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

*Kolbjørn Engeland*
*Hege Hisdal*
*Stein Beldring*

1
2006

REPORT

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

Norges vassdrags- og energidirektorat
2006

## Rapport nr

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

**Utgitt av:** Norges vassdrags- og energidirektorat

**Redaktør:**

**Forfattere:** Kolbjørn Engeland, Hege Hisdal, Stein Beldring

**Trykk:** NVEs hustrykkeri
**Opplag:** 50
**Forsidefoto:** Grunnfoss i Verdalselva – vinterlavvann. Foto: Arnt Bjøru
Grunnfoss i Verdalselva –sommerlavvann. Foto: Beate Sæther

**Sammendrag:** Estimater av lavvannsindekser er grunnlaget for mange avgjørelser i vannressursforvaltningen. I denne rapporten sammenlignes metoder for å bestemme alminnelig lavvannføring i umålte felt. Resultatene viser at regional regresjon, med alminnelig lavvannføring som avhengig variabel og klima- og feltegenskaper som uavhengige variable, gir det beste resultatet. Det anbefales derfor at denne metoden legges til grunn for et kartbasert GIS-system for estimering av ulike lavvannsindekser for nedbørfelt uten målinger.

**Emneord:** Lavvann, umålte felt, alminnelig lavvannføring, regresjon, HBV-modellen

Norges vassdrags- og energidirektorat
Middelthunsgate 29
Postboks 5091 Majorstua
0301 OSLO

Telefon: 22 95 95 95
Telefaks: 22 95 90 00
Internett: www.nve.no

Februar 2006

# Contents

# Forord

For å kunne vurdere virkningen av sterkt reduserte vannføringer i forbindelse med inngrep i vassdrag, må man kjenne den naturlige lavvannføringen i vassdraget. Slik kunnskap er viktig i forbindelse med nye vassdragskonsesjoner for eksempel for småkraftverk og revisjon av vilkår i gamle konsesjoner. I første del av prosjektet "Lavvannskart" har man sett på hvilken metode som egner seg best for å estimere lavvannsindeksen "alminnelig lavvannføring" i nedbørfelt med få eller ingen observasjoner. Denne rapporten beskriver arbeidet og konkluderer med at regional regresjon gir de beste resultatene. Metoden bør derfor benyttes i et automatisk GIS-basert kartsystem for estimering av lavvannsindekser. Utviklingen av lavvannskart støttes og finansieres gjennom Småkraftverkprogrammet i NVE.

Oslo, februar 2006

Morten Johnsrud
avdelingsdirektør

Sverre Husebye
seksjonssjef

# Sammendrag

Estimater av lavvannsindekser i umålte felt er grunnlaget for mange avgjørelser i vannressursforvaltningen. I Norge har for eksempel økningen i forespørselen om konsesjon for små vannkraftverk økt etterspørselen etter lavvannsindeksen ”alminnelig lavvannføring” i umålte felt. Alminnelig lavvannsføring blir ofte brukt som et utgangspunkt for å bestemme restvannføring når en konsesjon er nødvendig og som restvannføring når en konsesjon ikke behøves. Også i tilknytning til det europeiske rammedirektive for vann er det aktuelt å bestemme typiske verdier for små vannføringer.

Det er nødvendig med en standard prosedyrer som gir det best mulig estimatet av lavvannsindekser i umålte felt. Målet met dette prosjektet var å sammenligne statistisk regresjon med en regional nedbør-avløpsmodell. En region i Sørvest-Norge ble valgt som testområde. Dette er et område med store nedbørsgradienter og et vidt spekter av avrenningsregimer, fra sommerlavvann ved kysten og i lavlandet til vinterlavvann i innlandet og i fjellet.

Det ble satt fokus på å velge ut dataserier med god kvalitet på lave vannføringsverdier. Vannføringen blir beregnet fra den observerte vannstanden ved hjelp av en vannføringskurve. Kvaliteten på vannføringskurven ble vurdert for lave vannføringer. I tillegg ble informasjon fra stasjonsansvarlige brukt for å få et best mulig datagrunnlag.

