably contained in a large number of catchment characteristics and it is not straightforward to find a similarity measure that uses the information of the most relevant characteristics. Following the idea of Nathan and McMahon (1990), I selected relevant catchment characteristics by a stepwise multiple regression analysis between Q95 and the catchment characteristics and weighted them according to the coefficients in the regression model. I then assessed physiographic similarity of subject sites and possible donors by the Euclidean distance in the space of the weighted catchment characteristics.

In addition to physiographic catchment similarity, one can expect that similar catchments should lie in the same climatic region for similar impacts of climatic variation to occur. I adopted the classification of Austria into eight regions of section 2. These are regions that exhibit similar low flow seasonality, so one can assume that they are also suitable for identifying catchment similarity in terms of climatic impact. The selection of physiographically similar donors was then limited to gauges located in the same seasonality zone as the subject site. The stepwise regression mentioned above was performed independently for each of these regions. The procedure consists of the following steps:

(a) Select all gauges within the seasonality zone of the subject site as possible donors;

(b) Perform a stepwise regression between Q95 and catchment characteristics to determine the most relevant catchment characteristics for assessing physiographic similarity;

(c) Weight the selected catchment characteristics by the coefficients of the regression model;

(d) Calculate Euclidean distances between subject site and all possible donors in the space of weighted catchment characteristics;

(e) Select the most similar site (i.e. the site that exhibits the shortest Euclidean distance) as a donor.

4.3.2.4 Correlation of annual low flows

The fourth technique is based on the procedure of Robson (1999). Although the procedure of Robson (1999) was designed for adjusting flood characteristics there may be some similarity of climate variability effects with low flows. I therefore think it is worth applying the method of Robson (1999) to the case of low flows. The selection of donors proceeds in two main steps. The first step identifies potentially useful sites on the basis of spatial proximity and the possible gain of information from each donor. The second step refines the selection on the basis of the correlations of annual low flows between the subject and donor sites. Because of this, I term it the correlation technique. Among all donor selection techniques, the correlation technique appears to be most straightforward, since observed climatic variations of low flows are directly used for donor selection. However, one drawback of the method is that the estimation of correlation coefficients requires a sufficient number of years of concurrent observations at the subject site and possible donors. Hence, the application of this method is restricted to a minimum of 5 years of overlapping data (Robson, 1999). Correlations are estimated by the Spearman's rank correlation coefficient as a sample of only 5 values is still very small for a parametric estimation of correlations. The selection procedure uses the following quantities:

- The weight w of a possible donor which takes into account the distance d in kilometres between the subject site and donor, the length of the overlap period n_o in years between subject and donor sites and the additional years of data available for the donor (n_d - n_o, where n_d is the length of the donor site record):

$$w = \left(1 - \frac{d}{120}\right) n_o (n_d - n_o) \tag{1}$$

- The similarity of climatic variation of low flows at the subject and donor sites is assessed by the Spearman's rank correlation coefficient r between annual low flows Q95(yr.) at the subject and donor sites.
- The value v of a possible donor is based on the weight w and the Spearman's rank correlation r simply as:

$$v = wr \tag{2}$$

- The 95% lower confidence limit r_l of the correlation coefficient r is calculated as:

$$r_l = \frac{e^{2z - \frac{2}{\sqrt{n_o - 3}}} - 1}{e^{2z - \frac{2}{\sqrt{n_o - 3}}} + 1} \quad \text{where} \quad z = 0.5 \cdot \ln\left(\frac{1 + r_{\max}}{1 - r_{\max}}\right) \tag{3}$$

The procedure consists of the following steps:

(a) Select all gauges within a distance of 60 km from the subject site as possible donors that have longer records than the subject site and overlap with the subject site record;

(b) Calculate weight w, correlation coefficient r and the value v of each possible donor;

(c) Limit pool of possible donors by the following criteria:

i) $r > 0$ (positive correlation),

ii) $v \geq v_{max} / 2$ (where v_{max} is the maximum donor value amongst the candidate sites),

iii) a maximum of 30 donors (otherwise drop donors with lowest values v),

(d) Determine highest correlation r_{max} amongst all the candidate sites;

(e) Calculate the 95% lower confidence limit r_l of r_{max};

(f) Remove all sites that have correlations smaller than r_l;

(g) Classify the remaining sites according to the correlation significance level (*p-value*) using the following classes: (1) $p \leq 0.01$, (2) $0.01 > p \geq 0.05$, (3) $0.05 > p \geq 0.1$, (4) $0.1 > p \geq 0.2$, (5) *any positive correlation*

(h) Final selection of donors: Select either all sites significant at the same, highest possible level or single sites that are clearly better correlated than all other sites. Starting with the highest level, the level of significance is gradually reduced until either there are at least three donor sites significant at the selected level, or there is at least one site that is significant two levels above.

4.3.3 Record augmentation

Once a suitable donor or suitable donors have been identified, the second step of climate adjustment consists of calculating adjusted values of flow characteristics for the subject site by using information from the donor or donors. Two methods are examined here. The first method adjusts the low flow characteristic at the subject site by scaling it by the ratio of Q95 calculated from the entire observations period and Q95 calculated from the overlap period (e.g. Kroiß et al., 1996)

$$QS_{pred} = QS_o \left(\frac{QD}{QD_o} \right) \tag{4}$$

where QS_{pred} is the adjusted value of Q95 at the subject site, QS_o is Q95 at the subject site calculated from the overlap period, QD_o is Q95 at the donor site calculated from the overlap period and QD is Q95 at the donor site calculated from the entire observation period. In this study there is no need to introduce a minimum overlap period as, for all subject site – donor combinations, the overlap period is identical with the record length of the subject site. I term this method the unweighted record augmentation method.

The second method uses the same principle, but includes a weighting coefficient to account for the strength of correlation between subject site and donors. A large adjustment is made for subject site – donor combinations that are highly correlated and no adjustment is made for combinations that are uncorrelated (Robson, 1999). The formula of Robson (1999) for the case of a complete overlapping of subject site record and donor-site record is used:

$$QS_{pred} = QS_o \left(\frac{QD}{QD_o} \right)^{M(r)} \tag{5}$$

which is similar to the augmentation method proposed by Vogel and Stedinger (1985). The difference is that Vogel and Stedinger (1985) used $M(r)$ as a multiplicative factor while

Robson (1999) used it as an exponent as is Eq. 5. The weighting coefficient $M(r)$ is estimated by:

$$M(r) = \frac{(n_o - 3)r^3}{(n_o - 4)r^2 + 1} \tag{6}$$

$M(r)$ takes into account the degree of correlation of annual low flows as well as the length of the overlap period of the records. I term this method the weighted record augmentation method. The limitation of this method is that, for short overlap periods, the correlation coefficients can not be estimated very reliably.

4.3.4 Combining adjusted values from multiple donors

In case of the correlation technique, more than one donor is selected, so the adjusted values for each of the donors need to be combined into a single adjusted value. The adjusted values can be combined by a weighted arithmetic average but Robson (1999) recommended a weighted geometric average which appears to be more robust to the presence of outliers in the adjusted values than an arithmetic average. The weights w are calculated from the distance between subject site and donor, the length of the overlap period and the additional years of data provided by the donor by using Eq. 1. The weighting formula then is:

$$QS_{pred} = \prod_{i=1}^{n} \left(QS_{pred}^{(i)}\right)^{\frac{w_i}{\Sigma w_i}} \tag{7}$$

where w_i is a weight for the i^{th} donor and $QS_{pred}^{(i)}$ is Q95 at the subject site adjusted by the i^{th} donor.

4.4 Evaluation method

4.4.1 Variation of record length

For each technique, the value of different record lengths is assessed by using hypothetically shortened records. This emulates the case of only short records being available at the subject site. However, in this study I have the full record length for all subject sites, so I can compare the adjusted low flow characteristic $Q95_{pred}$ for hypothetically shortened records with the low flow characteristic $Q95_{obs}$ estimated from the complete records, which gives me a measure of the estimation error introduced by a record length that is shorter than the full period. To obtain shortened records of 15, 10, 5, 3 and 1 years of observation I sub-sampled the full observation period of 20 years. All shortened records were continuous, i.e. no gaps were allowed. The beginning of the shortened records was chosen at random to make the assessment of the techniques independent of the climatic variations during the 20 years standard period.

Two additional cases were considered, spot gaugings and the case of no local data which are the limiting cases as the record length approaches zero. Spot gaugings for determining some low flow characteristic are most efficient if taken during a low flow period or, more specifically, when the discharge measured at a close-by gauge at the same stream equals the characteristic low flow discharge. In a practical study, a hydrologist could monitor daily discharges of a stream gauge near the subject site, and once the discharge is close to Q95 he/she could go out into the field and measure the discharge at the subject site on the next day. I represent this setup in this study by choosing the daily discharge Q(S) from the stream flow time series of the subject site on the day after the occurrence of a discharge value close to Q95 at the nearest downstream gauge. The daily discharge Q(S) is then interpreted as a single measurement at the subject site.

For the spot gaugings, the same donor selection procedures were used as for the shortened records, whenever possible. The methods are downstream site and catchment similarity. The no donor option is not possible to apply as the spot gauging method needs an index stream gauge to identify the appropriate day to make the measurements. Similarly, it is not possible to calculate an annual correlation coefficient, so the correlation technique could not be used in the case of spot gaugings. By the same token, only the unweighted record augmentation method (Eq. 4) could be used. For the case of no stream flow data available at the subject site, only regional information can be used to estimate low flow characteristics. Two out of the four donor selection techniques transform into simple regionalisation methods as the record length approaches zero (i.e. no local data): The downstream site method corresponds to a regional transposition of specific discharges from the downstream gauge to the subject site, and the catchment similarity method corresponds to the regional transposition of specific discharges from the site that is physiographically most similar to the subject site. In both cases the assumption is that the specific low flow discharge at the subject site is the same as at the donor site. This is a method sometimes termed the drainage area ratio method (e.g. Stedinger et al., 1992). The errors of this simple regionalisation technique will be compared to errors of the various climate adjustment techniques for varying record lengths to assess the value of short stream flow records relative to regionalisation for estimating low flow characteristics.

