Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Hydrology 306 (2005) 264–286
Journal of Hydrology
www.elsevier.com/locate/jhydrol

Low flow estimates from short stream flow records—a comparison of methods

G. Laaha[a,*], G. Blöschl[b]

[a]Institut für Angewandte Statistik und EDV, Universität für Bodenkultur Wien, Gregor Mendel Str. 33, A-1180 Wien, Austria
[b]Institut für Hydraulik, Gewässerkunde und Wasserwirtschaft, Technische Universität Wien, Karlsplatz 13/223, A-1040 Wien, Austria

Received 20 January 2004; revised 3 September 2004; accepted 17 September 2004

Abstract

We compare a number of methods of adjusting Q95 estimates from short stream flow records for climate variability. Q95 is the discharge that is exceeded on 95% of all the time for one particular site. 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. 132 catchments in Austria are used with catchment areas ranging from 9 to 479 km^2. 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 1-year records, and from 86 to 93% for 3-year records when adjusting the estimates by the downstream site method. For 5 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 suggest that, on average over the study region, 1 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.

Keywords: Climate variability adjustment; Stream flow record augmentation; Regionalisation; Regional regression; Low flow seasonality index; Droughts

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

* Corresponding author.
E-mail addresses: gregor.laaha@boku.ac.at (G. Laaha), bloeschl@hydro.tuwien.ac.at (G. Blöschl).

0022-1694/$ - see front matter
doi:10.1016/j.jhydrol.2004.09.012

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 adjust low flow characteristics to longer-term climatic conditions, 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 we term the subject site) where a short stream flow record is available, based on stream flow data from other catchments (which we term donor sites) where long records are available. While a subject site provides information about the local specifics of the low flow regime, donor sites provide information about how the short record at the subject site fits into the long-term picture. 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; Laaha and Blöschl, 2004b) 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 (Laaha and Blöschl, 2004a; 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; Merz and Blöschl, 2004a,b). 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 Laaha and Blöschl (2004b).

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 95% of all the time) 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 our 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 paper therefore is to examine the relative performance of different climate adjustment techniques for estimating low flow characteristics from short stream flow records. We 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 the time 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 paper is organised as follows: Section 2 summarises the data used. Section 3 details the methods of climate adjustment examined in this paper 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. Results of the comparisons are presented in Section 5 and discussed in Section 6. Section 7 gives conclusions.

2. Data

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).

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, karst effects or lake storage during low flow periods (Laaha and Blöschl, 2004a).
- 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 km^2, which is about 60% of the national territory of Austria. Although a larger number of catchments are monitored in Austria, we have chosen to give priority to a consistent observation period to make all records comparable in terms of climatic variability. We use all of these 325 catchments as possible donor sites.

For the subject sites, i.e. the sites where we test the value of short stream flow records, we have chosen to only use those catchments that do not have an upstream neighbouring gauged catchment. We 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 (Laaha and Blöschl, 2004b). 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 we therefore excluded those catchments that did not have a downstream neighbour. What remained was a set of 132 gauged catchments which we used as subject sites in this paper. These are

Table 1
Characteristics of the 132 catchments used as subject sites

	Minimum	25%	50%	75%	Maximum	Mean
Q95 (m^3/s)	0.013	0.194	0.449	0.927	3.890	0.692
q95 ($l\ s^{-1}\ km^{-2}$)	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

Q95 are low flow discharges, q95 are specific low flow discharges, area is the catchment area. The percent values are the quantiles.

the sites for which we analyse the effects of record length and climate adjustment method on estimating low flow characteristics.

2.3. Low flow characteristics

The low flow characteristic chosen in this paper is the flow quantile Q95, i.e. the discharge equalled or exceeded during 95% of the observation period ($\Pr(Q \geq Q95 = 0.95)$). 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 1 along with those of the specific discharges q95 and the catchment areas.

2.4. Catchment characteristics

One of the investigated donor selection techniques is based on hydrological similarity of catchments. To define the similarity measures, we 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, which contains detailed information about mineral resources in Austria, distinguishes 670 geological classes. From these we derived nine hydrogeological classes we 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, we condensed the original classification of the Corine Landcover map. The Corine Landcover map (Coordination of Information on the Environment program of the EU, Aubrecht, 1998), originally consisted of 44 standardised soil cover and land use classes from satellite data at a 1:100,000 scale based on computer-aided visual image interpretation. These classes we reduced to nine main 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.

