Printed in Water Resources Management: Volume 23, Issue 12 (2009), Page 2567-2586.

A comparison of low flow estimates in ungauged catchments using regional regression and the HBV-model.

K. Engeland[1], H. Hisdal[2]

[1]Sintef Energy Research, N-07465 Trondheim, Norway, **kolbjorn.engeland@sintef.no**

[2]The Norwegian Water Resources and Energy Directorate, P.O. Box 5091 Majorstua, N-0301 Oslo, Norway

Abstract

Estimates of a low flow index in ungauged catchments calculated by a regional regression model and a regional hydrological model were compared for a study region southwestern Norway. The regression method was based on a relationship between the low flow index and an optimal set of catchment descriptors, established using stepwise linear regression for homogeneous subregions. Subregions were distinguished according to the season in which the lowest flow occurs, winter (May to October) or summer (November to April), and the average July temperature was found to be the best index for determining the low flow season for ungauged catchments. Catchment descriptors characterising the presence of lakes and bogs, in addition to catchment length and indicators of climatic conditions, were found to be important in the regression models. A cross-validation procedure was used to evaluate the predictive performance of the model in ungauged catchments. A gridded version of HBV, a daily rainfall-runoff model was also applied as a regional hydrological model and was calibrated using the average Nash-Sutcliffe coefficient for log-transformed streamflow as the calibration criterion. A comparison of the two methods in 21 independent catchments indicates that the regression method generally gives better estimates of Qc in ungauged catchments than does the HBV model , particularly in those catchments with the lowest Qc values.

Key words: Low flow index; ungauged catchment; regional regression; rainfall-runoff model

Introduction

Information about low flows is required in water resources management, for example to estimate hydropower energy production, to design abstraction schemes for public water supply, fish farming and irrigation and to estimate dilution of effluents. Low flows are often characterized by indices, i.e. single numbers describing an aspect of the low flow behaviour at a site or in a region. Assuming that a stationary flow record of a certain length undisturbed by human influence is available, various low flow indices can be calculated. Frequently applied indices are percentiles from the flow duration curve and mean annual minimum flows. In Hisdal *et al.* (2004) the derivation of various low flow indices and the interrelationships between indices are described.

Often the low flow indices are needed for ungauged river basins or at sites where data are incomplete, and regionalisation techniques are therefore essential in operational hydrology. In Norway increasing requests to build small hydropower plants has led to a growing demand for low flow data, especially for small ungauged catchments. The Norwegian Water Resources Act requires estimation of a specific low flow index, the “common low flow”, Q_C, if a hydropower plant is to be constructed. This index is often used as a starting point to set residual flow in the licensing procedure. Q_C is calculated as follows: a) use a flow record with a daily time resolution (preferably 15-20 years of data); (b) remove the 15 smallest values every year; (c) calculate the annual minimum series; and (d) rank the values in the annual minimum series and remove the 1/3 smallest values. The smallest value remaining is defined as the Q_C. Based on order statistics, it can be shown that for a sample of iid values, the r^{th} order statistics p_r is asymptotically normally distributed (David and Nagaraja, 2003):

$$p_r \sim N\left(\frac{r}{n+1}, \frac{r(n-r)}{(n+1)^2(n+2)^3}\right) \tag{1}$$

Using r = 16 and n = 365, and calculating the 1/3 quantile for p_r, we obtain p_{clf} = 0.96. This means that Q_C is approximately the 0.96 quantile of the flow duration curve, i.e. the flow that is exceeded 96 percent of the time. We will use Q_C as an example of a low flow index, even though this index is only used in Norway. The results and methodology are, however, of general interest since Q_C is closely related to Q_{95}, a low flow index that is widely applied. The conclusions would not have changed if Q_{95} was used instead of Q_C.

Many decisions in water resources management include some degree of subjectivity. In water resources administration, it is important to establish methods and procedures where the outcome does not depend on the individual officer in charge. A common procedure for estimation of low flow indices at an ungauged site is to select a donor catchment. This procedure includes subjectivity in the choice of donor catchments and how to transfer the low flow index from the donor catchment to the ungauged catchment. An objective method to estimate low flow indices at ungauged sites is therefore required.. In the literature, two basically different methods are presented, the stochastic or the deterministic approach (e.g. Smakhtin, 2001).

In the stochastic approach, the streamflow statistics at ungauged sites are conditioned on the streamflow statistics at gauged sites using either catchment descriptors or spatial distance as similarity measures. The streamflow statistics (e.g. a low flow index) can be related to catchment characteristics such as area, land use or geology via multiple regression (see Demuth and Young (2004) for an overview of these methods). Alternatively, geostatistical interpolation approaches can be used to explore the whole spatial-temporal correlation structure of the runoff field (e.g. Gottschalk *et al*., 2006, Skøien *et al*., 2006) although these methods are not frequently implemented for operational use. Interpolation, top-kriging, is compared to regional regression in Laaha *et al*. (2007), and they conclude that regression outperforms interpolation for small catchments and headwater catchments in regions with scarce station density. In this paper we, therefore, choose the regression approach since the density of the streamflow observations is very low compared to the correlation-length of low flow indices in most of Norway. The region studied has large precipitation and climatic gradients that make the variability in hydrology over short distances very large. In addition the interpolation method does not account for lakes, which are a very pronounced characteristic of the Norwegian landscape and are especially important for low flows.

The regression approach is widely applied to predict low flows in ungauged catchments. Smakhtin (2001) and Demuth (2004) give extensive lists of references to applications in Canada, USA, Australia, New Zealand, Slovenia, Slovakia, Greece, Japan, UK, Germany and Norway. Some of the earlier publications include Thomas and Benson (1970) in the USA, Leith (1978) in Canada, the low flow studies report by Institute of Hydrology (1980) in the UK, and Krokli (1988) in Norway. Recent publications include low flow estimation as a part of the StreamStats software of USGS USA (Ries, 2002), the Low Flows 2000 software (Gustard *et al*., 2004) in the UK, and a national procedure for low flow estimation in Austria (Lahaa and Blöschl, 2007).

In this paper a regional regression approach in which large or heterogeneous domains are grouped into homogenous regions with respect to low flow processes is applied. Regression equations are established for each region independently. It is necessary to define either geographically continuous regions or regions defined by catchment and climate characteristics. Smakhtin (2001) and Laaha (2006) give reviews of strategies to define homogeneous regions. These include weighted cluster analysis (Nathan and McHahon (1990) regression tree and residual pattern analysis (Laaha and Blöschl, 2006). The most appropriate classification procedure to use depends on the climate and landscape characteristics. Laaha (2006) show that for Austrian catchments, a grouping based on seasonality gives the best prediction of low flow indices, whereas Young *et al.* (2000) show that in the UK where soil classes, rather than seasonality, should be used to define homogeneous regions.

In the deterministic approach, a precipitation-runoff model is used to generate a continuous streamflow time series at ungauged sites from which the desired stream flow statistics can be extracted. Smakhtin (2001) reviews this method, and finds that applications of the method are rather limited in number. Previous examples include Smakhtin and Watkins (1997) in South Africa, Clausen and Rasmussen (1993) in Denmark, and Lanmen *et al.* (1993) in Europe. In order to use a rainfall runoff model for low flow estimations, calibration criteria that provide information about the quality of the low flow simulations are necessary (Smakhtin, 2001). To calculate the runoff at ungauged sites, the model parameters have to be transferred to the ungauged sites (e.g. Engeland, 2006).

The choice of regionalisation method, stochastic or deterministic, depends on data availability and the purpose of the regional estimates. The deterministic approach is preferred when more explicit knowledge of the hydrological processes is required, for example to obtain runoff statistics for climate change or land-use change scenarios, and the stochastic approach is most frequently applied for prediction at the ungauged site (Smakhtin, 2001).