Regresjonsmetoden bestemmer en sammenheng mellom alminnelig lavvannføring og fysikalske feltkarakteristika. Det første skrittet var å bestemme homogene under-regioner. I Norge er lavvann bestemt av to forskjellige prosesser. Årsaken til vinterlavvann er at nedbøren blir lagret som snø. Årsaken til sommerlavvann er økt fordampning. To homogene under-regioner ble derfor etablert, bestemt av hvilken måned de laveste vannføringene forekommer. For å avgjøre om et umålt felt har sommer, eller vinterlavvann ble det funnet at midlere juli temperatur er det beste kriteriet. Individuelle regresjonsligninger ble etablert for de to under-regionene. En skrittvis prosedyre ble brukt for å velge ut det optimale antallet feltkarakteristika. Innsjøer, myrer og klima var de viktigste feltkarakteristika. Kryssvalidering ble brukt for å avgjøre hvor god modellen er i umålte felt. Resultatene viser at regresjonsmodellen er mer presis i sommer-regionen enn i vinter-regionen.

En griddet versjon av HBV-modellen ble brukt som en regional nedbør-avløpsmodell. En gridstørrelse på 1x1km$^2$ ble brukt og hver gridrute ble tildelt modellparametre avhengig av arealbruksklasser. Modellen har tidsoppløsning på 1 døgn. Interpolert nedbør og temperatur fra met.no ble brukt som inngangsvariable. Modellen ble kalibrert basert på et utvalg nedbørfelt og validert på uavhengige felt for å kunne si noe om hvor god modellen er i umålte felt. Det ble valgt et kalibreringskriterium som legger størst vekt på de lave vannføringene. Fordelen med å bruke en slik modell er at man kan beregne hvilken som helst lavvannsindeks på grunnlag av en modellsimulering. Ulempen er den økende modellkompleksiteten som kan øke usikkerheten i estimatene.

Sammenligningen av resultatene basert på regresjon og HBV-modellen viser at regresjon gir de mest presise estimatene, spesielt på lave verdier for alminnelig lavvannføring. Regresjon gir en relativ feil, mens HBV-modellen gir en absolutt feil. Usikkerheten i estimatene fra HBV-modellen er ± 4.3 l/s/km$^2$ (observasjonene er i området 0.3-11.4

l/s/km$^2$). Dette tilsvarer en usikkerhet fra ca 35 til flere hundre prosent. Usikkerheten i estimatene fra regresjonsmetoden er ca ±35 %, dvs. ca. 4.0 l/s/km$^2$ for maksimalverdien og ca. 0.11 l/s/km$^2$ for minimumsverdien. Fremstillingen av lavvannskart vil derfor baseres på regresjon.

# Summary

Estimation of low flow indices at ungauged sites forms a basis for many decisions in the water resources management. In Norway, the European Water Framework Directive (WFD) and construction of small hydro power plants, has increased the demand for estimating low flow indices in heavily modified watersheds and at ungauged sites. In Norway the construction of small hydro power plants requires estimation of ‘common low flow’ (clf) in small ungauged catchments. ’Common low flow’ is approximately the 0.956 quantile of the flow duration curve and is often used a starting point to set residual flow when a licence is needed , and used as the residual flow if a licence is not needed.

Some standard procedures that give the best possible estimation of low flow indices in ungauged catchments are needed. The aim of this study is to compare the regression method and application of a regional precipitation-runoff model for estimating common low flow in ungauged catchments in Norway. South-western Norway was chosen as a study region. This is an area with large gradients in precipitation and a wide range of runoff regimes, from summer low flow regimes at the coast and in the low lands to winter low flow regimes in the inland and in the mountains.