3. Climate adjustment techniques

3.1. General concept

Our 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. We 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.

Table 2
Statistical summary of the characteristics of the 325 catchments used in this paper

Acronym	Variable description	Units	Minimum	Mean	Maximum
A	Catchment area	km^2	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}	Deep 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/km^2	160	790	1320

3.2. Donor selection

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.

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.

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), we 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. We 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. We adopted the classification of Austria into eight regions of Laaha and Blöschl (2004a). 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.

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. We 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, we 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 five 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:

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

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.

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. We 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:

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

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

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:

$$\mathrm{QS}_{\mathrm{pred}} = \prod_{i=1}^{n} (\mathrm{QS}_{\mathrm{pred}}^{(i)})^{w_i/\Sigma w_i} \quad (7)$$

where w_i is a weight for the ith donor and $\mathrm{QS}_{\mathrm{pred}}^{(i)}$ is Q95 at the subject site adjusted by the ith donor.

4. Evaluation method

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 we have the full record length for all subject sites, so we can compare the adjusted low flow characteristic $\mathrm{Q95}_{\mathrm{pred}}$ for hypothetically shortened records with the low flow characteristic $\mathrm{Q95}_{\mathrm{obs}}$ estimated from the complete records, which gives us 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 we 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. We 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.

4.2. 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):

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

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

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 $q95_{pred} = Q95_{pred}/A$ ($l\,s^{-1}\,km^{-2}$) 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 km^2. 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, we use the 5% trimmed RMSE instead. This means that 5% of the catchments (in our 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 we compared all results in this paper 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 $Q95_{obs}$:

$$re_C = \mathrm{RMSE}_C / m_C(Q95_{obs}) \tag{10}$$

where $m_c(Q95_{obs})$ 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.

For ease of comparison with other low flow studies we 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. We therefore only evaluated R^2 of q95 specific low flow discharges:

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

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), we 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

Table 3
Class limits and class means for estimating relative errors (m^3/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

adjusted. From this we estimated the gain in information by

$$\text{gain} = \left(\frac{n_{\text{eff}}}{n} - 1\right)100 = \left(\left(\frac{\text{RMSE}_{\text{no\_donor}}}{\text{RMSE}_{\text{adjusted}}}\right)^2 - 1\right)100 \quad (12)$$

where n_{eff} is the effective record length and n is the record length of the subject site. Eq. (12) is based on Eq. (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.

5. Results

5.1. Errors of unadjusted low flow characteristics

As a starting point we 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. 1 shows the relative errors (Eq. (10)) of Q95 for this case as a function of low flow discharge $Q95_{\text{obs}}$. 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. 1), the errors decrease from 30% to 16, 12, 5 and 3% as one moves from 1 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.

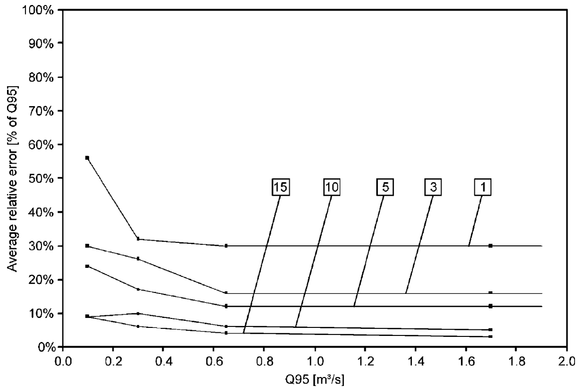

Fig. 1. Relative errors re_c of low flow discharge $Q95_{\text{pred}}$ estimated from records of less than 20 years as compared to 20 year records, plotted versus the Q95 low flow discharge. No climate adjustment. Numbers in boxes are the record lengths in years.