4.4.2 Statistical performance measures

To assess the performance of the various techniques, several statistical measures are calculated from the differences between adjusted low flow characteristics ($Q95_{pred}$) estimated from hypothetically shortened records and low flow characteristics ($Q95_{obs}$) estimated from the entire observation period of 20 years. Scatterplots of $Q95_{pred}$ vs. $Q95_{obs}$ are used for a visual assessment of the techniques and the role of record length. To facilitate the comparison, scatterplots for different techniques are grouped together for a given record length. The absolute errors for each technique and record length are assessed by the root mean squared error (*RMSE*):

$$MSE = \frac{1}{n}\sum(Q95_{pred} - Q95_{obs})^2 \tag{8}$$

$$RMSE = \sqrt{MSE} \tag{9}$$

where n is the number of subject sites. Absolute errors are calculated both for low flow discharges $Q95_{pred}$ (m^3/s) and for specific low flow discharges $q95_{pred} = Q95_{pred} / A$ ($ls^{-1}km^{-2}$) where A is the catchment area. The error of specific discharges gives more weight to smaller catchments. Note that the catchment areas of the subject sites range from 8.7 to 479 km². The mean squared error *MSE* generally constitutes an unbiased estimate of the expected error of one technique, except for the case that outliers (single sites that deviate from the bulk of the sites) are present. If one removes outliers manually, one obtains error estimates that are representative of the bulk of the data but this involves a subjective element. To obtain an objective and robust estimate of mean squared errors, I use the 5% trimmed *RMSE* instead. This means that 5% of the catchments (in my case six catchments) are disregarded in estimating *RMSE*. These are the catchments that exhibit the largest magnitudes of the differences $Q95_{pred} - Q95_{obs}$. In an exploratory analysis I compared all results in this section obtained from trimmed error statistics with untrimmed error statistics and the results only changed slightly but were less robust as indicated by somewhat more erratic error patterns.

The relative errors are estimated by dividing the absolute errors of $Q95_{pred}$ by the long term values $Q95_{obs}$. Since errors are expected to depend on the magnitude of low flow discharge, relative errors (re_C) are calculated for different classes of $Q95_{obs}$:

$$re_C = RMSE_C / m_C(Q95_{obs}) \tag{10}$$

where m_c is the class mean. The class limits have been set to the quartiles of $Q95_{obs}$ to give the same number of catchments in each class. The class limits and class means consistent with the quartiles are given in Table 3. re_c, again, is a 5% trimmed statistic.

Table 4.3. Class limits and class means for estimating relative errors (m³/s).

Class limits of $Q95_{pred}$	0.0 - 0.2	0.2 - 0.4	0.4 - 0.9	0.9 - 4.0
Class means $m_C(Q95_{obs})$	0.10	0.30	0.65	1.70

For ease of comparison with other low flow studies I also estimated the coefficient of determination R^2. Preliminary analysis indicated that R^2 of Q95 discharges are close to 100% for all techniques and all record lengths. I therefore only evaluated R^2 of q95 specific low flow discharges:

$$R^2 = \frac{s^2(q95_{obs}) - MSE(q95_{pred})}{s^2(q95_{obs})} \tag{11}$$

where s^2 is the variance of specific low flow discharges $q95_{obs}$ at all subject sites using the full record length and MSE is the mean squared error.

Following Vogel and Kroll (1991), I finally estimated the effective record length which is defined as the length of an unadjusted record that gives the same estimation error as a shorter record that is adjusted. From this I estimated the gain in information by

$$gain = \left(\frac{n_{eff}}{n} - 1\right) \cdot 100 = \left(\left(\frac{RMSE_{no\_donor}}{RMSE_{adjusted}}\right)^2 - 1\right) \cdot 100 \tag{12}$$

where n_{eff} is the effective record length and n is the record length of the subject site. Eq. 12 is based on equation 8 of Vogel and Kroll (1991) and assumes that the bias is small. A preliminary analysis of the data showed that the biases were indeed small as compared to *RMSE*. The gain provides an intuitive measure of the value of various climate adjustment procedures. If, say, an adjusted record of 10 years gives the same estimation error as an unadjusted record of 15 years the gain is 50 % in terms of the effective record length.

4.5 Results

4.5.1 Errors of unadjusted low flow characteristics

As a starting point I examined the errors of Q95 estimates from short stream flow records without applying any climate adjustment. This is the benchmark case against which the climate adjustment techniques are to be tested. All climate adjustment techniques should improve on this benchmark case. Fig. 4.1 shows the relative errors (Eq. 10) of Q95 for this case as a function of low flow discharge $Q95_{obs}$. For all record lengths there is a trend of relative errors to decrease with the Q95 discharge. Quite clearly, the errors also decrease with increasing record length from 1 to 15 years as would be expected. For the catchments with Q95 discharges larger than the median (two classes on the right hand side of Fig. 4.1), the errors decrease from 30% to 16, 12, 5 and 3% as one moves from 1 year to 3, 5, 10 and 15 years. For a record length of 20 years the error would be zero as this is the standard period the shortened records are compared with to estimate the errors. These changes of the errors with record length are a reflection of the effect of climatic variability on the low flow estimates.

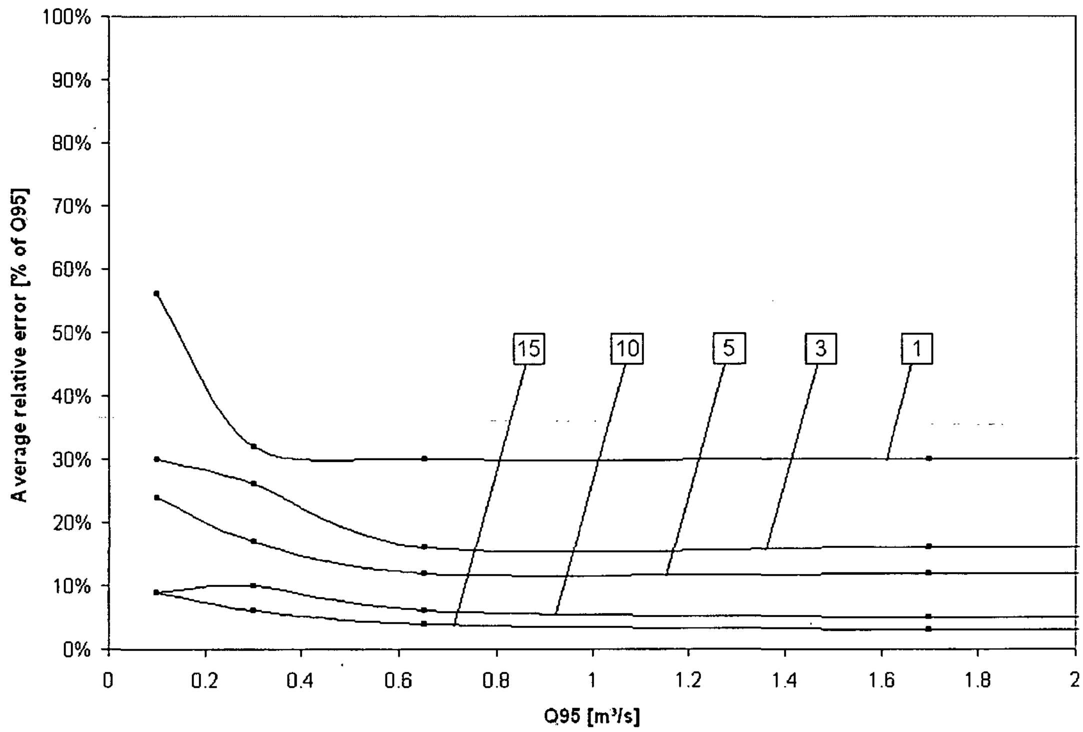

Fig. 4.1. Relative errors re_c of low flow discharge $Q95_{pred}$ estimated from records of less than 20 years as compared to 20 year records, plotted versus the Q95 low flow discharge. No climate adjustment. Numbers in boxes are the record lengths in years.

4.5.2 Relative performance of donor selection techniques

Three donor selection techniques were applied to records of variable lengths and the estimation errors were analysed by comparison with the full 20 year period. For less than 5 years the unweighted record augmentation method (Eq. 4) was used while for 5 years and more the weighted record augmentation method (Eq. 5) was used.

Three error measures are shown. Fig. 4.2 gives the absolute errors (*RMSE*) for discharges Q95 (m^3/s), Fig. 4.3 gives the absolute errors (*RMSE*) for specific discharge q95 ($ls^{-1}km^{-2}$) and Fig. 4.4 gives the coefficients of determination (R^2) for specific discharges q95. Each line represents one of the climate adjustment techniques. The line labelled "no donor" relates to the errors of unadjusted low flows as per Fig. 4.1. The minimum record length that can be used for the correlation method is 5 years. The downstream site technique and the catchment similarity technique can be used both for the case of a spot gauging (labelled S on the horizontal axis of Fig. 4.2) and the case of no local stream flow data where Q95 is estimated from the donors alone (labelled 0 on the horizontal axis of Fig. 4.2).

Figs. 2, 3, and 4 show similar results in terms of the relative performance of the methods although the magnitudes of the errors are different. The difference between the climate adjustment techniques is somewhat smaller in the case of specific low flows (Figs. 3 and 4) than for low flow discharges (Fig. 4.2). This is the result of a relatively better performance of large catchments in the downstream site method. The large catchments get more weight in RMSE calculated from Q95 than in RMSE calculated from q95. All three figures suggest that the downstream catchment method performs best. This is the case for all record lengths including spot gaugings and no data. The absolute errors of discharges decrease from 0.24 m^3/s to 0.19, 0.10, 0.08, 0.08, 0.03, 0.02 m^3/s as one moves from no data to spot gaugings, 1, 3, 5, 10 and 15 years. The absolute errors of specific discharges decrease from 2.3 $ls^{-1}km^{-2}$ to 2.1, 1.1, 0.9, 0.7, 0.4 and 0.3 $ls^{-1}km^{-2}$ and the coefficients of determination of specific discharges increase from 56% to 62, 89, 93, 96, 99 and 99%. The catchment similarity method, where the donors are physiographically similar catchments, performs second best. For no data, spot gaugings, 1 and 3 years of record there is a significant difference between the catchment similarity method and the downstream site method for all error measures. For record lengths of 5 years and more the two methods are more similar although, for the absolute errors of Q95 (Fig. 4.2), the downstream method still performs clearly better. The correlation method performs similar to the other methods in terms of the error measures based on specific low flows (Fig. 4.3 and 4.4) and it is slightly poorer for the error measure based on low flow discharges (Fig. 4.2). As compared to the benchmark case of no climate adjustment (no donor) the downstream site and the catchment similarity methods perform clearly better for record lengths of less than 5 years. For a 1 year record length the absolute errors of the down stream site method and the no donor case are 0.10 and 0.22 m^3/s, respectively, 1.1 and 2.1 $ls^{-1}km^{-2}$, respectively, and the coefficients of determinations of q95 are 89% and 63%, respectively. However, for 5 years and more the merits of using climate adjustments are relatively slim. In terms of the absolute errors of Q95, the downstream method does seem to improve the estimates while the other two methods don't. In terms of the absolute errors and the coefficient of determination, all methods exhibit some very minor improvement with the downstream method performing somewhat better than the others. It appears that climate adjustments are particularly useful for stream flow records shorter than five years but for longer records the gain of using these adjustment techniques is relatively modest.