A comparison of the two approaches for low flow estimation at the ungauged site is lacking in the literature. The aim of this paper is, hence, to evaluate and compare regression- and precipitation-runoff modelling methods to estimate low flow indices in small ungauged catchments (catchment area less than 2000 km^2). The study region is located in Southern Norway and includes 51 pristine catchments with suitable streamflow records. Regression equations were established between Q_C and catchment characteristics. A gridded version of the HBV-model was calibrated using objective criteria that ensure good model fits at low flows. The two methods were compared using a split sample test focussing on explained variance (R^2) and bias for the predicted Q_C.

This paper starts with a presentation of the streamflow and geographical data. Then the regression method and derivation of regional regression equations is described, followed by a presentation of the HBV model and a comparison of the two methods. Finally, the results are presented and discussed, and conclusions are drawn.

Study Region and Data Availability

The study region is the south-western part of Norway (Fig. 1). Precipitation in this region is mainly caused by depressions arriving from the south-west. Air masses are lifted when arriving at the mainland due to the presence of a mountain range. A maximum zone of precipitation is found 50-100 km from the coast, and on the leeward side of the mountains, the precipitation is lower. The measured average annual precipitation in the study region varies from 515 mm to 2800 mm (Førland, 1993). The average annual runoff varies from 10 ls^{-1}km^{-2} to 130 ls^{-1}km^{-2}. Close to the coast, monthly average temperatures are above 0 ºC, whereas in the mountains, six months of the year (November – April) have a monthly average temperature below 0ºC. The climatic differences lead to different hydrological regimes. In the inland and mountainous areas the low flow period is in the winter due to precipitation being stored as snow, whereas in the coastal lowlands the low flow period is in the summer due to increased evapotranspiration and slightly lower rainfall. The vegetation cover is mainly coniferous and deciduous forests in the low-land and grass and bushes in the mountains.

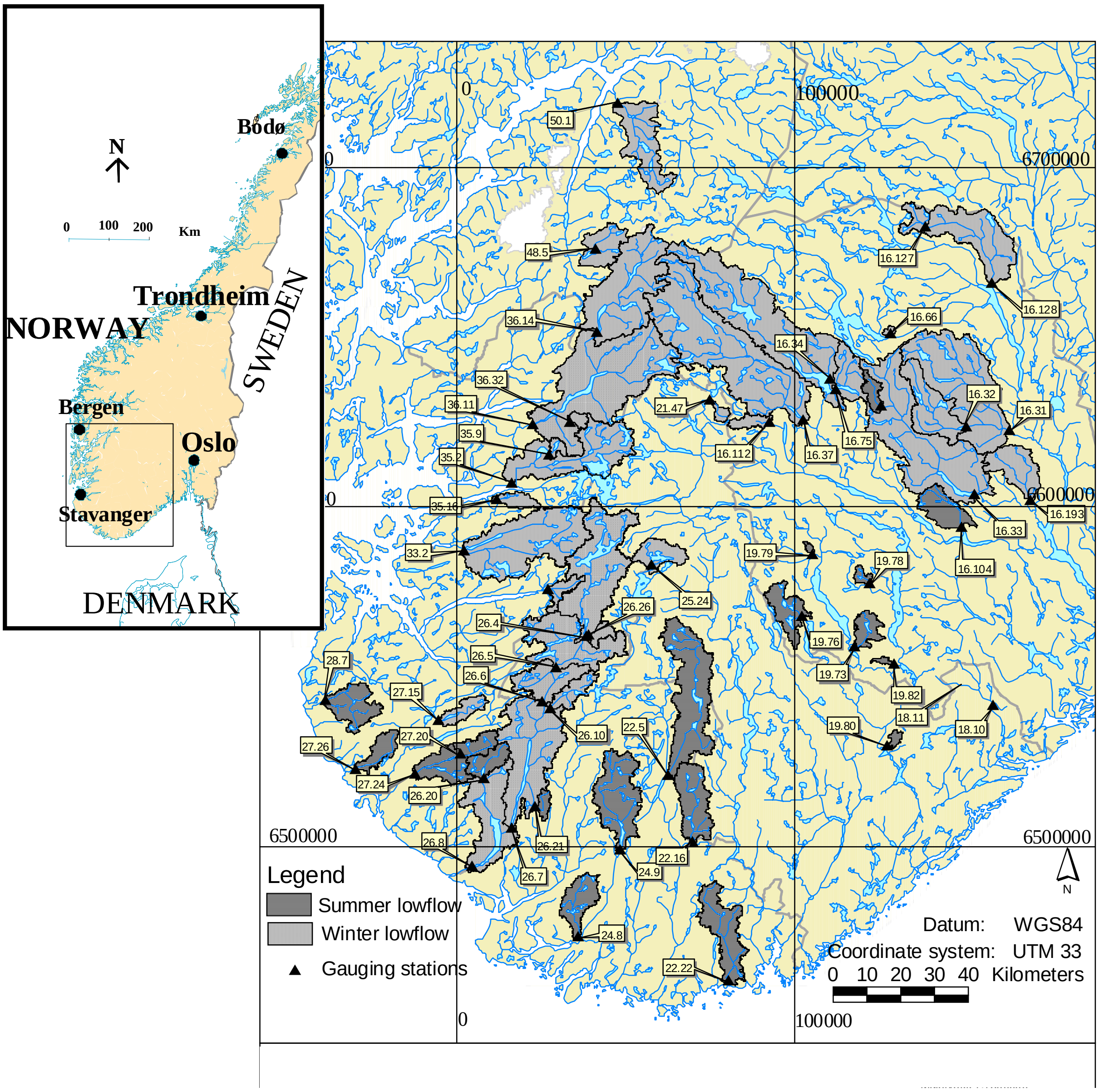

Fig. 1 Catchments and corresponding streamflow stations used in this study.

Agricultural and urban areas are of minor importance. The landscape also includes numerous lakes and mires that are of high importance for the hydrological response. Soils are mainly thin till deposits on bedrock and localised fluvial deposits in the valley bottoms.

Daily streamflow data were obtained for 51 stations with catchment areas less than 2000 km^2. The stations and their catchment boundaries are shown in Fig. 1. Table 1 lists the selected stations, record length, catchment area, Q_C, mean annual runoff and the dominant low flow season.

Tab. 1. Gauging stations used in the analysis. All stations were used to develop the regression equations. Stations marked with bold types were used to calibrate the HBV model and the regression parameters in a split-sample test of model performance. The remaining stations were used for validation.

Station	Period of measurements	Area (km^2)	Q_M* ($ls^{-1}km^{-2}$)	Q_C ($ls^{-1}km^{-2}$)	Low flow season
16.31 Omnesfoss	1921-1957	806	28.3	3.18	Winter
16.32 Hjartsjø	1919-1957	215	27.4	2.27	Winter
16.33 Seljordvatn	1912-1944	728	18.8	3.38	Winter
16.34 Totak	1895-1957	855	37.0	4.08	Winter
16.37 Vinjevatn	1919-1955	907	43.7	3.77	Winter
16.66 Grosettjern	**1949-2005**	**6.48**	**29.2**	**2.16**	**Winter**
16.75 Tannsvatn	**1955-2005**	**117**	**22.8**	**2.54**	**Winter**
16.104 Kilen	1962-2005	121	15.7	0.69x	Summer
16.112 Byrteåi	**1967-2005**	**37.3**	**50.2**	**1.55**	**Winter**
16.122 Grovåi	1972-2005	42.7	19.2	1.12x	Summer
16.127 Viertjern	**1977-2005**	**49.0**	**29.4**	**1.86**	**Winter**
16.128 Austbygdåi	**1976-2005**	**344**	**25.5**	**1.35**	**Winter**
16.193 Hørte	**1961-2005**	**156**	**15.5**	**2.24**	**Winter**
18.10 Gjerstad	1980-2005	237	25.1	0.62	Summer
19.73 Kilåi bru	**1968-2005**	**64.4**	**28.5**	**0.50**	**Summer**
19.76 Tovsliøytjønn	**1969-2002**	**115**	**32.8**	**2.67**	**Summer**
19.78 Grytå	1977-2005	18.7	24.2	1.76	Summer
19.79 Gravå	**1970-2005**	**6.31**	**22.1**	**0.32**	**Summer**
19.80 Stigvassåni	**1972-2005**	**14**	**27.4**	**0.43**	**Summer**
19.82 Rauåna	**1972-2005**	**8.93**	**23.9**	**0.34**	**Summer**
21.47 Lislefjødd	**1972-1995**	**19**	**35.8**	**1.32**	**Winter**
22.5 Austerhus	1922-1957	413	43.5	3.49	Summer
22.16 Myglevatn	**1951-2005**	**182**	**44.8**	**0.81**	**Summer**
22.22 Søgne	**1973-2005**	**210**	**29.9**	**1.45**	**Summer**
24.8 Møska	**1978-2005**	**121**	**50.2**	**2.50**	**Summer**
24.9 Tingvatn	**1922-2005**	**272**	**61.2**	**2.47**	**Summer**
25.24 Gjuvvatn	**1971-2005**	**97**	**65.4**	**7.41**	**Winter**
26.4 Fidjedalsvatn	1919-1969	506	80.7	5.42	Winter
26.5 Dorgefoss	1913-1969	808	76.4	4.53	Winter
26.6 Lindeland	**1913-1969**	**963**	**74.2**	**5.42**	**Winter**
26.7 Sirdalsvatn	1894-1964	1528	70.1	7.31	Winter
26.8 Lundevatn	1897-1964	1899	68.2	9.14	Winter
26.10 Liland	1933-1970	72.7	64.2	2.31	Winter
26.20 Årdal	**1970-2005**	**77.3**	**68.1**	**5.02**	**Summer**
26.21 Sandvatn	1970-2005	27.5	62.1	4.76	Summer
26.26 Jogla	**1973-2005**	**31.1**	**70.5**	**2.73**	**Winter**
27.15 Austrumdal	**1980-2005**	**60.5**	**95.8**	**11.42**	**Winter**
27.20 Gya	**1933-2005**	**60.7**	**97.1**	**4.51**	**Summer**