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. 2 gives the absolute errors (RMSE) for discharges Q95 (m^3/s), Fig. 3 gives the absolute errors (RMSE) for specific discharge q95 ($l\,s^{-1}\,km^{-2}$) and Fig. 4 gives

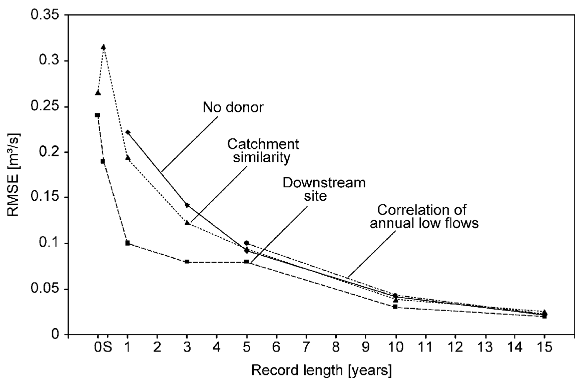

Fig. 2. Absolute errors RMSE (m^3/s) of low flow discharge $Q95_{\text{pred}}$ estimated from records of less than 20 years as compared to 20-year records. Various climate adjustment techniques are used. 0, no local stream flow data; S, spot gaugings.

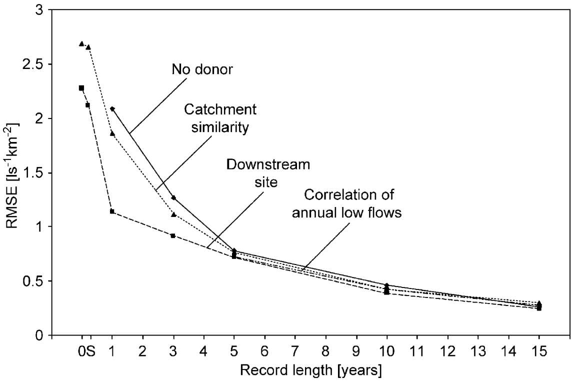

Fig. 3. Absolute errors RMSE (l s^{-1} km^{-2}) of specific low flow discharge $q95_{pred}$ estimated from records of less than 20 years as compared to 20-year records. Various climate adjustment techniques are used. 0, no local stream flow data; S, spot gaugings.

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. 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. 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. 2).

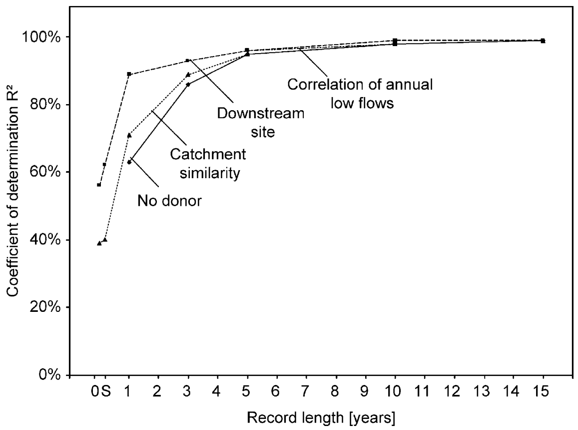

Fig. 4. Coefficient of determination R^2 (%) of specific low flow discharge $q95_{pred}$ estimated from records of less than 20 years as compared to 20-year records. Various climate adjustment techniques are used. 0, no local stream flow data; S, spot gaugings.

Figs. 2–4 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. 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 to 2.1, 1.1, 0.9, 0.7, 0.4 and 0.3 l s^{-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. 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 (Figs. 3 and 4) and it is slightly poorer for the error measure based on low flow discharges (Fig. 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 l s^{-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 do not. 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 5 years but for longer records the gain of using these adjustment techniques is relatively modest.

5.3. Relative performance of record augmentation techniques

We 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. 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

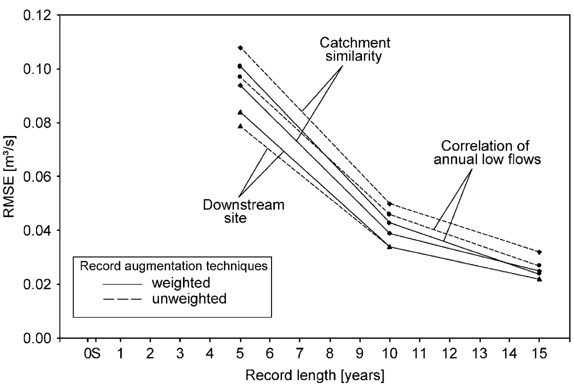

Fig. 5. Absolute errors RMSE (m^3/s) of low flow discharge $Q95_{pred}$ estimated from records of less than 20 years as compared to 20-year records. Three donor selection techniques are combined with two record augmentation methods (weighted: Eq. (5); unweighted: Eq. (4)).