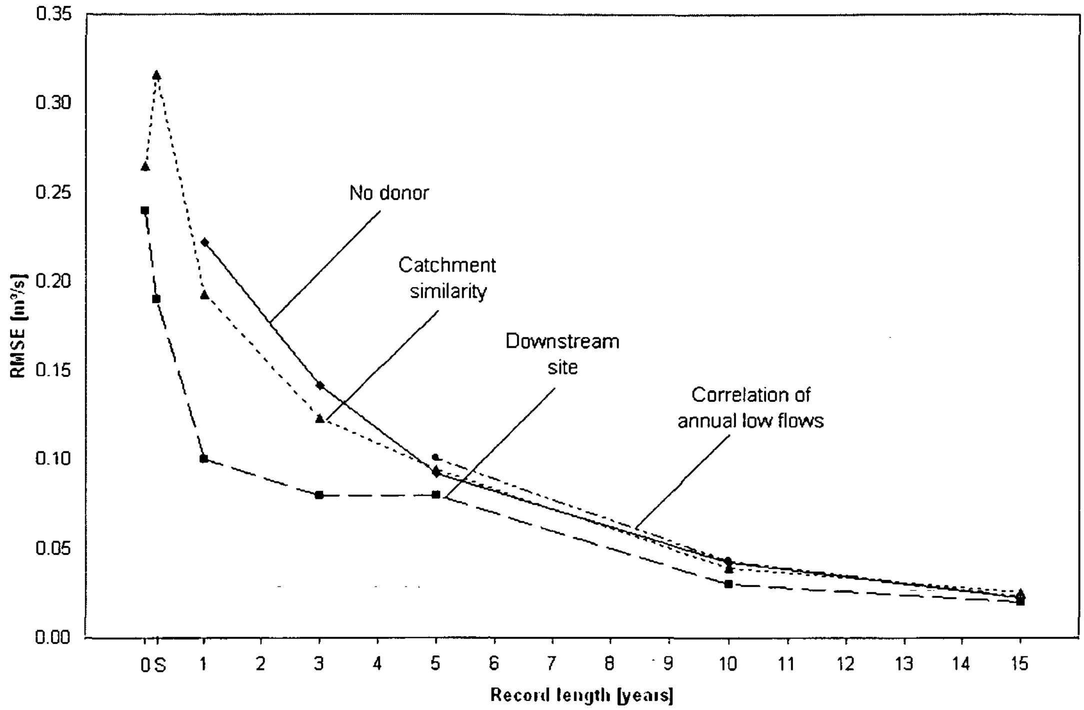

Fig. 4.2. Absolute errors RMSE (m^3/s) of low flow discharge $Q95_{pred}$ estimated from records of less than 20 years as compared to 20 year records. Various climate adjustment techniques are used. 0 = no local stream flow data, S = spot gaugings.

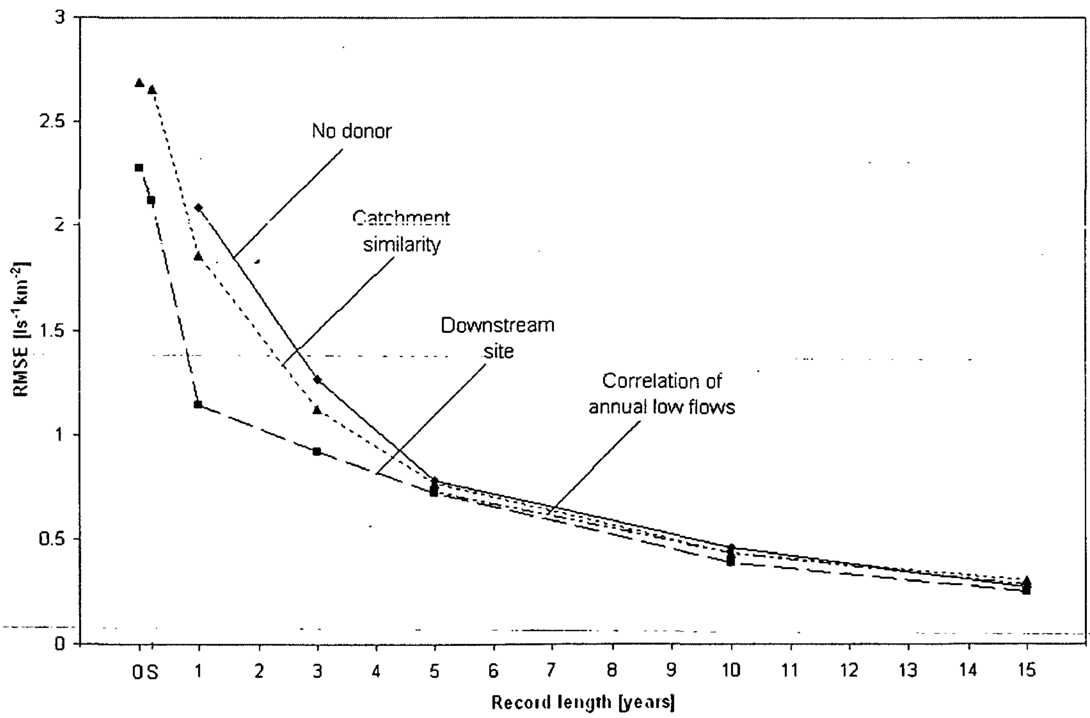

Fig. 4.3. Absolute errors RMSE ($ls^{-1}km^{-2}$) of specific low flow discharge $q95_{pred}$ estimated from records of less than 20 years as compared to 20 year records. Various climate adjustment techniques are used. 0 = no local stream flow data, S = spot gaugings.

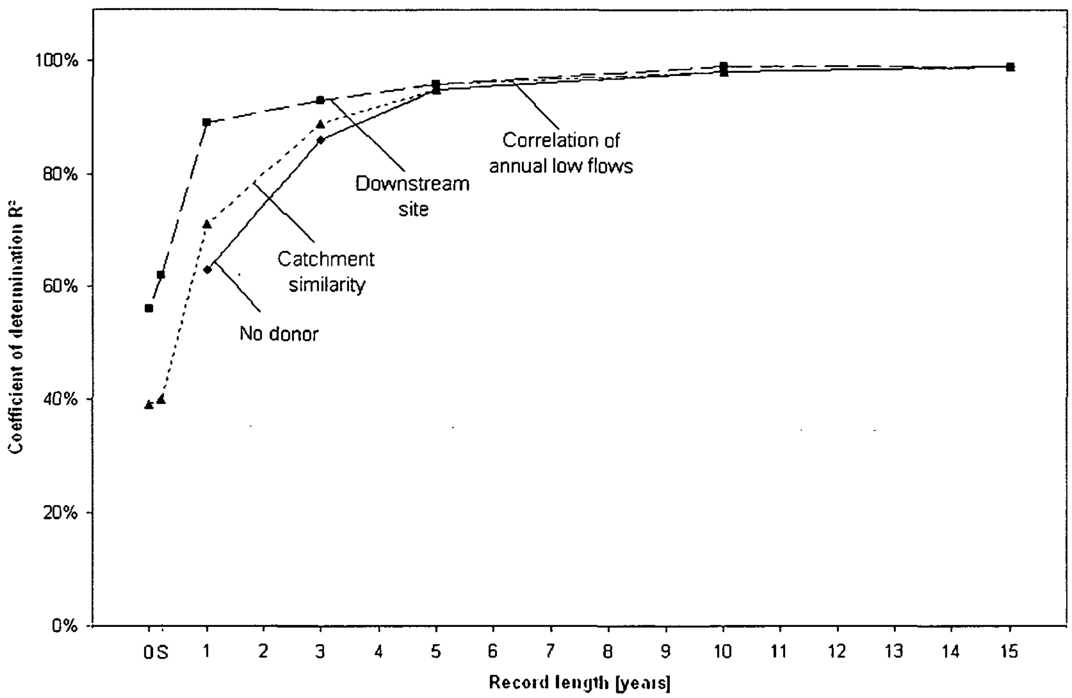

Fig. 4.4. Coefficient of determination R^2 (%) of specific low flow discharge $q95_{pred}$ estimated from records of less than 20 years as compared to 20 year records. Various climate adjustment techniques are used. 0 = no local stream flow data, S = spot gaugings.

4.5.3 Relative performance of record augmentation techniques

I now compare the performance of the two record augmentation techniques for each of the donor selection methods. The first method (Eq. 4) is an unweighted scaling of the Q95 at the subject site using the low flows from the donor while the second method (Eq. 5) is a weighted scaling where the weights are related to the correlation between the annual low flows at the subject and donor sites. The results are shown in Fig. 4.5.

For the downstream site method the two record augmentation techniques give very similar results. For the correlation method there is a slight improvement when using the weighted augmentation procedure and for the catchment similarity method there is a significant improvement. This is interesting as the weighting moves the catchment similarity method from the poorest rank to an above average rank. It appears that the value of record augmentation significantly depends on the donor selection procedure. It should be noted, however, that the choice of the donor selection method is the more important part in using climate adjustment procedures given that the differences between the donor selection methods are larger than the differences between the record augmentation methods.

It is also of interest to examine the relative gain in effective record length for each donor selection – record augmentation combination as per Eq. 12. Table 4 shows the gain (%) in effective record length based on estimated low flow discharges $Q95_{pred}$ and Table 5 shows the corresponding values for specific low flow discharges $q95_{pred}$. This comparison clearly highlights that the downstream site method yields the largest gain of all combinations both when examining discharges and specific discharges. When expressed in terms of effective record length for q95, the gain is 236% for the one year record and 91% for the three year record. For five years, the gain is either 17 or 40%, depending on the augmentation method, which means that the adjusted 5 year record is equivalent to an unadjusted 5.9 year or 7 year record. The downstream method gains 53% for a ten year record, as compared to the 20 year reference period if measured in terms of Q95 discharge, and 40% if measured in terms of q95 specific discharge. The downstream method gains 0% for a fifteen year record, as compared to the 20 year reference period if measured in terms of Q95 discharge, and 24% if measured in terms of q95 specific discharge. For 5 years and more, some of the other methods yield negative gains when using the unweighted augmentation method. This means that the estimation errors are larger than those of the unadjusted estimates. The weighting significantly reduces the occurrence of negative gains. This would be expected as poorly correlated donors get less weight than well correlated donors. Clearly, donors need to be selected with much care if they are to improve low flow estimates at the subject site.