Tab. 1 continues

27.24 Helleland	1896-2005	186	79.5	9.85	Summer
27.26 Hetland	**1970-2005**	**69.5**	**58.5**	**3.15**	**Summer**
28.7 Haugland	**1918-2005**	**140**	**49.8**	**3.31**	**Summer**
31.2 Lysedalen	**1953-1984**	**47.2**	**90.5**	**11.33**	**Winter**
33.2 Tveid	1896-1956	513	88.9	10.69	Winter
35.2 Hauge bru	**1905-1980**	**394**	**87.0**	**5.88**	**Winter**
35.16 Djupadalsvatn	**1990-2005**	**45.4**	**70.48**	**5.84**	**Winter**
35.9 Osali	**1982-2005**	**22.6**	**86.6**	**4.78**	**Winter**
36.11 Stråpa	1904-1964	1307	73.5	5.33	Winter
36.14 Røldalsvatn	1913-1964	496	73.2	3.22	Winter
36.32 Lauvastøl	1985-2005	20.5	105.1	4.88	Winter
48.5 Reinsnosvatn	**1918-2004**	**121**	**76.50**	**4.34**	**Winter**
50.1 Hølen	**1923-2004**	**232**	**53.22**	**2.72**	**Winter**

* Q_M is mean annual runoff for the period 1961-1990 from Beldring *et al*. (2002).

We chose to model *Qc* normalized with respect to catchment area (units $ls^{-1}km^{-2}$). All other fluxes used in the regression equations (mean annual runoff, precipitation) were also specified in equivalent units (length / time).

The stations were selected according to their record length and the quality of low flow measurements. A minimum of 20 years with streamflow measurements, if possible covering the period 1960-2000, was required. A few stations do not have any observations within this period, and these are mainly catchments that have been heavily modified due to construction of reservoirs for hydropower production. In this study, only data predating hydropower regulation period are used. The second selection criterion was the low flow data quality. The streamflow is derived from measured river stage via the rating curve. The uncertainty in the rating curve for low flows is dependent on the number of flow measurements at low water levels, and on the shape and stability of the river profile. The quality of the rating curve was evaluated by a procedure based on a Bayesian estimation of credibility intervals around the annual minimum flow (Petersen-Øverleir *et al*., 2008). The relative uncertainty measured as the average ratio between width of the 95% credibility intervals and the estimated annual minimum flow, was used to classify the stations into five classes: very good (0-20%), good (20-40%), satisfactory (40-60%), bad (50-80%), very bad (>80%). The numbers in parentheses indicate the relative uncertainty. The stations classified as ‘very bad’ were excluded from the dataset. In addition to the rating curve evaluation, a subjective quality control was performed. Personal communication with field hydrologists provided information about small regulations not included in national databases, unstable profiles, problems with leaking V-notch weirs, and difficult ice conditions at the gauging stations.

In inland and high elevation areas with winter low flow, the quality of the low flow measurements depends on the ice conditions. Ice can cause the water level to rise without an increase in runoff. ‘Ice reduction’ procedures are carried out in order to reduce the increased streamflow and obtain more correct values. Ice can also influence the measurement instruments themselves. Winter low flow measurements might therefore be of poorer quality than summer low flow measurements.

The physiographic catchments descriptors were obtained from a GIS system. Table 2 lists physiographic together with climatic descriptors. All the land cover percentages were based on the national N50 maps (Scale 1:50 000). All the gradients were based on a digital elevation model with a resolution of 100x100 m. A digital river network was used to calculate the river gradients. The mean annual runoff Q_M was obtained from the runoff map of Norway (Beldring *et al*., 2002) for all locations. Observed values of Q_M were not used as the aim is to test the model performance at ungauged sites. The average precipitation P_A (annual), P_S (summer) and P_W (winter) as well as temperature T_A (annual), T_S (summer) and T_W (winter) were provided by the Norwegian Meteorological Institute. They were given as average values for the period 1961-1990 on a regular grid with a resolution of 1x1km. Catchment averages were estimated based on the gridded values.

Tab. 2. The catchment characteristics included in the regression analysis.

Symbol	Group	Description
A	1	Catchment area (km^2)
R_L	1	Length of main river (km) from the outlet to the most distant river string.
C_L	1	Catchment length (km) from outlet to the mots distant point at the water divide
C_W	1	Catchment width (km)
Q_M	2	Mean annual runoff (l/s km^2) from the runoff map of Norway (Beldring *et al.* (2002)
P_A	2	Annual precipitation (mm)
Ps	2	Summer precipitation (mm)
Pw	2	Winter precipitation (mm)
R_G	3	River gradient (m/km)
G_{1085}	3	River gradient excluding the 10 % lowest parts and the 15% highets parts1085 (m/km)
C_G	3	Catchment gradient (m/km)
D_H	3	Elevation gradient (m)
H_{max}	4	Maximum elevation (masl)
H_{min}	5	Minimum elevation (masl)
$U_{\%}$	6	Urbanised areas (%)
$A_{\%}$	7	Agricultural areas (%)
$F_{\%}$	8	Forested area (%)
$B_{\%}$	9	Bogs (%)
$M_{\%}$	10	Mountainious areas (%)
$L_{\%}$	11	Lake percentage (%)
L_{eff}	11	Effective lake percentage (%)
T_A	12	Average annual temperature (^oC)
Ts	12	Average summer temperature (^oC)
Tw	12	Average winter temperature (^oC)

Methods

Regional regression analysis

The regional regression analysis was performed in two steps. The first step was to divide the data into regions that can be regarded as homogeneous with respect to their low flow behaviour. In the second step the independent variables for the regression equations were selected for each region using a stepwise procedure.

Laaha and Blöschl (2006) investigated several catchment grouping strategies when developing regression equations to estimate low flow indices in Austria. The results showed that a catchment grouping based on seasonality gave the best performance. The reason is the large differences in low flow processes in Austria: winter low flows due to precipitation stored as snow and summer low flows due to high soil moisture deficit caused by evapotranspiration losses. Since the climate in Norway is similar, the catchments were divided into two groups: summer- and winter low flow catchments. The summer season was defined as May to October and the winter season as November to April. The average flows for the three winter months and the three summer months with the lowest streamflow were used to determine the dominant low flow season (Table 1). In ungauged catchments, it is necessary to use climate and geographical data for this classification. Climate statistics describing mean monthly, seasonal and annual temperatures and precipitation were compared to the hydrograph-based classification.