We have focussed on selecting data series with good quality for low streamflows. The streamflow is calculated from the observed water level using a rating curve. The quality of the rating curve was assessed for low streamflow values. In addition the hydrometrists performed a subjective quality control

The regression method aims to establish a relationship between the common low and some catchment characteristics. The first step was to establish homogeneous sub-regions. In Norway the low flow is governed by two different processes. Winter low flow is caused by precipitation being stored as snow, whereas summer low flow is caused by high evaporation losses. The homogeneous sub-regions were classified according to in which season the lowest flow takes place, winter or summer. For ungauged catchments a rule was established for deciding to which sub-region they belong, summer or winter low flow. Individual regression equations were established for each sub-region between common low flow and catchment characteristics. A step-wise procedure was used for selecting the optimal number of catchment descriptors. Lakes and bogs turned out to be important catchment descriptors in addition to climatic conditions. A cross-validation procedure was used for evaluating the predictive performance of the model in ungauged catchments. The predictive power was higher for the summer catchments than for the winter catchments.

A gridded version of the HBV model was applied as a regional precipitation-runoff model. A grid size of 1x1km$^2$ was used and each model element was assigned parameter values according to predefined land use classes. Interpolated precipitation and temperature from the Norwegian meteorological institute was used as input. The model was calibrated to a subset of the catchments and validated on independent catchments. The calibration criterion was selected to fit the low flow parts of the stream flow record. The gain in using this method is that you can calculate any runoff characteristics from one model. The loss is the increased model complexity that might increase the estimation uncertainties.

Comparison of the predicted clf by the regression method and by the HBV model shows that the regression method, on average, gives the best estimates, especially fro low clf values. The regression method gives a relative error whereas the HBV model gives an absolute error. The uncertainty in the estimates from the HBV model is ± 4.3 l/s/km$^2$ (the observations are in the interval 0.3-11.4 l/s/km$^2$). This corresponds to an uncertainty from about 35 up to several hundred percent. The uncertainty in the estimates from the regression method is about ±35 %, e.g. about 4.0 l/s/km$^2$ for the maximum value and 0.11 l/s/km$^2$ for the minimum value. The low flow map will therefore be based on regression.

# 1 Introduction

Estimation of low flow indices at ungauged sites is needed for many decisions in water resources management. In Norway an increasing demand for low flow data, especially for small catchments, is related to the increasing request to build small hydropower plants. Also related to other water management issues such as river pollution and ecological aspects, irrigation, reservoir design and management, drinking water supply and fish farming, there is a need to estimate low flow indices. The European Water Framework Directive (WFD) requires that all inland waters must reach a good status by 2015. It defines how this should be achieved through the establishment of environmental objectives and ecological targets for surface waters. Some minimum flow requirements will be important when implementing the WFD.

In Norway the construction of small hydropower plants requires estimation of 'common low flow' (clf) in small ungauged catchments. This is due to the Norwegian Water Resources Act § 10 which states (Lovdata, www):

- By water abstractions and flow regulations that change the streamflow in perennial rivers and streams, at least 'common low flow' has to remain in the river or stream.
- A licence is needed if a flow regulation or water use implies that the streamflow becomes less than the 'common low flow'.

'Common low flow', a low flow index, is often a starting point to set residual flow when a licence is needed , and it is often used as the residual flow if a licence is not needed.

Common low flow is defined as follows (preferably based on at least 15-20 years of data):

- Remove the 15 smallest values every year in a daily streamflow record
- Calculate the annual minimum series
- Rank the values in the annual minimum series and remove 1/3 of the smallest values.

The smallest value left is defined as the clf, and it is approximately the 0.956 quantile of the flow duration curve, i.e. the flow that is exceeded 95.6 percent of the time. This low flow index is used only in Norway. Low flow is controlled by climatic conditions, the storage capacity of the upstream catchment and other physiographic catchment characteristics. The low flow follows the main patterns of average precipitation and runoff. In addition the low flow is modified by local catchment characteristics, e.g. catchments with many lakes or large groundwater reservoirs will have a higher low flow. The clf is defined for all year. For inland and mountain catchments the low flow period will be during the winter caused by precipitation stored as snow. In lowland and coastal areas the low flow period is during the summer, mainly caused by higher evaporation losses. The clf calculated for a catchment with winter low flow, will therefore not be very useful for ecological purposes since it gives no information about natural low flow during the summer season when the river is ecological active. An alternative approach is then to calculate flow quantiles for the desired season.