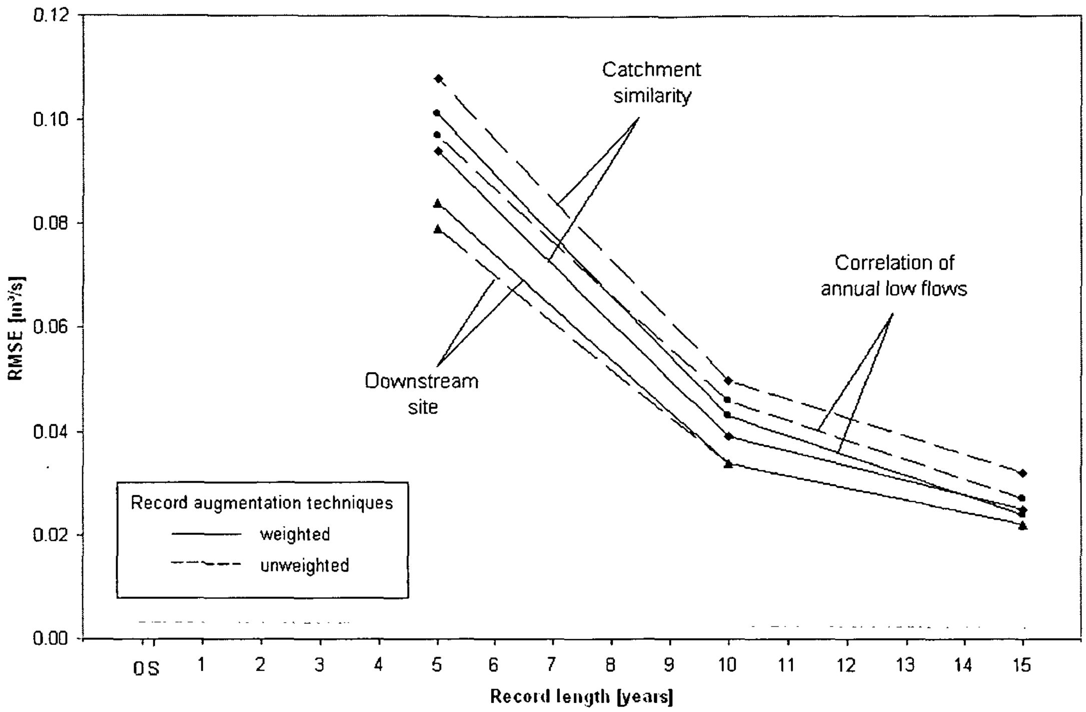

Fig. 4.5. Absolute errors RMSE (m^3/s) of low flow discharge $Q95_{pred}$ estimated from records of less than 20 years as compared to 20 year records. Three donor selection techniques are combined with two record augmentation methods (weighted: Eq. 5; unweighted: Eq.4).

Table 4.4. Gain (%) in effective record length by various climate adjustment methods based on estimated low flow discharges $Q95_{pred}$. See Eq. 12. Negative gains imply that the climate adjustment procedure is poorer than the case without adjustment. See Eq. 12. (w=weighted: Eq. 5; uw=unweighted: Eq.4).

Record length (yrs)	1	3	5	10	15
Downstream site (w)			20%	53%	0%
Similarity (w)			-4%	16%	-23%
Correlation (w)			-17%	-5%	-16%
Downstream site (uw)	403%	200%	36%	53%	0%
Similarity (uw)	32%	33%	-27%	-29%	-53%
Correlation (uw)			-10%	-17%	-34%

Table 4.5. Gain (%) in effective record length by various climate adjustment methods based on estimated specific low flow discharges $q95_{pred}$. See Eq. 12. Negative gains imply *that the climate adjustment procedure is poorer than the case without* adjustment. (w=weighted: Eq. 5; uw=unweighted: Eq.4).

Record length (yrs)	1	3	5	10	15
Downstream site (w)			17%	40%	24%
Similarity (w)			1%	16%	-15%
Correlation (w)			15%	16%	-2%
Downstream site (uw)	236%	91%	40%	40%	24%
Similarity (uw)	27%	30%	-23%	-14%	-43%
Correlation (uw)			18%	-10%	-21%

4.5.4 Heteroscedasticity and outliers

The error measures examined in the previous sections are 5% trimmed error statistics, i.e. they reflect the performance of the various methods for the bulk of the catchments. However, it is also of interest to analyse outliers and the performance of individual catchments. I therefore plotted the low flow discharges estimated for various record lengths ($Q95_{pred}$) against the low flow discharges estimated for the full record length of 20 years ($Q95_{obs}$) in Figs. 6 to 10. These scatter plots also allow me to examine the estimates for heteroscedasticity, i.e. whether the variance of the differences $Q95_{pred}$ - $Q95_{obs}$ changes with the magnitude of $Q95_{obs}$. Again, for less than 5 years the unweighted record augmentation method (Eq. 4) was used while for 5 years and more the weighted record augmentation method (Eq. 5) was used.

Fig. 4.6 suggests that the 15 year estimates for all methods are very similar to the 20 year estimates. The errors are very small and there is essentially no difference between the methods discernable in Fig. 4.6. There are two or three catchments in all methods that are not exactly on the 1:1 line most of which are the same catchments in all methods. Scatter plots for the five year records (Fig. 4.7) still indicate very high correlations for all techniques, although there is some decrease in the performance relative to 15 years as one would expect. Again, all methods are rather similar although the correlation technique produces slightly more outliers than the other methods, particularly for the large low flow discharges. For one year of observation (Fig. 4.8), only three techniques remain to be compared. Both climate adjustment techniques (downstream site and catchment similarity) improve the accuracy of low flow estimates over the case without adjustment (no donor). For the downstream method, the increase in performance is very significant while for the catchment similarity method it is not. There appear to exist two groups of catchments, catchments with Q95 of less than 1.5 m^3/s and those with Q95 of more than 1.5 m^3/s. For the former group the catchment similarity method gives almost the same scatter as the no donor case, so there is no improvement, while the downstream site method gives significantly less scatter. For the latter group, the catchment similarity method does seem to slightly decrease some of the scatter over the no donor benchmarking case but the downstream method is clearly better.

For the case of using spot gaugings for estimating low flows (Fig. 4.9) there are again two groups of catchments. In the lower discharge group the scatter is relatively small, particularly for the downstream site method although there are a few outliers. The scatter of this group is similar to that of the one year case in Fig. 4.8, with the exception of the outliers. For the upper discharge group the scatter is larger and, again, the downstream site method performs better than the similarity method.

In the final case of no local information, i.e. regionalisation of Q95 (Fig. 4.10), the scatter of the low discharge group increases significantly, particularly for the downstream site method. For the upper discharge group, there is a slight increase in the scatter. It is interesting that the catchment similarity method tends to underestimate low flows in the upper discharge group for the no data case while there was consistent overestimation for the spot gauging case. This explains the larger RMSE in Fig. 4.2 for the spot gauging case than for the no data case. From a visual comparison of Figs. 10 and 9 it appears that the spot gauging does improve the performance of both methods over the no data case. This is not fully borne out by the error statistics in Figs. 2, 3 and 4 that only showed a slight improvement. It is therefore interesting to examine what is the reason of the lack of significant improvement by the spot gaugings which will be done in the following section.

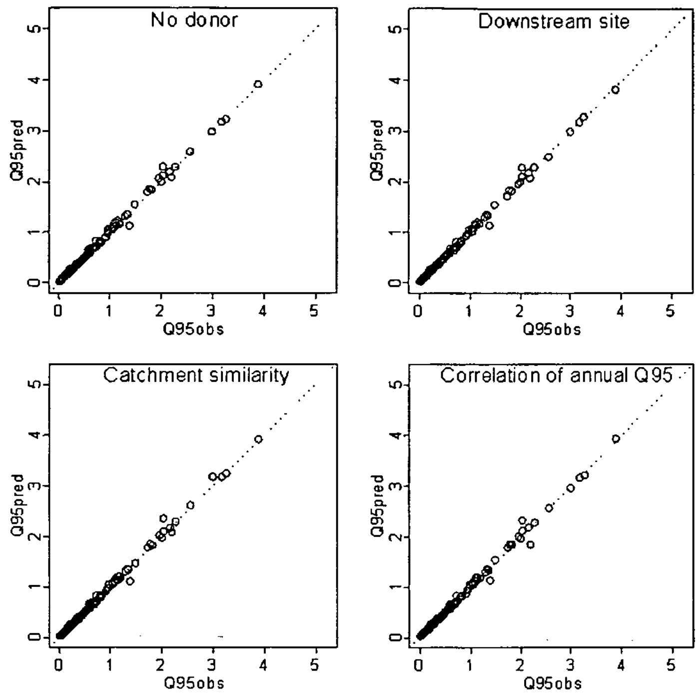

Fig. 4.6. Adjusted low flows $Q95_{pred}$ (m^3/s) estimated from 15 year records plotted versus low flows $Q95_{obs}$ (m^3/s) estimated from the full 20 year period. Each point represents a catchment and the panels relate to different donor selection methods.

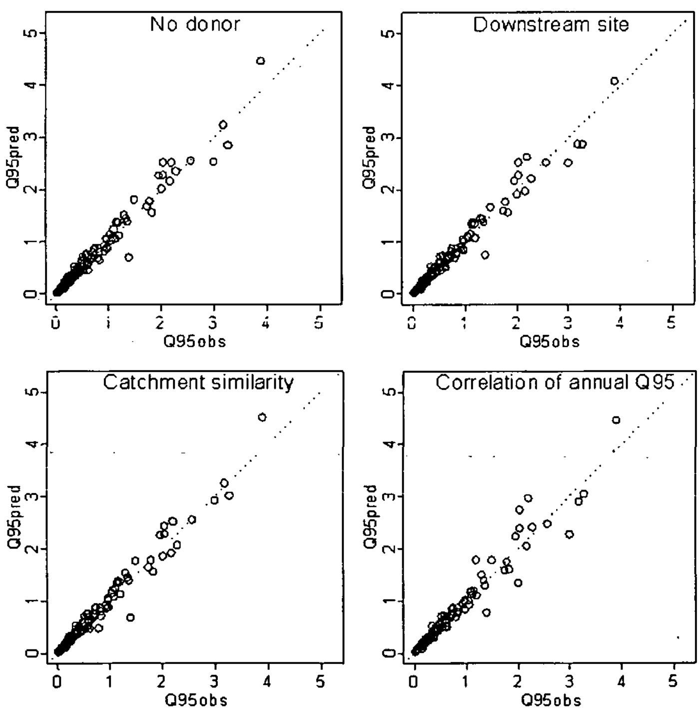

Fig. 4.7. Adjusted low flows $Q95_{pred}$ (m^3/s) estimated from 5 year records plotted versus low flows $Q95_{obs}$ (m^3/s) estimated from the full 20 year period. Each point represents a catchment and the panels relate to different donor selection methods.