In the second step multiple linear regression was used to obtain relationships between the low flow index, Q_C, and catchment characteristics for the winter and summer regions separately. In total, 24 catchment characteristics (Tab. 2) were potential candidates for the regression equation. A stepwise procedure (Draper and Smith, 1998) was used to select the most important characteristics explaining low flows. Since the aim is model prediction at ungauged sites, a cross-validation test based on the cross-validated explained variance R^2_{CV} was used to select the independent variables:

$$R^2_{CV} = 1 - \frac{\sum_{i=1}^{n} \left(Q_{C,obs,i} - Q_{C,pred,i}\right)^2}{\sum_{i=1}^{n} \left(Q_{C,obs,i} - \overline{Q}_{C,obs,i}\right)^2} \qquad (2)$$

where i is the index for the observation site, n is the number of observations, $Q_{c,pred,i}$ the predicted *low flow index* at site i estimated by leaving the observed value $Q_{C,obs,i}$ at this site out from the parameter estimation. If the model fit is perfect, R^2_{CV} is equal to one, and very poor models can give slightly negative values. Note that the explained variance, R^2, has zero as minimum value, whereas for cross-validated values, R^2_{CV} can be negative. To evaluate R^2_{CV}, each site was successively left out in the estimation of the regression parameters. The Q_C was then predicted at the independent site. New variables were included if they increased R^2_{CV}. We also required that the regression coefficients should be significant at a 5% level. Independent variables that had a high correlation were pooled into groups (Tab. 2), and from each group the variable giving the highest R^2_{CV} was selected. In addition a subjective selection procedure was carried out to obtain more robust equations. The value of the regression coefficient should be reasonable, e.g. the regression coefficient for lake percentage should be positive since increasing lake percentage should lead to increasing Q_C.

As a part of the regression analysis, it is important to check whether the necessary assumptions of multiple regression are fulfilled:

- Homoscedasticity: does the variance of the residuals depend on the predicted value?
- Bias: does the bias of the residuals depend on the predicted value?

In order to perform a statistical inference it is also necessary to test if the residuals are normally distributed. Further, to obtain the best possible predictions, it is useful to check if the relationship between the dependent and independent variable is linear. In many cases a transformation of the independent variable can make the system more linear. In this study we allowed each variable to be either untransformed or log-transformed and select the one giving the highest R^2_{CV}.

We tested six alternative models (M1-M6) in order to investigate these requirements (Tab. 3). Several transformations of Q_C were tested in order to obtain homoscedasity and normally distributed residuals, and the log-transformation was found to be the best alternative. Q_C was therefore log-transformed in M2 - M5 before performing the linear regression. M2 was based on the regression equation developed by Væringstad and Hisdal (2005). In this paper, the same independent variables were used, but new regression coefficients were calculated as a slightly different set of streamflow records were included. All variables were log-transformed. Prior to the log-transformation 0.1 was added for the land cover variables and 10 for the temperatures. For the models M3 - M5 different transformations of the independent variables were considered in order to test the linearity requirement. To test the effect of dividing the data into one summer- and one winter region, a global regression equation was developed assuming that all data belonged to the same region (M6) using the same stepwise procedure as for M5. In order to check the regression requirements, diagnostic plots of observed versus predicted values and qq-plots of regression residuals versus normal quantiles were produced.

To evaluate the predictive capability of the model, cross validation tests were carried out for the summer and winter regions separately.

Tab. 3. The regression models.

Name	**Model**
M1	Untransformed variables.
M2	Model from Væringstad and Hisdal (2005).
M3	All variables log-transformed.
M4	Only the Q_C is log transformed.
M5	The Q_C is log-transformed, the model chooses between untransformed and log-transformed independent variables.
M6	Like M5, but the winter and summer regions are merged.

The HBV-model

A gridded version of the Norwegian HBV-model (Sælthun, 1996; Beldring *et al*., 2002; Beldring *et al*., 2003) was used. The model has previously been used to calculate a water balance map for Norway (Beldring *et al*., 2002), and to assess climate change impacts (Beldring *et al.*, 2008) and in combination with ecological modelling (L'Abée-Lund *et al*., 2004). The HBV-model operates on a daily time step. In this study, the model calculated the water balance for grid-cells of 1x1km. For each grid-cell the percentage of lake and glacier was determined in addition to the proportion of the two dominant out of five land use classes

(Tab. 4). Some of the model parameters were common for the whole region whereas others were determined for each land use class. The same process parameterisations were applied to all grid-cells (Fig. 2). The interception storage has to be filled up before the precipitation falls to the ground. Water evaporates at the potential rate from the interception storage. Sub-grid scale distribution of snow is accounted for in calculating the snow melt. The snowmelt or throughfall (in snow free areas) might either infiltrate into the soil moisture zone or percolate into the upper zone. The separation between infiltration and percolation is controlled by the soil moisture content (Fig. 2). From the soil moisture zone water evaporates from the snow free part of the area. The evaporation is reduced when the soil moisture is low. The upper zone generates runoff as a piecewise linear reservoir, but some water can percolate at a constant rate to the groundwater zone. The groundwater zone is a linear reservoir, but water can also be drawn up to the soil moisture zone when the soil moisture is low. The grid-cells are not connected through routing and the total catchment runoff is the sum of runoff from all of the individual cells. This does not introduce large errors in catchments with small lakes since for low flows the hillslope response and not the channel network response, will be the factor controlling catchment runoff response. However, if lakes are present in the river network, they can have an important influence. The estimated low flows from the HBV-model might therefore have been improved if the effects of lakes in the channel network were explicitly included in the model structure and not implicitly accounted for via the model calibration.

Tab. 4. The vegetation classes used in the GWB model.

No	**Description**
1	Areas above the tree line with sparse vegetation.
2	Areas above the tree line with grass, heather, shrubs or dwarfed trees.
3	Areas below the tree line with sub alpine forest.
4	Lowland areas with coniferous or deciduous forest.
5	Non-forested areas below the tree line.

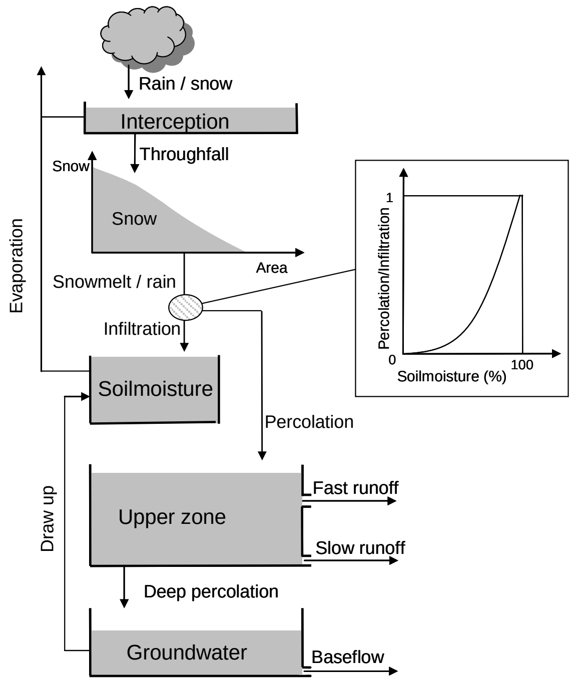

Fig. 2 The structure of the HBV model.

Daily precipitation and temperature observations were provided by the Norwegian Meteorological Institute. They were interpolated to each grid-cell using an inverse distance weighting routine with elevation correction to account for temperature and precipitation dependence on altitude. The temperature gradients were based on physical considerations. The precipitation gradients were calibrated according to the procedure described in Beldring *et al.* (2002). The gradients were between 8 % and 12 % per 100 meter up to 1200 m above sea level. For higher elevations the gradients were 4% to 6% per 100 meter. The gradients were defined for 29 points covering Norway, and for each grid cell a unique elevation gradient was obtained by an inverse distance weighting of the 3 closest of the 29 gradient points.