Table 4
Gain (%) in effective record length by various climate adjustment methods based on estimated low flow discharges $Q95_{pred}$

	Record length (years)				
	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

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)).

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 1-year record and 91% for the 3-year record. For 5 years, the gain is either 17 or 40%, depending on the augmentation method, which means

Table 5
Gain (%) in effective record length by various climate adjustment methods based on estimated specific low flow discharges $q95_{pred}$

	Record length (years)				
	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

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)).

that the adjusted 5-year record is equivalent to an unadjusted 5.9- or 7-year record. The downstream method gains 53% for a 10-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 15-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.

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. We therefore plotted the low flow discharges estimated for various record lengths ($Q95_{pred}$) against the low flow discharges estimated

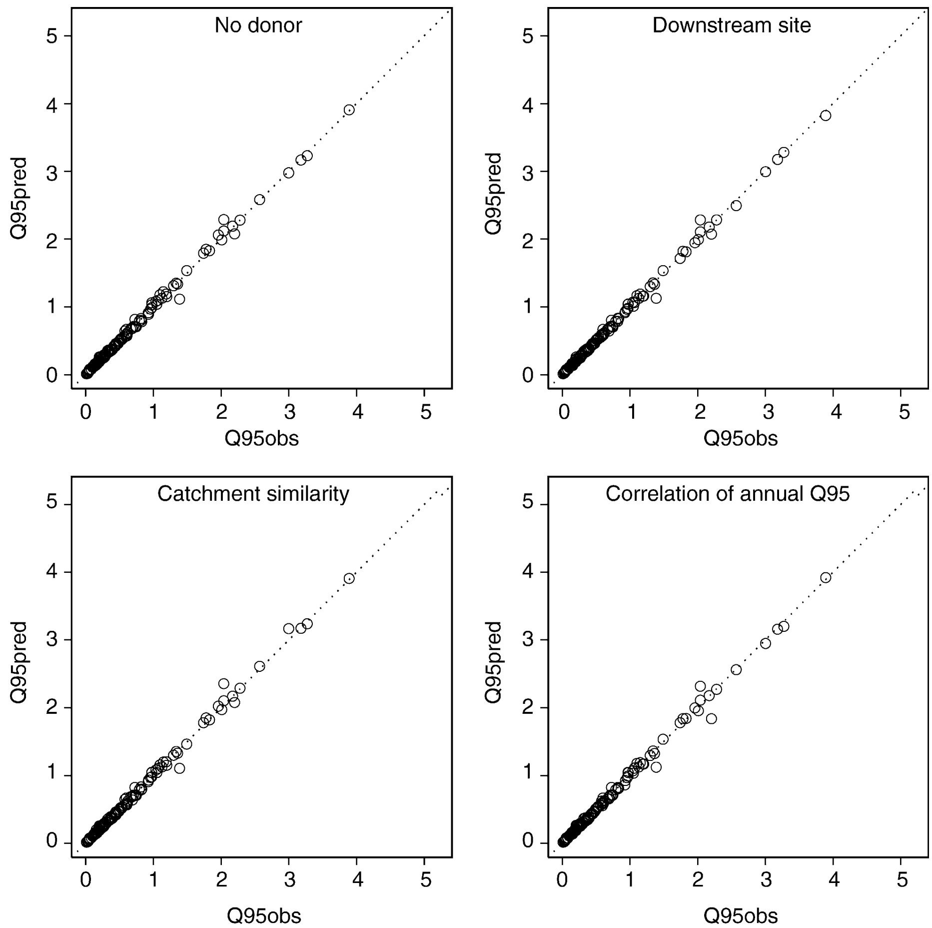

Fig. 6. Adjusted low flows $Q95_{pred}$ (m^3/s) estimated from 15-year records plotted versus low flows $Q95_{obs}$ (m^3/s) estimated from the full 20-year period. Each point represents a catchment and the panels relate to different donor selection methods.