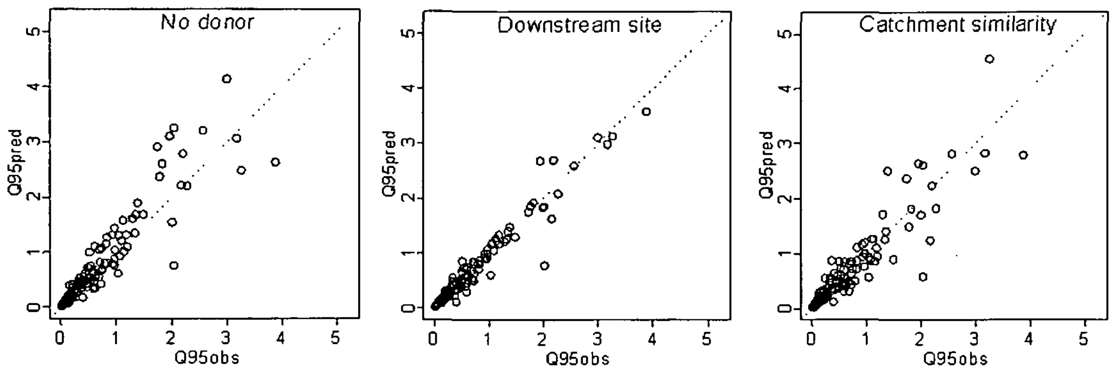

Fig. 4.8. Adjusted low flows $Q95_{pred}$ (m^3/s) estimated from 1 year records plotted versus low flows $Q95_{obs}$ (m^3/s) estimated from the full 20 year period. Each point represents a catchment and the panels relate to different donor selection methods.

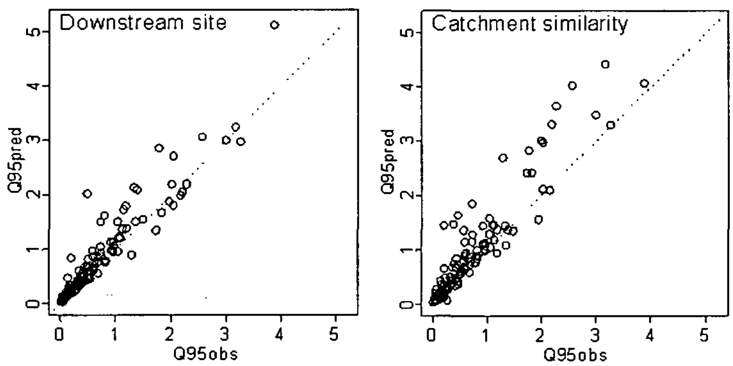

Fig. 4.9. Low flows $Q95_{pred}$ (m^3/s) estimated from spot gaugings plotted versus low flows $Q95_{obs}$ (m^3/s) estimated from the full 20 year period. Each point represents a catchment and the panels relate to different donor selection methods.

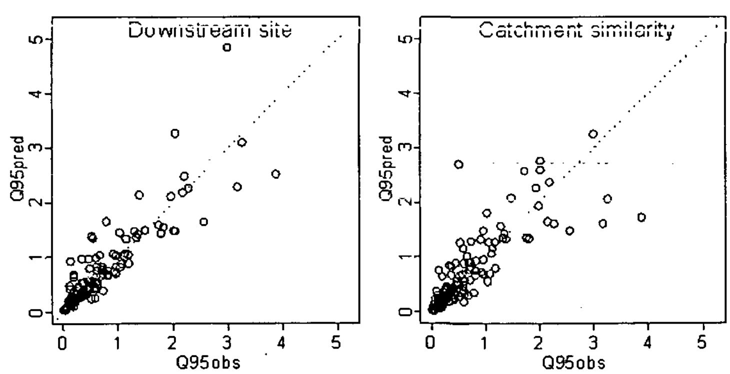

Fig. 4.10. Low flows $Q95_{pred}$ (m^3/s) estimated from a simple regionalisation model (i.e. no local stream flow data) plotted versus low flows $Q95_{obs}$ (m^3/s) estimated from the full 20 year period. Each point represents a catchment and the panels relate to different donor selection methods.

4.5.5 Spot gaugings

To analyse the error sources of the spot gauging and no donor (regionalisation) cases I calculated ratios of specific discharges at the subject and donor sites. In this comparison q95(S) is the specific low flow discharge exceeded 95 of all days at the subject site estimated from the 20 year record at the subject site. q95(D) is the analogous value for the donor site, and q(S) is the specific discharge "measured" by the spot gauging at the subject site.

The ratio q95(S)/q95(D) is a measure of the spatial variability of low flows in the region. A unit ratio represents spatially uniform low flows and values lower or larger than one reflect spatial variability. The no data (regionalisation) case is consistent with assuming q95(S)/q95(D) = 1, values much larger or smaller than one indicate large errors in the simple regionalisation procedure. The ratio q(S)/q95(S) is a measure of how well the spot gauging captures the q95 at the subject site. A unit ratio indicates that the spot gauging perfectly captures the q95 at the subject site and values lower or larger than one indicate that the spot gauging was not performed on a suitable day. The ratio q(S)/q95(D) can be thought of as the climate adjustment in the case of the spot gaugings, i.e. it reflects how different the spot gaugings are from the q95 at the donor site. This ratio can also be thought of as a reflection of the combined sources of variability or uncertainty, spatial variability (expressed as q95(S)/q95(D)) and unsuitable timing of the spot gaugings (expressed as q(S)/q95(S)).

Fig. 4.11 shows the cumulative distribution functions of these three ratios for both donor selection methods. The slope of the cumulative distribution functions at a ratio of one is an indication of the spread of the distribution and hence a measure of uncertainty. Fig. 4.11 top (downstream site method) indicates that the uncertainty introduced by the spatial variability (dashed line) is about the same as the uncertainty introduced by the timing of the spot gaugings (dashed dotted line). The combined effect of the two (solid line) shows a still larger spread and hence larger uncertainty. The interesting thing in this figure is that the additional information gained by a spot gauging is small as it tends not to be very representative of the Q95 low flow. Because of this, the spot gauging method does not improve the Q95 estimate much over the case of no data (regionalisation). On closer examination, the q(S)/q95(S) distribution shows a slightly smaller spread or random variability as indicated by the slope of the cumulative distribution function around the mean but it shows a significant bias as indicated by the location of the cumulative distribution function. The procedure emulated here of taking base flow measurements the day after the discharge at a nearby gauged site is close to q95 is clearly a biased procedure.

It is also interesting to compare the catchment similarity method (Fig. 4.11 bottom) with the downstream site method of donor selection (Fig. 4.11 top). The catchment similarity method is associated with a wider spread in the q95(S)/q95(D) distribution (dashed dotted line) indicating the donors are less similar than for the down stream case. There is also a larger spread in the q(S)/q95(S) distribution indicating that the spot gaugings are less representative of q95 as the timing of the gaugings is not picked well. The combined effect of the two (solid line) shows a still larger spread, pointing to the larger uncertainties of the catchment similarity method than the downstream site method.

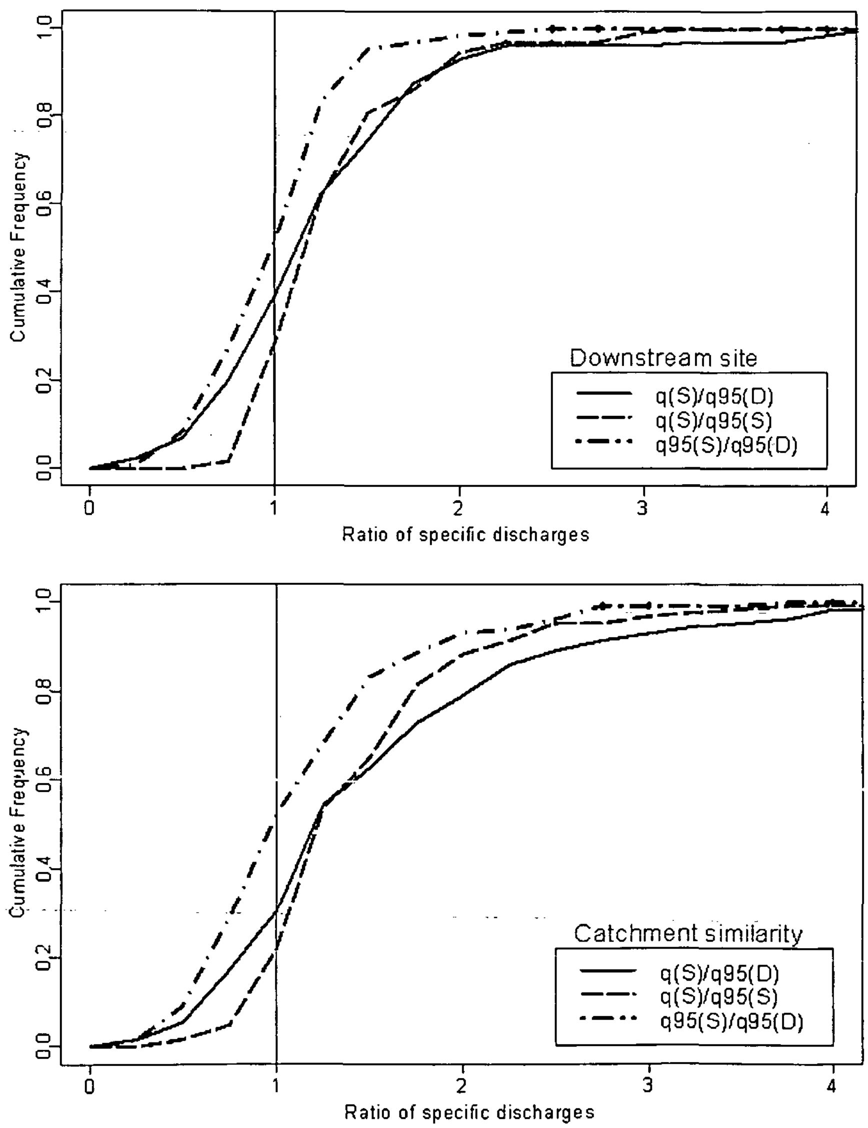

Fig. 4.11. Cumulative frequency distribution of specific discharge ratios. q95(S) is the specific low flow discharge exceeded on 95% of all days at the subject site, q95(D) is the analogous value for the donor site, and q(S) is the specific discharge measured by spot gauging at the subject site.

4.5.6 Effect of discharges

A final assessment in this section (Fig. 4.12) examines the performance of the best method, i.e., downstream site donor selection, as a function of the magnitudes of the Q95 discharges at the subject site. For all record lengths, there is a trend of relative errors to decrease with the Q95 discharge. Quite clearly, the errors decrease with increasing record length from no data to spot gaugings, 1, 3, 5, 10 and 15 years as would be expected. For the largest Q95 class the errors decrease from 28 to 22, 14, 10, 13, 4 and 4%. For the lowest Q95 class the errors decrease from 98 to 64, 25, 20, 22, 12 and 8%. The five year curve slightly crosses over some of the other curves which likely is an random artefact of the data and not a significant pattern.