Calibration and validation

To evaluate the predictability of Q_C in ungauged catchments using the HBV-model, a split sample test was applied (e.g. Klemeš, 1986). The dataset was divided into two groups, daily streamflow observations from 30 stations were used for calibration and 21 as independent stations for validation (Tab. 1). Only stations with observations in the period 1961-1990 were selected. The software PEST (Doherty, 2004) was used for automatic calibration of the model. The HBV-model was calibrated using the average root means square error for daily runoff values measured in mm for selected catchments all over Norway (Beldring *et al*, 2002). This calibration, referred to as the first calibration, places a relatively high weight on higher streamflow values. Therefore another calibration, referred to as the second calibration, was performed using the average Nash-Sutcliffe coefficient R_{eff} for log-transformed streamflow as a calibration criterion.

$$R_{eff} = \frac{1}{m}\sum_{j=1}^{m}\left[1-\frac{\sum_{i=1}^{n}\left(q_{obs,i,j}-q_{sim,i,j}\right)^2}{\sum_{i=1}^{n}\left(q_{obs,i,j}-\bar{q}_{obs,j}\right)^2}\right]; \quad q=\ln(Q) \qquad (3)$$

where i is an index for time, j is an index for catchment, n is number of time steps, m is number of catchments, and q_{obs} and q_{pred} is the log-transformed observed and simulated streamflow. This criterion was applied to obtain a better fit at the lowest flows. To reduce the number of parameters for calibration, the parameters were not calibrated for each land use class. Instead a common calibration factor was applied. For example, for calibration of the evaporation parameter, a factor was calibrated with which the evaporation parameter for each individual class was multiplied.

The Q_C was calculated both for the calibration and the validation catchments and compared to observed values. The explained variance R^2 and bias were calculated both for the calibration and the validation sets.

In order to compare the prediction of Q_C using the regression method and the HBV-model in a proper way, a split sample test was performed also for the regression method. The same 30 catchments were used to estimate the coefficients in the best regression model established by stepwise regression. The regression equations were established separately for the summer and the winter low flow regions. The estimated coefficients were then used to obtain the predicted Q_C in the 20 independent catchments. The explained variance and bias were calculated both for the calibration and the validation set.

Results and discussion

Regression model

A classification rule was sought and climate statistics describing mean monthly, seasonal and annual temperatures and precipitation were compared to the hydrograph-based classification. Among these variables, the average July temperature performed best in reproducing the initial classification. If this temperature is higher than 10.4 ^{o}C the catchment has summer low flow. Fig. 3 shows how the July temperature 10.4 ^{o}C divides the catchments into two groups. Using this criterion only two stations were not classified according to the hydrograph-based classification. Station 16.122 Grovåi shifted from the summer to the winter region. Inspection of the hydrograph showed that this station has a mixed regime with low flow periods during both summer and winter. The error would therefore not have been large, if it had been included in the winter region. Station 16.193 Hørte shifted from the winter to the summer region. Further inspection of the summer stations indicated that this station has the lowest ratio of winter precipitation divided by summer precipitation (Fig. 3), i.e. summer precipitation is very high compared to winter precipitation. Inspection of the hydrograph indicated that this station has a mixed regime with winter as the most pronounced low flow period. For the final classification, it was decided to assign catchments with average July temperature higher than 10.4 °C and the ratio between winter- and summer precipitation larger than 0.65 to the summer low flow region, and the others to the winter low flow region (Fig. 3). In total, 32 catchments were assigned to the winter region and 19 to the summer region.

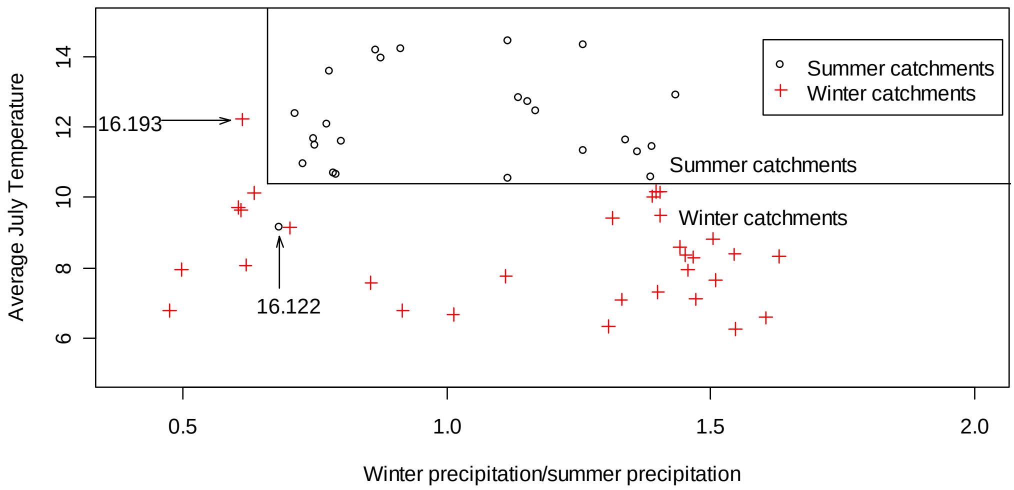

Fig. 3 Classification of summer and winter catchments. The circles and crosses indicate the summer and winter catchments, respectively, according to the initial classification, whereas the lines indicate the limits according to the classification based on climatic conditions.

We assume that for Norway the processes controlling low flows are snow cover formation and evapotranspiration. The analysis described above, shows that a temperature index is the best way to determine the dominating low flow process. A high temperature indicates that the snow-covered period is short and that evapotranspiration is high. These catchments will therefore have summer low flows. A low temperature implies low evapotranspiration, a long period with precipitation being stored as snow, and the winter as the dominating low flow period. Any temperature index will describe the importance of snow cover and evapotranspiration, and the average July temperature was the best for reproducing the initial classification. One possible explanation for why the July temperature was chosen among all the temperature indices is the difference in how low flows respond to temperature during the winter and summer seasons. Winter low flows are controlled by a threshold temperature, 0^oC, and it is of no importance how far below this threshold the temperature is. For summer low flows, however, it is reasonable to assume that the magnitude of the temperature is also important since higher temperature leads to higher evapotranspiration losses. It is interesting to note that the average July temperature is the temperature index that has the lowest correlation with the winter temperature indices (0.51 on average).

The estimated regression coefficients for the different regression models are shown in Tab. 5. For all models, except M6, separate regression equations were established for the summer- and the winter regions. Tab. 6 lists the results of the cross validation test using equation (2). The values of R^2_{CV} are shown for both Q_C and $\ln(Q_C)$. Note that exactly the same regression coefficients were used to evaluate R^2_{CV} for Q_C and $\ln(Q_C)$. The only difference is the log-transformation of the observed and predicted values. R^2_{CV} was calculated for the summer and winter region separately and for all observations and all predictions pooled into one region.

Fig. 4 shows diagnostics for the fit of M5 that gave the best results according to R^2_{CV}. There are four plots for the summer catchments and four plots for the winter catchments. In the upper plots Q_C is untransformed, whereas the lower plots show the results for log-transformed Q_C. Note that exactly the same regression coefficients were used to evaluate Q_C and $\ln(Q_C)$. The only difference lies in the log-transformation of the observed and predicted values.

The first plot shows predicted versus observed Q_C for the summer region. A good model fit is achieved if the points are close to the diagonal line. The second plot is a qq-plot for the residuals versus standard normal quantiles for the summer region. For normally distributed residuals, the points should lie on the diagonal line. The third and fourth plots show the same results, but for the winter region.

Tab. 5. The estimated regression coefficients. Q_c is given in ($ls^{-1}km^{-2}$).