for the full record length of 20 years ($Q95_{obs}$) in Figs. 6–10. These scatter plots also allow us 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. 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. 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 5-year records (Fig. 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 1 year of observation (Fig. 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

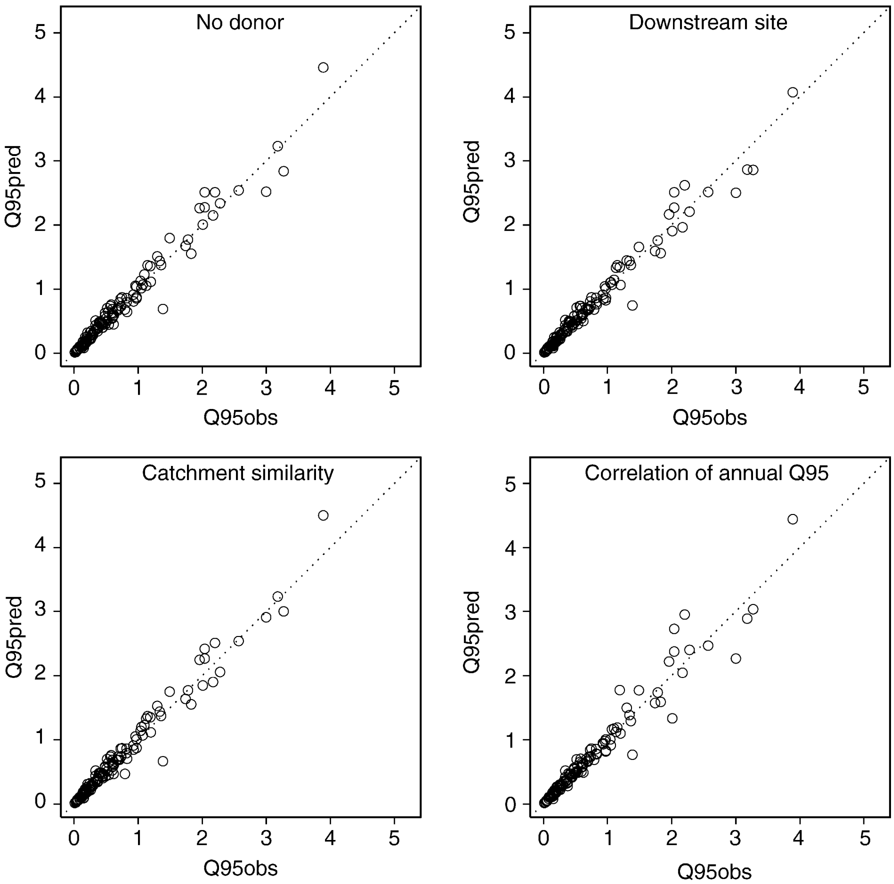

Fig. 7. Adjusted low flows $Q95_{pred}$ (m^3/s) estimated from 5-year records plotted versus low flows $Q95_{obs}$ (m^3/s) estimated from the full 20-year period. Each point represents a catchment and the panels relate to different donor selection methods.

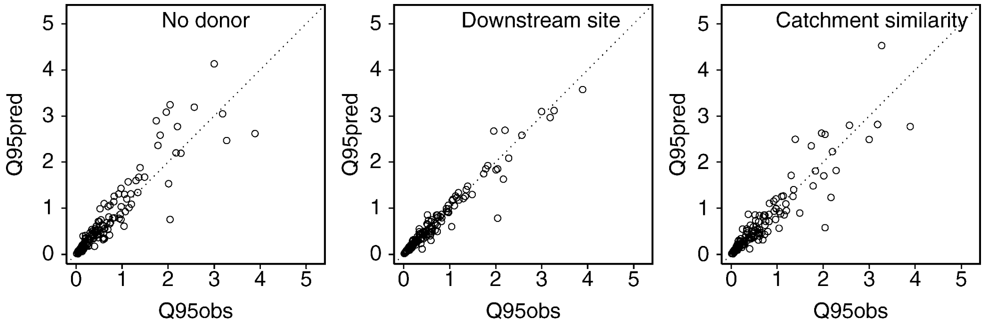

Fig. 8. Adjusted low flows $Q95_{pred}$ (m^3/s) estimated from 1-year records plotted versus low flows $Q95_{obs}$ (m^3/s) estimated from the full 20-year period. Each point represents a catchment and the panels relate to different donor selection methods.