Fig. 4.12 is a similar representation as Fig. 4.1 but the difference is that Fig. 4.1 is without climate adjustment while Fig. 4.12 is with climate adjustment based on the nearest downstream site. The degree to which the errors in Fig. 4.12 are smaller than those in Fig. 4.1 is a measure of the value of the climate adjustment procedure as a function of low flow discharge. The error pattern in Fig. 4.12 is similar to that in Fig. 4.1 but all errors are significantly smaller indicating that this climate adjustment method significantly enhances the accuracy of the Q95 estimates for short stream flow records.

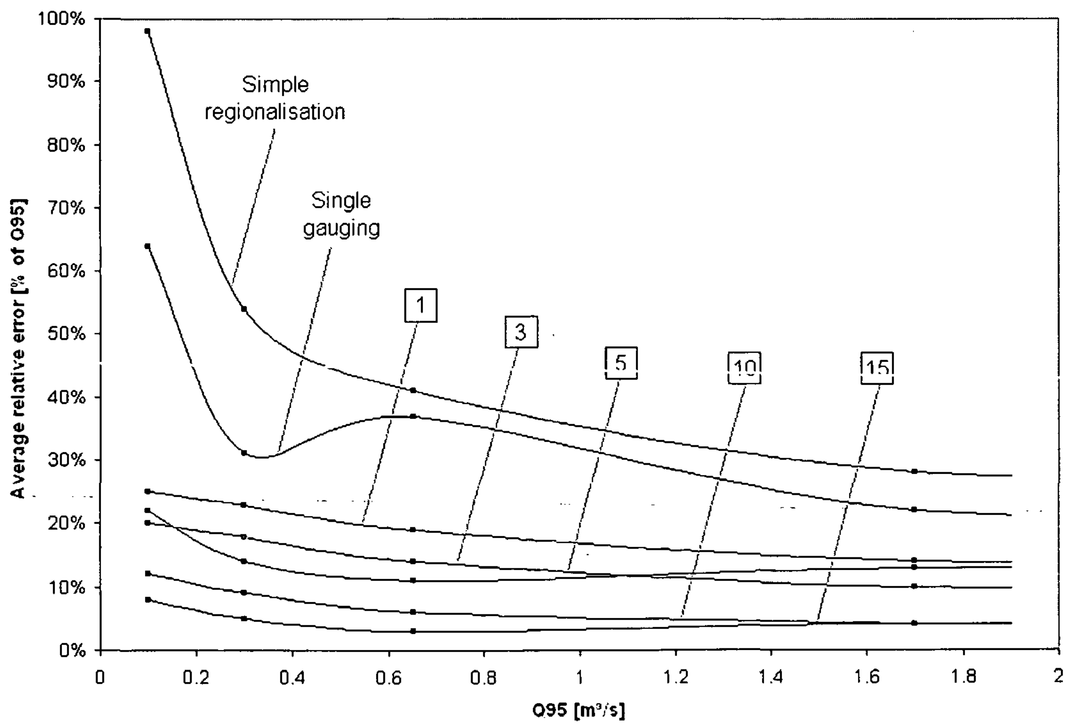

Fig. 4.12. Relative errors re_c of low flow discharge $Q95_{pred}$ estimated from records of less than 20 years as compared to 20 year records, plotted versus the Q95 low flow discharge. Climate adjustment based on the downstream site donor selection technique. Numbers in boxes are the record lengths in years.

4.6 Discussion

4.6.1 Assessment of climate adjustment methods

The comparisons have shown that the downstream donor selection method performs best on all scores. Part of the strength of using the nearest gauge at the same stream as a donor is probably related to the spatial proximity which, apparently, is associated with a significant similarity in the response to climate variation. Another, perhaps more important, reason of the good performance of this method is that the subject site catchment is a part of the donor catchment, so the subject site catchment actually drains to the donor site. If the ratio of donor and subject site catchment areas is not too large, the downstream site donor selection method is certainly the preferred choice. The catchment similarity method performs somewhat poorer. In this case, the donor selection is based on physiographic similarity and location in the same seasonality region. There are two possible reasons for the relatively poorer performance. The first may be that the physiographic catchment characteristics are not very representative of the climate impacts on low flows dynamics. The second reason may be related to the way the similarity measure was defined in this study. It is possible that a hydrologically more relevant combination of the catchment characteristics than the Euclidean distance used in this section will enhance the performance of the method. More work is needed along these lines. One possibility is the use of regression trees that have been shown to be promising in the context of regionalising low flow characteristics (section 3).

The performance of the correlation technique of donor selection, overall, is similar to the catchment similarity method. This is counterintuitive as one would expect annual correlations of low flows to be the most efficient similarity measure of climatic variation impacts both in the donor selection and record augmentation procedures. There may be a number of reasons for the somewhat lower performance than that of the downstream site method. It appears that the correlation coefficients cannot be estimated very well for short overlap periods. Indeed, the record lengths used in this section are significantly shorter than most of the record lengths examined in Vogel and Kroll (1991). It is likely that, if I compared the value of, say, 20 years of record relative to 40 years of record or more, the relative performance of the correlation method increased as the correlation coefficients can be estimated more reliably from larger samples. It should also be noted that, while the original development of the correlation method applies to both low flows and floods (Vogel and Stedinger, 1985; Vogel and Kroll, 1991), the refined version used here was specifically geared towards floods (Robson, 1999) which may also explain part of the lower than expected performance.

It is interesting that the comparison indicated that the weighted record augmentation procedure that uses information on the correlations of annual low flows significantly improved the estimates for the case of the catchment similarity donor selection procedure, slightly improved the correlation donor selection procedure, but hardly made any difference for the down stream site procedure. The combination of catchment similarity and weighted augmentation, however, does not give significantly better results than the unadjusted case. This means that, for the record lengths examined here, there is little practical value in this combination. It appears that the value of record augmentation significantly depends on the donor selection procedure. It should be noted, however, that the choice of donor selection method is the more important part of climate adjustments given that the differences between the donor selection methods found here were larger than the differences between the record augmentation methods. In a region that is hydrologically as diverse as the study area, the suitable choice of donor sites clearly is of utmost importance. Donors need to be selected with care if they are to improve low flow estimates at the subject site.

Both the unadjusted low flows (Fig. 4.1) and the low flows adjusted by the best method (downstream method, Fig. 4.12) have shown a trend of relative errors to decrease with the Q95 discharge. This may be due to a number of reasons. A first obvious reason may be measurement errors which are relatively more important for small low flow discharges. For the case of unadjusted low flows this trend may also suggest that climate variability is more important in small and/or dry catchments than it is in large catchments. For adjusted low flows this may also be the case although there is probably an additional scale effect. The large discharges tend to stem from large catchment areas. This can be seen from Table 1, as the relative range of discharges is significantly larger than the relative range of specific discharges. An examination of the distribution of the catchment areas for the case of the downstream site method (not shown here) indicates that for large subject catchments, the ratio of donor and subject catchment areas tends to be somewhat smaller than for small subject catchments. There is therefore a tendency for the downstream site method to perform better for the large catchments than for the small catchments which tends to give smaller errors for the larger Q95 classes. An additional interpretation offered here for the trend of errors to decrease with Q95 is that the regional transposition from donor to subject site may be more accurate in the wet catchments (large q95) than in dry catchments (small q95). Merz and Blöschl (2004a), for example, found significantly smaller regionalisation errors in wet catchments (large specific flood discharges) than in dry catchments (small specific flood discharges) for the case of flood frequencies.

4.6.2 Effect of record lengths

The comparisons show that the value of climate adjustment methods is very significant for record lengths shorter than 5 years. For the downstream site method, the coefficient of determination of q95 specific low flows increases from 63 to 89% for the one year record, and from 86 to 93 % for the three year record but the increase is much smaller for five years (from 95 to 96%) and still smaller for larger record lengths. When expressed in terms of effective record length for q95, the gain is 236% for the one year record and 91% for the three year record. For five years the gain is either 17 or 40%, depending on the augmentation method which means that the adjusted 5 year record is equivalent to an unadjusted 5.9 year or 7 year record.

Overall, the root mean squared errors *RMSE*, approximately, decrease with the inverse of the square root of the record length as one would expect for the mean of an uncorrelated sample. Correlations are likely present in the discharge times series but do not appear to have a significant effect. It should be noted that the *RMSE* is calculated from a regional comparison of a large number of catchments rather than from the statistical characteristics of a single site which may be part of the reason of the small effect of correlations.

It is now of interest to compare the results of the value of short records to the literature. Vogel and Kroll (1991), based on 23 catchments in Massachusetts suggested that the gain in effective record length depends on both the actual record length and the low flow quantity examined. They also noted that the serial correlations present in the discharge series may decrease the effective record length. They then examined the net effect of record augmentation and serial correlations for various low flow characteristics $Q_{d,10}$ which are the low flow values over *d* days associated with a return period of 10 years. For their set of six catchments with record lengths of about 18 years, the gain in effective record length was +7% for $Q_{1,10}$, +18% for $Q_{7,10}$, and +34% for $Q_{30,10}$. In this section, the downstream method gains 53% for a ten year record, as compared to the 20 year reference period if measured in terms of Q95 discharge, and 40% if measured in terms of q95 specific discharge (Tables 4 and 5). The downstream method gains 0% for a fifteen year record, as compared to the 20 year reference period if measured in terms of Q95 discharge, and 24% if measured in terms of q95 specific discharge. The Q95 low flow characteristic examined in this section has a similar order of magnitude as $Q_{7,10}$ for the study region (Kresser et al., 1985). The reference record length of Vogel and Kroll (1991) was longer than in this section and they had significantly fewer catchments than in this section, so the results are not strictly comparable. However, they do indicate that the order of magnitude of the gains in effective record length obtained by climate adjustments in the two studies are similar.

4.6.3 Value of short time series compared to regionalisation

In a practical application, there are two alternative ways of estimating low flow characteristics at sites without long-term observations, either from short records with or without record augmentation procedure or from regional information alone without making use of the local stream flow data. In this section, adjustment techniques for short records have been compared to simple regionalisation approaches. Overall, the results indicate that the spot gaugings slightly improve the low flow estimates over the simple regionalisation method and the one year record significantly improves the estimates over the spot gaugings. However, this comparison is based on a simple regionalisation method of assuming that the specific low flow discharge at the subject site is the same as at the donor site. More accurate low flow regionalisation methods exist.