Model	Equation
M1-Winter	$Q_c = 6.289 + 0.0484 Q_M - 0{,}0312 R_L + 0.873 T_W$
M1-Summer	$Q_c = -0.802 + 0.0766 Q_M + 0.115 L_{\%} - 0.192 B_{\%}$
M2-Winter	$Q_c = \exp\left[-0.570 + 0.770 \ln(Q_M) - 0.202 \ln(H_{\min})\right]$
M2-Summer	$Q_c = \exp\begin{bmatrix} -2.080 + 1.166 \cdot \ln(Q_M) - 0.534 \ln(R_G) \\ -0.368 \ln(B_{\%} + 0.1) + 0.153 \ln(M_{\%} + 0.1) \end{bmatrix}$
M3-Winter	$Q_c = \exp\begin{bmatrix} -6.387 + 0.835 \ln(Q_M) - 0.391 \ln(M_{\%} + 0.1) \\ -0.175 \ln(F_{\%} + 0.1) + 0.274 \ln(C_L) + 2.350 \ln(T_A + 10) \end{bmatrix}$
M3-Summer	$Q_c = \exp\begin{bmatrix} -4.288 + 1.282 \ln(Q_M) + 0.379 \ln(L_{\%} + 0.1) \\ -0.272 \ln(B_{\%} + 0.1) \end{bmatrix}$
M4-Winter	$Q_c = \exp\begin{bmatrix} -0.00758 + 0.00767 C_L + 0.0330 L_{\%} \\ +0.0204 Q_M - 0.00609 M_{\%} + 0.0894 T_A \end{bmatrix}$
M4-Summer	$Q_c = \exp\begin{bmatrix} -2.983 + 0.00285 P_S + 0.0976 L_{\%} \\ +0.0116 M_{\%} + 0.0150 C_L \end{bmatrix}$
M5-Winter	$Q_c = \exp\begin{bmatrix} -3.3325 + 0.0102 C_L + 0.03298 \ln(L_{eff}) + 0.026485 Q_M \\ +1.601 \ln(T_S) - 0.215 \ln(F_{\%} + 0.1) - 0.0173 M_{\%} \end{bmatrix}$
M5-Summer	$Q_c = \exp\begin{bmatrix} -4.734 + 1.301 \ln(Q_M) - 0.448 \ln(B_{\%} + 0.1) \\ +0.102 L_{\%} + 0.0130 C_L \end{bmatrix}$
M6-All	$Q_c = \exp\left[-1.355 + 0.0238(Q_M) + 0.380 \ln(C_W) + 0.243 \ln(L_{\%} + 0.1)\right]$

Tab. 6 The cross validated R^2_{CV} for Q_C and $\ln(Q_C)$ (in brackets) Q_c is given in ($ls^{-1}km^{-2}$).

Model	R^2_{CV} Summer	R^2_{CV} Winter	R^2_{CV} All catchments
M1	0.447 [-]	0.565 [-]	0.587 [-]
M2	0.467 [0.689]	0.480 [0.561]	0.537 [0.712]
M3	0.659 [0.792]	0.667 [0.695]	0.703 [0.803]
M4	0.676 [0.845]	0.607 [0.696]	0.670 [0.829]
M5	0.757 [0.820]	0.711 [0.816]	0.755 [0.855]
M6	-	-	0.520 [0.692]

The model that gave the best fit according to the R^2_{CV}, is Model 5. For this model Q_C was log-transformed before the regression equation was established, and the stepwise procedure chose between untransformed and log-transformed independent variables.

From Tab. 6 we clearly see that the R^2_{CV} is much smaller for Model 6 than for Model 5, which distinguishes summer and winter regions. We therefore concluded that two regions should be used.

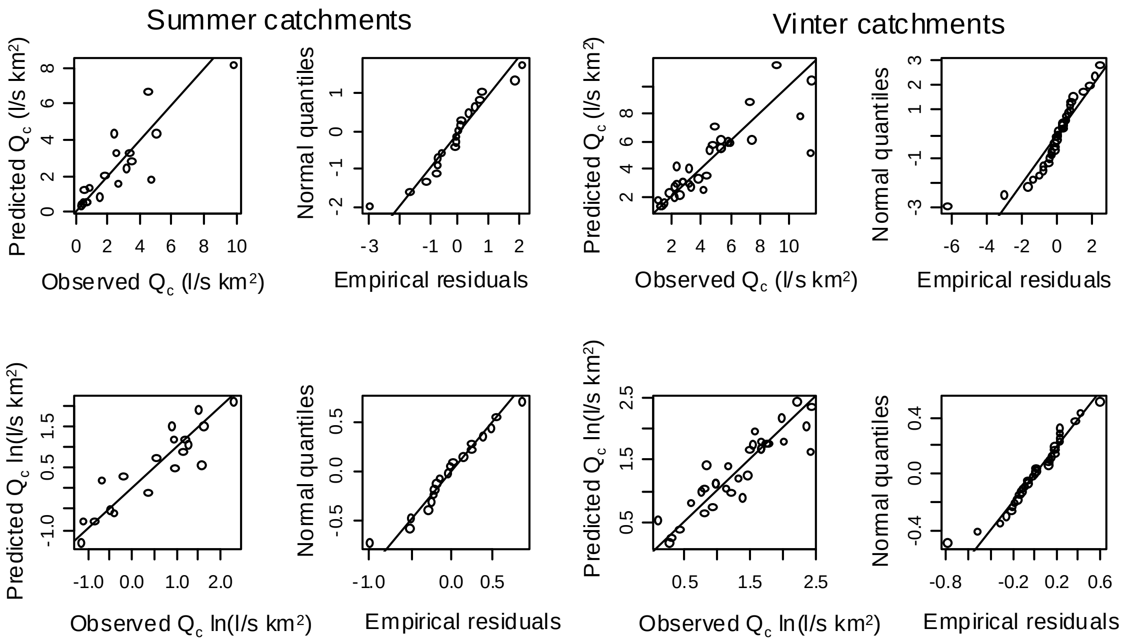

Fig. 4 Cross-validation of Model 5. The upper plots show the results as untransformed Q_C ($ls^{-1}km^{-2}$), whereas the lower plots show the results with Q_C log-transformed.

The qq-plots of the residuals indicate that the residuals from the log-transformed Q_C are close to normally distributed. For untransformed residuals, the normal distribution did not fit so well. We also see that for untransformed values, the estimation error depends on the predicted value. We therefore concluded that Q_C should be log-transformed to obtain normally distributed and homoscedastic residuals.

The bias of the residuals was centred on zero, but for many of the models the highest low flow values were underestimated. For a few of the models the lowest low flow values were overestimated. This was seen both for the log-transformed and the re-transformed low flow index.

The results show that better predictions were obtained for the summer region than for the winter region. A likely reason, apart from that the regression model performs worse, is that in Norway the low flow data are more uncertain in the winter than in the summer. During winter instruments may freeze up, or the low flows may have been estimated flow values based on an ice reduction procedure.

The different models include different combinations of independent variables, but some common features are seen. In all models either average runoff or summer precipitation was included. Q_C increased with increasing average runoff. Q_C also increased with lake percentage (included in 7 of the 11 equations). Bogs had the opposite effect of lakes. Increasing bog percentage gave decreasing Q_C. This is consistent with previous papers showing that the base flow from peat land is relatively small compared to other soils (e.g. Bragg, 2002; Bullock and Acreman, 2003; Evans *et al*., 1999, Shantz and Price, 2006). For the winter region Q_C increased with increasing temperatures. This is reasonable for winter catchments where snow accumulation and snow melt, highly influenced by temperature, control the magnitude of the low flow. The temperature was selected in 4 of 5 equations. Q_C increased with catchment length or width. This indicates that larger catchments have larger Q_C. Note that since we model Q_C in l / s km^2, this is an effect independent of Q_M that is given in the same units. Catchment geometry was included in 5 of 11 equations.

It is difficult to explain why Q_C decreased with increasing forest cover in M5 for the winter region. It might be that forest cover increases the evaporation from intercepted snow. It is, however, also possible that the correlation between the independent variables have given regression coefficients that can produce misleading interpretations, and they should be used with care. The correlation between TS and F% is 0.50, between T_S and $M_{\%}$ -0.42 and between $F_{\%}$ and $M_{\%}$ -0.70.