A standard procedure that gives the best possible estimation of low flow indices in ungauged catchments is needed. At least five methods can be used.

- *Percentage of mean annual runoff:* The simplest method is to set the clf equal to 10% of the mean annual runoff. In Norway the mean annual runoff can be obtained from the runoff map of Norway (Beldring *et al.*, 2002). Væringstad and Hisdal (2005) show that this method is not very precise, and that alternative methods give better results.
- *Donor catchment:* The most common method is to choose a donor catchment. The low flow index in the target- and the donor catchments are assumed to be proportional to the catchment area. The donor catchment is a gauged catchment similar to the target catchment with regards to climatic conditions and catchment response. In Norway the climate is mainly controlled by elevation distribution and the catchment response is mainly controlled by lake percentage. This method works well for areas with a dense station network. In Norway where we have large local variations and the station network is rather sparse, it is often difficult to find a good donor station.
- *Multiple regressions:* Regression techniques aim to establish regression equations between low flow indices and catchment characteristics. An overview is given in Demuth and Young (2004). Norwegian studies are presented in Krokli (1988), Skaugen *et al.* (2002) and Væringstad and Hisdal (2005). In heterogeneous areas, it is necessary to establish individual regression equations for sub regions that are homogeneous with regards to the low flow generating processes (e.g. Laaha and Blöschl, 2006). In Norway a major difference is between catchments with dominating summer low flow and winter low flow (Væringstad and Hisdal, 2005). A cross-validation procedure is often used to evaluate the predictive performance of the model. The catchment characteristics considered will depend on the characteristics that are assumed to be of hydrological importance and the data-availability. Characteristics that describe geology, soils, land-use, topography, and climate are commonly used.
- *Precipitation-runoff modelling:* Regional precipitation-runoff modelling can also be used to estimate low flow indices in ungauged catchments. The deterministic model produces streamflow series from which the desired low flow indices can be calculated. This method requires good procedures to transfer model parameters from gauged to ungauged catchments and for interpolation of the meteorological input variables (temperature, precipitation, etc.) The gain in using this method is that you can calculate any low flow index from one model. The loss is the increased model complexity that might increase the estimation uncertainties.
- *Interpolation:* Applying interpolation it is assumed that low flow is smoothly varying, and that the proximity in space is more important than similarity in catchment attributes. It is also possible to use interpolation procedures that accounts for the river network structure (Gottschalk *et al.*, 2004). Interpolation procedures require a rather dense streamflow gauging network preferably in nested catchments with few lakes, and are therefore not considered appropriate for Norwegian conditions.

The aim of this study is to evaluate and compare regression- and precipitation-runoff modelling methods for estimation of clf in ungauged catchments in Norway. This is achieved by estimating regression equations between clf and catchments characteristics for 56 catchments in southern Norway. The performance of the regression equations was evaluated by a cross-validation procedure. A gridded version of the HBV-model was calibrated to a subset of the 56 catchments using objective criterions that give good fit to low flows. The calibration results were validated on an independent set of catchments. The two methods were compared according to the root mean square error, explained variance ($R^2$) and bias for the predicted low flow.

This report starts with a presentation of the streamflow and geographical data. Then the regression method is described and regional regression equations are derived. It is followed by a presentation of the HBV model and how the two methods were compared before the results are presented and some conclusions are drawn.

# 2 Data

## 2.1 Study region

The study region is the south-western part of Norway (Fig. 1). The precipitation is mainly produced by low pressures arriving from south-west. The air masses are lifted when arriving at the main land due to a mountain range. Large precipitation amounts are produced, and a maximum zone of precipitation is found 50-100 km from the coast. On the leeward side of the mountains, the precipitation is lower. The highest measured average annual precipitation in the study region is 2800 mm in Maudal, Rogaland and the smallest is 515 mm in Mår in Telemark, (Førland, 1993). The runoff varies from 10 $ls^{-1}km^{-2}$ to 130 $ls^{-1}km^{-2}$ in the southwest. Close to the coast, none of the months have average temperatures below 0 ºC, whereas in the mountains six months (November – April) have average temperatures below 0ºC. The climatic differences lead to different hydrological regimes. In the inland and the 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. Agricultural and urban areas are of minor importance. The landscape is covered by several lakes and mires that are of high importance for the hydrological response. The soils are mainly thin till deposits on bedrock. Fluvial deposits are mostly found in the valley bottoms.