In section 3, for example, I have compared a number of low flow regionalisation methods in the same study area. This analysis was based on discharges of catchments without upstream gauges, as in this section, but there I also included discharges of residual catchments between subsequent gauges with a total of 325 catchments rather than 132 catchments as in this section. I assessed the regionalisation errors, among other measures, by the cross validated coefficient of determination. This is similar to the coefficient of determination in this section although I did not trim the statistics in section 3, so the coefficients of determination of section 3 likely are a little lower than trimmed coefficients of determination as used in this section. The global regression model of section 3 included eight variables representing topography, precipitation and catchment geology, yielded a coefficient of determination of R^2=57% which is similar to the best simple regionalisation model of this section (R^2=56%, Fig. 4.4). The optimal regionalisation model of section 3, however, yielded R^2=70%. This model was based on separate regressions in eight seasonality zones. If I compare this to the results of this section I can see from Fig. 4.4 that one year of observations gives an R^2=89% while the spot gauging gives R^2=62% if the downstream site method is used. This means that one year of stream flow data clearly outperforms the regionalisation method while the regionalisation of section 3 performs better than the spot gauging method. It should also be noted that the performance of the regionalisation method of section 3 varied significantly between the zones, ranging from R^2=51% to R^2=89%. For the zones associated with the lower performance it is likely that even the spot gauging will improve on the regionalisation estimate while for the zones associated with the higher performance the climate adjustment based on one year of continuous data will give similar results as the regionalisation.

The spot gaugings perform somewhat better than the simple regionalisation procedure on all scores. However, the difference is not very large. The analysis in Fig. 4.11 has indicated that the uncertainty introduced by the timing of the spot gaugings is about the same as the uncertainty introduced by the spatial variability. This is the case for both the downstream site and the catchment similarity methods. This means that the additional information gained by a spot gauging is small as it tends not to be very representative of the Q95 low flow. Because of this, the spot gauging method does not improve the Q95 estimate much over the case of no data (regionalisation). Also, the procedure emulated here of taking base flow measurements the day after the discharge at a nearby gauged site is close to q95 is clearly a biased procedure. This may be related to the temporal dynamics of stream flow. Increases in discharge tend to be steeper than the recessions which may bias the spot gaugings if performed on the day after the occurrence of Q95. Based on the experience of the case study of Kroiß et al. (1996), my expectation was that the spot gauging significantly improves over the simple regionalisation but this is apparently not the case when examined on a larger data base. It should also be noted that the relative performance of the spot gauging method depends on the uncertainty introduced by the timing of the spot gaugings relative to the spatial low flow variability. The study area of Kroiß et al. (1996) was in the north of Austria where the low flows are very heterogeneous over a small region. It is likely that in a very heterogeneous region, the value of spot gaugings increases. The error statistics provided in this section are averages over 132 catchments in a large region and the relative performance in subregions may be different from the general trend. The relatively low performance is also consistent with an assessment of the use of individual measurements for estimating low flows in other climate regions which suggests that the method can be subject to considerable error when only a few discharge measurements are used (Stedinger et al., 1992). It is likely that the performance of this method increased if I extended the sampling to a number of spot gaugings during more than one low flow period.

4.7 Conclusions

The comparisons have shown that the downstream donor selection method performs best on all scores. This method yields the smallest root mean square errors, the largest coefficients of determination, and the fewest outliers if the adjusted Q95 and q95 low flow estimates from shortened records are compared to estimates from the full 20 year record. The catchment similarity and correlation donor selection methods yield larger errors on most scores. The performance of these two methods is similar.

The relative performance of the record augmentation methods depends on the donor selection method. The more sophisticated augmentation method that uses correlations of annual low flows increases the performance in the case of catchment similarity donor selection. This performance, however, is not significantly better than the unadjusted case. For the other donor selection methods the two record augmentation methods yield similar performances. Overall, the choice of donor site appears to be more important than the choice of record augmentation method.

The value of the climate adjustment methods is very significant for record lengths shorter than 5 years. For the downstream site method, the coefficient of determination of q95 specific low flows increases from 63 to 89% for one year records, and from 86 to 93% for three year records. When expressed in terms of effective record lengths of q95, the gain is 236% for the one year record and 91% for the three year record. The value of the climate adjustment methods is much smaller for records of five years and more. For five years, using the downstream site donor selection, the gain is either 17 or 40%, depending on the augmentation method, and is smaller or non-existent for other donor selection methods.

The method that uses spot gaugings of stream flow during a low flow period performs slightly better than a simple regionalisation procedure in terms of predicting Q95 at an otherwise ungauged site. The additional information gained by spot gaugings is small mainly because they are not very representative in terms of their timing. This uncertainty has a similar magnitude as the uncertainty introduced by the spatial variability of low flows.

Comparisons of the accuracy of q95 specific discharge estimates from short stream flow records in this section with more sophisticated regionalisation procedures from section 3 suggest that, on average over the study region, one year of continuous stream flow data clearly outperforms the more sophisticated regionalisation method but the spot gauging method provides less accurate low flow estimates than the sophisticated regionalisation method.

5. Overall conclusions

The objective of Section 2 was to examine the value of different seasonality indices for low flow regionalisation. In a first step, three seasonality indices were compared. Seasonality histograms are the most detailed indices, but classification techniques are needed to compare seasonality among a large number of catchments. The cyclic seasonality index is a more compact index and the spatial patterns can be delineated by visual inspection of a vector map. The seasonality ratio 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, low lands and hilly landscapes. In a second step, three catchment classification methods that are based on seasonality have been examined. Cluster analyses of seasonality histograms resulted in a first 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 regionalisation was examined by comparing three multiple regression approaches which include the seasonality classifications in different ways, to the global regression model which does not include seasonality. The overall coefficient of determination of specific low flow discharges, q95, in cross-validation mode does not change much between the seasonality approaches. Fitting separate models to three regions (summer, winter and mixed seasonality) performs best (R^2_{CV} = 60%), followed by separate models fitted to two regions (R^2_{CV} = 59%). Including different calibration coefficients in each of the eight seasonality regions resulted in R^2_{CV} = 58% and hence performs only slightly better than the global regression model (R^2_{CV} = 57%). The models for the summer regions (R^2_{CV} = 66% and 60%), however, clearly perform better than the models for the winter regions (R^2_{CV} = 51%). The model for the catchments of the mixed seasonality type (R^2_{CV} = 35%) does not nearly perform as well. The residual maps of predicted minus observes q95 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. This is one of the issues explored in Section 3.

In Section 3, the value of catchment grouping for low flow regionalisation has been assessed. Four catchment grouping methods are evaluated in terms of their performance in predicting specific low flow discharges q_{95}. These methods are the residual pattern approach, weighted cluster analysis, regression trees and an approach based on the eight seasonality regions identified in Section 2. In a first step the part of the variance (R^2) of specific low flows q_{95} is examined that can be explained by the grouping alone without using regressions. In this comparison, the regression tree approach performs best and explains 62% of the spatial variance. This means that the regression tree is an excellent classification method if one is interested in finding groups that are most distinct in terms of both catchment characteristics and low flow catchment response. In a second step the goodness-of-fit of the regressions between catchment characteristics and q_{95} for each of the groups identified by the various grouping methods is compared. Here, the seasonality regions approach provides the best goodness-of-fit of the component regression models and explains between 51 and 89% of the spatial variance of q_{95}, depending on the region. A global regression model explains 62% of the variance in q_{95}. It uses range of altitude, fraction of rock, fraction of Flysch, and average winter precipitation as the predictor variables. In a third step the predictive power of the regional regressions based on each of the grouping methods using leave-one-out cross-validation is examined. The cross-validation represents the regionalisation error of the low flows one has to expect for the case of ungauged sites. Among the grouping methods tested here, the grouping based on seasonality regions performs best and explains 70% of the variance in a cross-validation mode. The favourable performance of this grouping method is likely related to the striking differences in seasonal low flow processes in the study domain. Winter low flows are a result of the retention of solid precipitation in the seasonal snow pack of the catchments and of freezing processes in the soils while summer low flows are related to the relatively large moisture deficits in the lowland regions of Austria during summer. The regression tree grouping performs second best (explained variance of 64%) and the performance of the residual pattern approach is similar (explained variance of 63%). The weighted cluster analysis only explains 59% of the spatial variance of q_{95} which is only a minor improvement over the global regression model, i.e. without using any grouping, in a cross-validation mode (explained variance of 57%). An analysis of the sample characteristics of all methods suggests that, again, the grouping method based on the seasonality regions has the most favourable characteristics although all methods tend to underestimate specific low flow discharges in the very wet catchments.

In section 4, the value of short stream flow records in regional low flow estimation is explored. A number of methods of adjusting Q95 estimates from short stream flow records for climate variability are compared. The comparisons have shown that the downstream donor selection method performs best on all scores. This method yields the smallest root mean squared errors, the largest coefficients of determination, and the fewest outliers if the adjusted Q95 and q95 low flow estimates from shortened records are compared to estimates from the full 20 year record. The catchment similarity and correlation donor selection methods yield larger errors on most scores. The performance of these two methods is similar. The relative performance of the record augmentation methods depends on the donor selection method. The more sophisticated augmentation method that uses correlations of annual low flows increases the performance in the case of catchment similarity donor selection. This performance, however, is not significantly better than the unadjusted case. For the other donor selection methods the two record augmentation methods yield similar performances. Overall, the choice of donor site appears to be more important than the choice of record augmentation method. The value of the climate adjustment methods is very significant for record lengths shorter than 5 years. For the downstream site method, the coefficient of determination of q95 specific low flows increases from 63 to 89% for one year records, and from 86 to 93% for three year records. When expressed in terms of effective record lengths of q95, the gain is 236% for the one year record and 91% for the three year record. The value of the climate adjustment methods is much smaller for records of five years and more. For five years, using the downstream site donor selection, the gain is either 17 or 40%, depending on the augmentation method, and is smaller or non-existent for other donor selection methods. The method that uses spot gaugings of stream flow during a low flow period performs slightly better than a simple regionalisation procedure in terms of predicting Q95 at an otherwise ungauged site. The additional information gained by spot gaugings is small mainly because they are not very representative in terms of their timing. This uncertainty has a similar magnitude as the uncertainty introduced by the spatial variability of low flows. Comparisons of the accuracy of q95 specific discharge estimates from short stream flow records in this section with more sophisticated regionalisation procedures from Section 3 suggest that, on average over the study region, one year of continuous stream flow data clearly outperforms the more sophisticated regionalisation method but the spot gauging method provides less accurate low flow estimates than the sophisticated regionalisation method.

The analyses of the regional low flow processes and the comparisons of the regionalisation methods in this thesis suggest that process understanding can indeed assist in regionalising low flow characteristics to provide more accurate estimates than existing standard methods. There are a number of logical extensions of the work of this thesis. The most obvious, and perhaps most important, extension is to examine regionalisation methods for other low flow characteristics. While this thesis has examined low flows associated with a certain exceedance probability, Q95, low flow characteristics associated with a certain duration and characteristics representing the stream flow dynamics are also of interest in water resources management and engineering. It is likely that the process based methods of this thesis can be profitably used in these extensions.