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. 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 1 year case in Fig. 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. 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. 2 for the spot gauging case than for the no data case. From a visual comparison of Figs. 9 and 10 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. 2–4 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.

5.5. Spot gaugings

To analyse the error sources of the spot gauging and no donor (regionalisation) cases we calculated

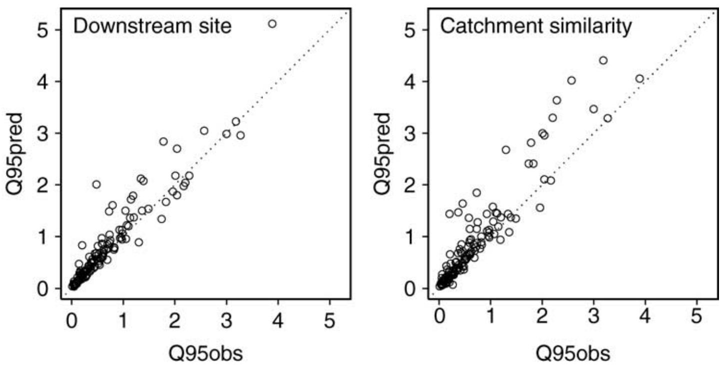

Fig. 9. Low flows $Q95_{pred}$ (m^3/s) estimated from spot gaugings plotted versus low flows $Q95_{obs}$ (m^3/s) estimated from the full 20-year period. Each point represents a catchment and the panels relate to different donor selection methods.

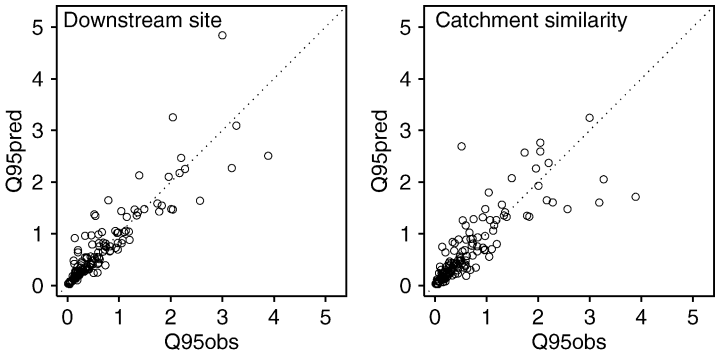

Fig. 10. Low flows $Q95_{pred}$ (m^3/s) estimated from a simple regionalisation model (i.e. no local stream flow data) plotted versus low flows $Q95_{obs}$ (m^3/s) estimated from the full 20-year period. Each point represents a catchment and the panels relate to different donor selection methods.

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. 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. 11 top (downstream site method) indicates that the uncertainty introduced by the spatial variability (dotted line) is about the same as the uncertainty introduced by the timing of the spot gaugings (dashed line). The combined effect of

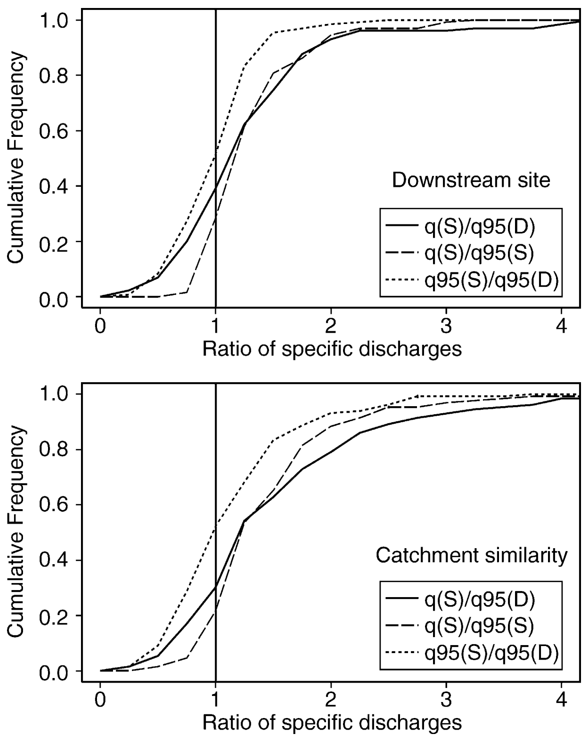

Fig. 11. Cumulative frequency distribution of specific discharge ratios. q95(S) is the specific low flow discharge exceeded on 95% of all days at the subject site, q95(D) is the analogous value for the donor site, and q(S) is the specific discharge measured by spot gauging at the subject site.