The HBV model

The results of the first calibration show that the Q_C is overestimated with a bias of 1.21 $ls^{-1}km^{-2}$ and an explained variance of only 0.29 (Fig. 5b). In the second calibration the bias is closer to zero (-0.05 $ls^{-1}km^{-2}$) and the explained variance increases to 0.59 (Fig. 5c). Hence, the calibration of the HBV-model with more weight on the low flow values improved the results. For the validation catchments, however, the first calibration gives better results than the second calibration if one considers both bias (0.07 $ls^{-1}km^{-2}$ in the first calibration and -0.99 $ls^{-1}km^{-2}$ in the second calibration) and explained variance (0.45 in the first calibration and 0.32 in the second calibration). If the three highest observed values are excluded from the validation set, the second calibration performs the best.

The use of the HBV-model to calculate Q_C demands high performance during the recession period. The regression method shows that it is very likely that, in addition to climatic descriptors, lakes and bogs are important landscape characteristics controlling the low flow. In this version of the HBV-model, individual lakes were not included as explicit elements in the model, and the bogs were not included in the model parameterisation at all. Better results might have been obtained with an improved interpolation of precipitation, an improved representation of lake elements and the introduction of soil and land use classes that are important for the recession.

Comparison between regression model and HBV model

Model M5, which gave the best fit according to the R^2_{CV} , was used in the split sample test for comparison with the results from the HBV model. The same independent variables as in M5 were used to re-estimate the regression coefficients using data from the 31 calibration catchment, and Q_C was predicted for the 21 independent catchments using the re-estimated regression coefficients. Fig. 5 shows the observed, HBV-estimated, and regression-estimated Q_C for the calibration and validation catchments.

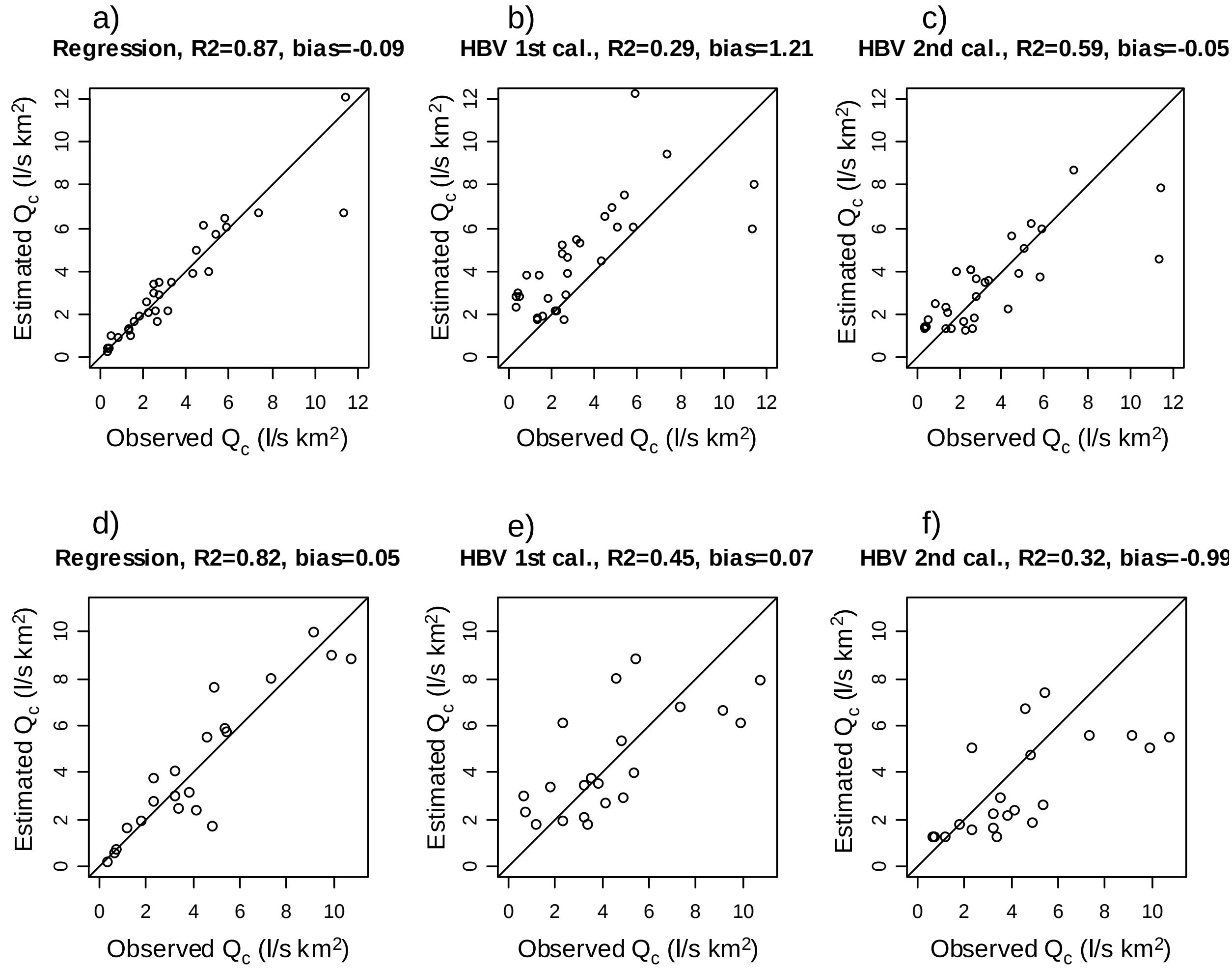

Fig. 5 The observed and simulated Q_C ($ls^{-1}km^{-2}$) for the calibration catchments for a) the regression method; b) the HBV-model, first calibration; c) the HBV-model, second calibration with high weights on low streamflow values, and for the validation catchments for d) the regression method; e) the HBV-model, first calibration; f) the HBV-model, second calibration with high weights on low streamflow values.

The regression method in general gave better prediction of Q_C in ungauged catchments than did the HBV-model. The regression method was especially superior to the HBV-model for the lowest Q_C -values. The predictive power for Q_C -values less than 2-3 $ls^{-1}km^{-2}$ for the HBV-model was rather limited.

The rainfall-runoff models introduce some additional intermediate steps, and thus some additional error sources, in the regional estimates of streamflow statistics. Most importantly, the structure of the models is constructed to describe all important hydrological processes, as opposed to the regression model that only contains explicit equations for the statistics of interest. Secondly, the hydrological model is calibrated to fit the hydrograph whereas the regression model is calibrated to fit the statistics of interest. The hydrologic model provides a lot of information that is actually not needed, and the principle of parsimony tells us that the simplest model should be preferred.

The error in the regression estimated Q_C is dependent on the predicted level. For log-transformed regression, the errors are proportional to the magnitude of streamflow. This means that the absolute error was small for lower predicted values of Q_C and larger for higher predicted values of Q_C. This was not the case when using the HBV-model. The absolute error seemed to be independent of the magnitude of Q_C, implying that the precision of the lowest Q_C predictions was low.

None of the methods accounted for the correlation in Q_C along rivers as few nested catchments were included in the dataset. If measurements are available from the same river, these should be used to obtain estimates of Q_C., e.g. by interpolation (Laaha et al, 2007).

Conclusions

The motivation for this study is the need for an objective method to estimate a low flow index at an ungauged site in order to improve operational procedures in water resources management. We have tested and compared two objective methods, a multiple regression analysis and a rainfall-runoff model. The regression method was used to establish a relationship between catchment characteristics and the low flow index. The rainfall-runoff model was used to simulate daily runoff series in ungauged catchments and to calculate the low flow index from these time series. The calculations were performed for Q_C, but the results are of general interest since Q_C is closely related to Q_{95}, a widely applied low flow index. Based on the results, the following conclusions can be drawn:

- The regression method gave better estimates of Q_C in ungauged catchments than did the HBV-model, especially for low values of Q_C.
- For the regression method, a catchment grouping based on the dominant low flow season was an effective method for obtaining homogeneous sub-regions in Norway where winter low flow and summer low flow are controlled by different processes. The average July temperature was the best index for determining the low flow season for ungauged catchments.
- Important catchment characteristics controlling low flows in the region were found to be average runoff, lakes, bogs, catchment area and average air temperature.
- For the regression method, the best results in this study were obtained when the Q_C was log-transformed and some of the independent variables were log-transformed.