6. References

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

Aschwanden H. & Kan C. (1999) Le débit d'étiage Q347 – Etat de la question. *Communications hydrologiques* 27, Service hydrol. et géol. 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. Institut für Hydraulik Gewässerkunde und Wasserwirtschaft, Technische Universität Wien, Forschungsbericht 11, Wien.

Blöschl G., U. Piock-Ellena und R. Merz (2000) Abflußtypen-Klassifizierung als Basis für die Regionalisierung von Hochwässern. Endbericht an die Österreichische Akademie der Wissenschaften (HÖ-12/97), Feb. 2000, Institut für Hydraulik, TU Wien.

Breiman L., Friedman J.H., Olshen R., Stone C.J. (1984) Classification and Regression Trees. Wadsworth International Group, Belmont, California.

Burn D.H. (1997) Regionalisation of catchments for regional flood frequency analysis. J. Hydrol. Engng 2 (2), 76-82.

Clark, L.A. and Pregibon, D. (1991) Tree-based models. In: Chambers, J.M. and Hastie, T.J. (Eds.) Statistical models in S. Chapman and Hall/CRC, New York.

Clausen, B. and Pearson, C.P. (1995) Regional frequency analysis of annual maximum streamflow drought. J. Hydrol. 173, pp. 111–130.

Demuth, S., Young, A., Hisdal, H., Tallaksen, L.M. (2004) Regional estimation methods. To appear in: Tallaksen, L.M. and van Lanen, H.A.J. (eds.), 2004. Hydrological Drought: Processes and Estimation Methods for Streamflow and Groundwater. Developments in Water Science no. 48, Elsevier, NL.

Dingman, S.L. and Lawlor, S.C. (1995) Estimating low flow quantiles from drainage-basin characteristics in New Hampshire and Vermont. Water Resour. Bull. 31 2, pp. 243–256.

Dobesch, H. (2003) Mean annual potential evaporation. In: BMLFUW (ed.) Hydrologischer Atlas Österreichs, Vienna, Austria.

Efron, B., and Tibshirani, R.J. (1993) An introduction to the bootstrap. Monographs on statistics and applied probability 57, Chapman and Hall, New York.

Elder, K. (1995) Snow distribution in alpine watersheds, *PhD dissertation*, University of California, Santa Barbara, 309pp.

Fiering, M.B. (1963) Use of correlation to improve estimates of the mean and variances. U.S. Geol. Surv. Prof. Pap., 434-C, C1-C9.

Fuchs, G. and Fürst, J. (2003a) Mean depth to groundwater table. In: BMLFUW (ed.) Hydrologischer Atlas Österreichs. Österreichischer Kunst- und Kulturverlag, Wien.

Fuchs, G. and Fürst, J. (2003b) Mean annual fluctuation of the groundwater table. In: BMLFUW (ed.) Hydrologischer Atlas Österreichs. Österreichischer Kunst- und Kulturverlag, Wien.

Fürst, J. (2003) Gewässernetzdichte (stream density). In: BMLFUW (ed.) Hydrologischer Atlas Österreichs. Österreichischer Kunst- und Kulturverlag, Wien.

Grayson, R. and Blöschl, G. (2002) (Eds.) Spatial patterns in catchment hydrology - Observations and modelling. Cambridge University Press.

Gustard, A., Bullock, A., Dixon, J.M. (1992) Low flow estimation in the United Kingdom. Institute of Hydrology, Report No. 108, 88 pp., append.

Gustard, A., Irving, K.M. (1994) Classification of the low flow response of European soils. FRIEND: Flow Regimes from International Experimental and Network Data. IAHS Publication No. 221, pp. 113–117.

Gustard, A., Young, A.E., Rees, G., Holmes, M. (2004) Low flows, drought and operational hydrology. To appear in: Tallaksen, L.M. and van Lanen, H.A.J. (eds.), 2004. Hydrological Drought: Processes and Estimation Methods for Streamflow and Groundwater. Developments in Water Science no. 48, Elsevier, NL.

Haines, A.T., Finlayson, B.L., McMahon, T.A. (1988) A global classification of river regimes. Appl. Geog., 8: 255-272.

Hartigan, J. A. (1975) Clustering Algorithms. New York: Wiley.

Hayes, D.C. (1992) Low flow characteristics of streams in Virginia. U.S. Geological Survey, Water Supply Paper 2374.

Holmes, M.G.R., Young, A.R., Gustard, A. and Grew, R. (2002) A region of influence approach to predicting flow duration curves within ungauged catchments. *Hydrology and Earth System Sciences*, vol 6, 721 – 733.

Hydrographisches Zentralbüro (1996) (Ed.) Hydrographisches Jahrbuch von Österreich1996. 104. Band, Bundesministerium für Land- und Forstwirtschaft, Wien.

Institute of Hydrology (1980) Low Flow Studies Report. Crowmarsh Gifford, Wallingford, Oxon, UK.

Kaufman, L. and Rousseeuw, P.J. (1990) Finding Groups in Data: An introduction to cluster Analysis. Wiley, New York.

Kresser, W. (1965) Österreichs Wasserbilanz. *Österr. Wasserwirtschaft*, 17, 213-221.

Kresser, W., Kirnbauer, R., Nobilis, F. (1985) Überlegungen zur Ermittlung von Niederwasserkenngrößen. Mitteilungsblatt des Hydrographischen Dienstes Nr. 54. Wien.

Kroiß, H., Gutknecht, D., Schmid, C., Blöschl, G., Fenz, R., Hinteregger, J., Honsowitz, H. (1996) Methodik und Kritik der ökologischen und ökonomischen Bewertung von regionalen Lösungen zur Abwasserentsorgung am Beispiel des Lainsitztales. Wiener Mitteilungen, Band 130, TU Wien.

Kundzewicz, Z.W. and Robson, A. (eds.) (2000) Detecting trend and other changes in hydrological data. World climate programme - Data and monitoring, WCDMP-45, WMO/TD-No. 1013, CEH Wallingford, UK.

Laaha G., Blöschl G. (2003) Saisonalität von Niederwasserspenden in Österreich. Mitteilungsblatt des Hydrographischen Dienstes in Österreich, 82.

Laaha, G. (2000) Statistik-Experten-Tools für Niederwasserfragestellungen. Forschungsbericht beim BMLFUW, Abt. IV A 3.

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 (Proceedings of the Fourth International Conference held in Cape Town, South Africa, March 2002). IAHS Publ. no. 274, 289-295 .

Lorenz, P. and Skoda, G. (1999) Ermittlung von Flächenmitteln des Niederschlags aus punktuellen Messungen, Mitteilungsblatt des Hydrographischen Dienstes in Österreich, Nr. 78, p. 47-66.

Mardia, K. (1972) Statistics of directional data: Probability and mathematical statistics. Academic Press, London.

Matalas, N.C. and Jacobs, B. (1964) A correlation procedure for augmentating hydrologic data. U.S. Geol. Surv. Prof. Pap., 434-E, E1-E7.

Merz, R. and Blöschl, G. (2004a) Flood frequency regionalisation – spatial proximity vs. catchment attributes. To appear in *J. Hydrol.*

Merz, R. and Blöschl, G. (2004b) Regionalisation of catchment model parameters. To appear in *J. Hydrol.*

Merz, R., Piock-Ellena, U., Blöschl, G. & Gutknecht, D. (1999) Seasonality of flood processes in Austria. IAHS Publ. no. 255, 273-278.

Nathan R.J. & McMahon T.A. (1990) Identification of homogeneous regions for the purpose of regionalization. J. Hydrol. 121, 217-238.

NERC (1975) Flood Studies Report (in five volumes). Natural Environment Research Council, London.

Pearson, C.P., 1995. Regional frequency analysis of low flows in New Zealand rivers. J. Hydrol. NZ 30 2, pp. 53–64.

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: 13-21, Baden.

Rieger, W. (1996) Accuracy of Slope Information Derived from DEM-Data. XVIII. ISPRS Congress, Vienna, Commission IV/4, Int. Arch. Photogr. & Rem. Sens., Vol. 31/B4: 690-695.

Robson, A. (1999) Statistical procedures for flood frequency estimation. In: Flood estimation handbook, Volume 3, Institute of Hydrology, Wallingford, UK.

Rosenthal W. & Dozier J. (1996) Automated mapping of montane snow cover at subpixel resolution from the Landsat Thematic Mapper. Wat. Ressour. Res. 32, 115-130.

Schreiber, P. and Demuth, S. (1997) Regionalisation of low flows in southwest Germany. Hydrol. Sci. J. 42 6, pp. 845–858.

Skoda, G. and Lorenz, P. (2003) Mean annual precipitation - modelled with uncorrected data. In: BMLFUW (ed.) Hydrologischer Atlas Österreichs. Österreichischer Kunst- und Kulturverlag, Wien.

Skop, E. and Loaiciga, H.A. (1998) Investigating catchment hydrology and low-flow characteristics using GIS. Nordic Hydrol. 29 2, pp. 105–128.

Smakhtin V.U. (2001) Low flow hydrology: a review. J. Hydrol. 240, 147-186.

Stedinger, J.R., Vogel, R.M. and Foufoula-Georgiou, E. (1992) Frequency analysis of extreme events. In Handbook of Hydrology, Ed. By Maidment, D.R. McGraw Hill New York, 18.1-18.66.

Tallaksen L.M. and 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 (ed. by A. Gustard, S. Blazkova, M. Brilly, S. Demuth, J. Dixon, H. van Lanen, C. Llasat, S. Mkhandi and E. Servat), 141-150. IAHS Publ. No. 246.

Tallaksen, L.M. and van Lanen, H.A.J. (eds.), 2004. Hydrological Drought: Processes and Estimation Methods for Streamflow and Groundwater. Developments in Water Science no. 48, Elsevier, NL.

Tucci, C., Silveira, A., Sanchez, J. and Albuquerque, F. (1995) Flow regionalisation in the upper Paraguay basin, Brazil. Hydrol. Sci. J. 40 4, pp. 485–497.

Veanbles, W.N., and Ripley, B.D. (1999) Modern applied statistics with S-Plus. Third edition. Springer, New York.

Vogel, R.M. and Kroll, C. N. (1991) The value of streamflow record augmentation procedures in low-flow and flood-flow frequency analysis. J. Hydrol., 125, (3-4), pp 259-276.

Vogel, R.M. and Kroll, C.N. (1992) Regional geohydrologic-geomorphic relationships for the estimation of low-flow statistics. Water Resour. Res. 28 9, pp. 2451–2458.

Vogel, R.M. and Stedinger, J.R. (1985) Minimum variance streamflow record augmentation procedures. Water Resources Research, Vol.21, No.5, p.715-723.

Weisberg, S. (1985) Applied linear regression. John Wiley and sons, New York.

Young A.R., C.E. Round, A. Gustard (2000) Spatial and temporal variations in the occurrence of low flow events in the UK. Hydrology and Earth System Sciences, 4(I), 35-45.