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. 11 bottom) with the downstream site method of donor selection (Fig. 11 top). The catchment similarity method is associated with a wider spread in the q95(S)/q95(D) distribution (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.

5.6. Effect of discharges

A final assessment in this paper (Fig. 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 5-year curve slightly crosses over some of the other curves

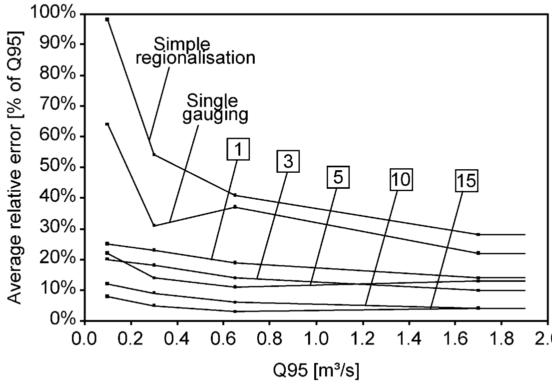

Fig. 12. Relative errors re_c of low flow discharge $Q95_{pred}$ estimated from records of less than 20 years as compared to 20-year records, plotted versus the Q95 low flow discharge. Climate adjustment based on the downstream site donor selection technique. Numbers in boxes are the record lengths in years.

which likely is an random artefact of the data and not a significant pattern.

Fig. 12 is a similar representation as Fig. 1 but the difference is that Fig. 1 is without climate adjustment while Fig. 12 is with climate adjustment based on the nearest downstream site. The degree to which the errors in Fig. 12 are smaller than those in Fig. 1 is a measure of the value of the climate adjustment procedure as a function of low flow discharge. The error pattern in Fig. 12 is similar to that in Fig. 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.

6. Discussion

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. One would therefore expect the value of using a donor site to decrease with the ratio of donor and subject site catchment areas. An additional analysis (not shown in this paper) was carried out to examine what the maximum ratio of catchment areas would be for climate adjustments to be of use. At-site estimation errors were plotted versus the ratio of areas for a given record length. Results for different record lengths indicated a sudden increase of estimation errors for a ratio larger than about 10 (i.e. where downstream donor catchments were more than 10 times the size of the subject site catchments). For the conditions in the study area this is likely an upper limit for the downstream donor sites to be of use for climate adjustments. 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 paper 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 (Laaha and Blöschl, 2004b).

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 paper are significantly shorter than most of the record lengths examined in Vogel and Kroll (1991). It is likely that, if we 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. 1) and the low flows adjusted by the best method (downstream method, Fig. 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.

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 1-year record, and from 86 to 93% for the 3-year record but the increase is much smaller for 5 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 1-year record and 91% for the 3-year record. For 5 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- 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 paper, the downstream method gains 53% for a 10-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 15-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 paper 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 paper and they had significantly fewer catchments than in this paper, 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.

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 paper, 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 1-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.

Laaha and Blöschl (2004b), for example, have compared a number of low flow regionalisation methods in the same study area. Their study was based on discharges of catchments without upstream gauges, as in this paper, but they also included discharges of residual catchments between subsequent gauges with a total of 325 catchments rather than 132 catchments as in this paper. They assessed the regionalisation errors, among other measures, by the cross validated coefficient of determination. This is similar to the coefficient of determination in this paper although they did not trim their statistics, so their coefficients of determination likely are a little lower than trimmed coefficients of determination as used in this paper. Their global regression model that 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 paper ($R^2=56\%$, Fig. 4). Their optimal regionalisation model, however, yielded $R^2=70\%$. This model was based on separate regressions in eight seasonality zones. If we compare this to the results of this paper we can see from Fig. 4 that 1 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 1 year of stream flow data clearly outperforms the regionalisation method while the regionalisation of Laaha and Blöschl (2004b) performs better than the spot gauging method. It should also be noted that the performance of their regionalisation method varied significantly between the zones, ranging from $R^2=51$ to 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 1 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. 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), our 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 paper 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 we extended the sampling to a number of spot gaugings during more than one low flow period.