It should be emphasized that these conclusions are based on a dataset characterized by many relatively small catchments, most of them unnested, in a landscape with many small lakes. Different results might have been obtained in a region with a denser gauging network and different landscape characteristics. _E.g. in partially gauged catchments, the interpolation method should be considered.

As an extension of this study, regression equations have now been developed for the whole of Norway (Engeland *et al.*, 2008) for estimating low flows. In order to deliver this method as a standard tool, GIS-based software has been developed to automatically retrieve catchment boundaries and required physiographic and climatic catchment characteristics upstream from a user-selected point in a river (Voksø *et al.*, 2008). The regression equations are implemented in an interactive web-application, the Norwegian low flow map, and the low flow index is calculated.

Refrerences

Beldring S, Engen-Skaugen T, Førland EJ, Roald LA (2008) Climate change impacts on hydrological processes in Norway based on two methods for transferring regional climate model results to meteorological station sites. Tellus 60A (3): 439-450.

Beldring S, Engeland K, Roald LA, Sælthun NR, Voksø A (2003) Estimation of parameters in a distributed precipitation-runoff model for Norway. Hydrol Earth Syst Sc 7(3): 304-316.

Beldring S, Roald LA, Voksø A (2002) Runoff map for Norway, average annual runoff 1961-1990 (in Norwegian), NVE Dokument no. 2/2002, Norwegian Water Resources and Energy Directorate, Oslo, Norway.

Bragg, OM (2002) Hydrology of peat-forming wetlands in Scotland, The Science of the Total Environment 294: 111-129.

Bullock A, Acreman M (2003) The role of wetlands in the hydrological cycle. Hydrol Earth Syst Sc 7(3): 358-389.

Clausen B, Rasmussen KR (1993) Low flow estimation and hydrogeology in a chalk catchment. Nordic Hydrology 24 (5): 297–308.

David H A, Nagaraja HN (2003) Order Statistics (3rd Edition). Wiley, New Jersey, pp 458.

Demuth S, Young AR (2004) Regionalisation Procedures. In: Tallaksen LM, van Lanen HAJ (Eds), Hydrological Drought Processes and Estimation Methods for Streamflow and Groundwater. Developments in Water Sciences vol 48. Elsevier Science Publisher, the Netherlands: 307–343.

Doherty J (2004) PEST Model-Independent Parameter Estimation. User Manual: 5th Edition. Watermark Numerical Computing (http://www.simulistics.com/files/documents/pestman.pdf - accessed December 1 2008)

Draper NR, Smith H (1998) Applied regression analysis, 3rd ed, Wiley, New York.

Engeland K, Hisdal H, Orte NK, Petersen-Øverleir A, Voksø A (2008) A low flow map for Norway (in Norwegian), Oppdragsrapport A nr 5/2008 The Norwegian Water Resources and Energy Directorate, Oslo, Norway.

Engeland K, Braud I, Gottschalk L, Leblois E (2006) Multi-objective regional modelling, Journal of hydrology, 327 (3-4): 339-351.

Evans MG, Burt TP, Holden J, Adamson JK (1999) Runoff generation and water table fluctuations in blanket peat: evidence from UK data spanning the dry summer of 1995. J Hydrol 221: 141-160.

Førland EJ (1993) Precipitation normals, standard period 1961-1990 (in Norwegian), DNMI-rapport nr. 39/93 Klima, The Norwegian Meteorological Institute, Oslo, Norway.

Gottschalk L, Krasovskaia I, Leblois E, Sauquet E (2006) Mapping mean and variance of runoff in a river basin. Hydrol Earth Syst Sc 10 (4): 469-484.

Gustard A, Young AR, Rees G, Holmes MGR (2004) Operational Hydrology, In: Tallaksen LM, van Lanen HAJ (Eds), Hydrological Drought Processes and Estimation Methods for Streamflow and Groundwater. Developments in Water Sciences vol 48. Elsevier Science Publisher, the Netherlands: 455-484.

Hisdal H, Tallaksen LM, Clausen B, Peters E, Gustard A (2004) Hydrological drought characteristics, In: Tallaksen LM, van Lanen HAJ (Eds), Hydrological Drought Processes and Estimation Methods for Streamflow and Groundwater. Developments in Water Sciences vol 48. Elsevier Science Publisher, the Netherlands: 139-198.

Institute of hydrology (1980) Low flow studies report, Wallingford UK.

Krokli B 1988. Analysis of low streamflows (in Norwegian), NVE Publication No. V14, Oslo, Norway.

Laaha G, Blöschl G (2007) A national low flow estimation procedure for Austria, Hydrological Sciences Journal, 52(4): 635-644.

Laaha G, Blöschl G (2006) A comparison of low flow regionalisation methods—catchment grouping. J Hydrol 323: 193-214.

Laaha G, Skoien J, Blöschl G (2007) A comparison of Top-kriging and regional regression for low flow regionalisation. Geophysical Research Abstracts 9, 07015, SRef-ID: 1607-7962/gra/EGU2007-A-07015

L'Abée-Lund JH, Vøllestad LA, Beldring S (2004) Spatial and temporal variation in the grilse proportion of Atlantic salmon in Norwegian rivers. Transactions of the American Fisheries Society 133: 743-761.

Lanen HAJ van, Clausen B, Kasparek, L (1993) Interaction between low flows and hydrogeology. In: Gustard A (Ed.), Flow Regimes from International Experimental and Network data (FRIEND), vol. 1. Hydrological Studies, Wallingford: 21–56.

Leith, RM (1978) Streamflow Regionalization in British Columbia, No. 4: Regression of Low Flows on Physiographic Parameters. Report Series No. 57, 26 pp.

Nathan RJ, McMahon TA (1990) Identification of homogeneous regions for the purpose of regionalisation, Journal of Hydrology, 121: 217-238.

Petersen-Øverleir A, Soot A, Reitan T (2008) Bayesian rating curve inference as a streamflow data quality assessment tool. Wat Res Manag. DOI: 10.1007/s11269-008-9354-5

Ries KG, (2002) STREAMSTATS: A U.S. Geological Survey web site for stream information. In: Hydroinformatics 2002 (Proceedings of the Fifth International Conference on Hydroinformatics, Cardiff, UK). IWA Publishing.

Shantz MA, Price JS (2006) Characterization of surface storage and runoff patterns following peatland restoration, Quebec, Canada. Hydrological Processes 20: 3799-3184

Smakhtin VY, Watkins DA (1997) Low-flow estimation in South Africa. Water Research Commission Report No. 494/1/97 Pretoria, South Africa.

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

Skøien JO, Merz R, Blöschl G (2006) Top-kriging - geostatistics on stream networks.

Hydrol Earth Syst Sc 10(2): 277-287.

Sælthun NR (1996) The “Nordic” HBV model. Description and documentation of the model version developed for the project Climate Change and Energy Production. NVE publication no 7-1996, Norwegian Water Resources and Energy Directorate, Oslo, Norway.

Thomas, DM, Benson, MA (1970) Generalization of streamflow characteristics from drainage-basin characteristics. USGS Water Supply Paper 1975.

Voksø A, Orthe NK, Hisdal H, Engeland K (2008) Low flow index map for Norway – interaction using GIS-software and analysis. In: Sveinsson ÓGB, Gardarsson SM, Gunnlaugsdóttir S (Eds.) Northern Hydrology and its Global Role, XXV Nordic Hydrological Conference, NHP Report No. 50: 154-159.

Væringstad T, Hisdal H, Drageset TA (2005) Estimation of low flow indices at ungauged sites, a case study from Norway. In: Proceedings XXIII Nordic Hydrological Conference 2004, 8-12 August 2004, Tallinn, Estonia. NHP report 48: 61-70.

Young AR, Round CE, Gustard A (2000) Spatial and temporal variations in the occurence of low flow events in the UK, Hydrology and earth system Sciences 4(1), 35-45.