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 1-year records, and from 86 to 93% for 3-year records. When expressed in terms of effective record lengths of q95, the gain is 236% for the 1-year record and 91% for the 3-year record. The value of the climate adjustment methods is much smaller for records of 5 years and more. For 5 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 paper with more sophisticated regionalisation procedures from Laaha and Blöschl (2004b) suggest that, on average over the study region, 1 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.

This latter finding has important implications for network planning in the face of reduced budgets of hydrographic agencies. It is sometimes argued that suitable regionalisation techniques (e.g. based on catchment attributes) can be used to make up for a loss of information resulting from downsized stream gauge networks. The results of this study suggest that there is indeed a lot of information in actual stream flow data, even if the record lengths are short. Regionalisation techniques can be used for estimating low flow characteristics but typical stream flow record lengths of significantly more than 1 year allow for much better accuracies than what can be achieved by regionalisation. Small differences in low flow estimates may have dramatic economic consequences, e.g. when used for estimating environmental flows. Downsizing hydrographic networks may therefore not be a wise decision when the economic value of accurate low flow estimates is factored in.

Austria is a relatively well gauged country and in most countries of the world the stream gauge density is lower. It would be interesting to examine what is the effect of network density on the value of short record lengths for low flow estimation, relative to regionalisation methods. Ultimately, an approach that combines climate adjustment techniques, based on short records, and regionalisation methods may exploit the maximum information available at a particular site. This is the subject of ongoing work.

Acknowledgements

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

References

Andrews, D.F., 1972. Plots of high dimensional data. Biometrics 28, 125–136.

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.

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

Fiering, M.B., 1963. Use of correlation to improve estimates of the mean and variances. US Geological Survey Professional Paper 434-C, C1–C9.

Hayes, D.C., 1992. Low flow characteristics of streams in Virginia. US Geological Survey Water Supply Paper 2374 1992.

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

Kresser, W., 1965. Österreichs Wasserbilanz. Österreichs 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., 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 2003; 82.

Laaha, G., Blöschl, G., 2004a. Seasonality indices for regionalising low flows. Hydrological Processes 2004; (Submitted for publication).

Laaha, G., Blöschl, G., 2004b. A comparison of low flow regionalisation methods—catchment grouping. Journal of Hydrology 2004; (Submitted for publication).

Lorenz, P., Skoda, G., 1999. Ermittlung von Flächenmitteln des Niederschlags aus punktuellen Messungen. Mitteilungsblatt des Hydrographischen Dienstes in Österreich 78, 47–66.

Matalas, N.C., Jacobs, B., 1964. A correlation procedure for augmentating hydrologic data. US Geological Survey Professional Paper 434-E, E1–E7.

Merz, R., Blöschl, G., 2004a. Flood frequency regionalisation—spatial proximity vs. catchment attributes. Journal of Hydrology 2004; (in press).

Merz, R., Blöschl, G., 2004b. Regionalisation of catchment model parameters. Journal of Hydrology 287, 95–123.

Merz, R., Piock-Ellena, U., Blöschl, G., Gutknecht, D., 1999. Seasonality of flood processes in Austria. IAHS Publication 255, 273–278.

Nathan, R.J., McMahon, T.A., 1990. Identification of homogeneous regions for the purpose of regionalization. Journal of Hydrology 121, 217–238.

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

Robson, A., 1999. Statistical procedures for flood frequency estimation. In: Flood Estimation Handbook, vol. 3, Institute of Hydrology, Wallingford, UK.

Smakhtin, V.U., 2001. Low flow hydrology: a review. Journal of Hydrology 240, 147–186.

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

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

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

Vogel, R.M., Stedinger, J.R., 1985. Minimum variance streamflow record augmentation procedures. Water Resources Research 21 (5), 715–723.