## 2.2 Streamflow data

Daily streamflow data was obtained from 56 stations in the study area. The stations and their catchment boundaries are shown in Fig. 1. Table 1 lists the selected stations, record length, catchment area, clf, mean annual runoff and the dominant low flow season. Summer was defined as May to October and winter as November to April. The average flows for the three winter months and the three summer months with the lowest streamflow were used to find he dominant low flow period.

The stations were selected according to the record length and the quality of the low flow measurements. A minimum of 20 years with streamflow measurements, if possible covering the period 1960-2000, was required. For many stations, however, data previous to 1960 were used. This mainly concerns catchments that have been heavily modified due to construction of reservoirs for hydropower production. These historical streamflow data were included since it is assumed that the temporal variation in clf is much less important than the spatial variation. 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 in the low flow part amongst others depends on how many flow measurements that have been done 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 developed at NVE (Petersen-Øverleir, 2005). The procedure is based on a Bayesian estimation of credibility intervals around the annual minimum flow. 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, good, satisfactory, bad, very bad (Table 2). The stations classified as 'very bad' were excluded from the dataset. In addition the hydrometrists performed a subjective quality control.

In inland and high elevation areas were the low flow takes place during winter, the quality of the low flow measurements depends on the ice conditions. For many locations ice causes the water level to rice without an increase in runoff. ‘Ice reduction’ procedures are carried out in order to reduce the increased streamflow and obtain correct values. Ice might also influence the measurement instruments themselves.

Table 1. The gauging stations applied. Stations marked with bold types were used to obtain regression equations and to calibrate the HBV model.

| Station | Period of measurements | Area (km$^2$) | $Q_M$ obs (ls$^{-1}$km$^{-2}$) | $Q_M$* (ls$^{-1}$km$^{-2}$) | Common low flow (ls$^{-1}$km$^{-2}$) | Low flow season |
|---|---|---|---|---|---|---|
| 16.31 Omnesfoss | 1921-1957 | 806 | 29.8 | 28.3 | 3.18 | Winter |
| 16.32 Hjartsjø | 1919-1957 | 215 | 27.4 | 27.4 | 2.27 | Winter |
| 16.33 Seljordvatn | 1912-1944 | 728 | 18.8 | 18.8 | 3.38 | Winter |
| 16.34 Totak | 1895-1957 | 855 | 37.0 | 37.0 | 4.08 | Winter |
| 16.37 Vinjevatn | 1919-1955 | 907 | 43.7 | 43.7 | 3.77 | Winter |
| **16.66 Grosettjern** | **1949-dd.** | **6.48** | **19.8** | **29.2** | **2.16** | **Winter** |
| **16.75 Tannsvatn** | **1955-dd.** | **117** | **22.5** | **22.8** | **2.54** | **Winter** |
| 16.104 Kilen | 1962-dd. | 121 | 17.4 | 15.7 | 0.69 | Summer |
| **16.112 Byrteåi** | **1967-dd.** | **37.3** | **52.0** | **50.2** | **1.55** | **Winter** |
| **16.122 Grovåi** | **1972-dd.** | **42.7** | **25.8** | **19.2** | **1.12** | **Summer** |
| **16.127 Viertjern** | **1977-dd.** | **49.0** | **21.1** | **29.4** | **1.86** | **Winter** |
| **16.128 Austbygdåi** | **1976-dd.** | **344** | **21.8** | **25.5** | **1.35** | **Winter** |
| **16.193 Hørte** | **1961-dd.** | **156** | **29.9** | **15.5** | **2.24** | **Winter** |
| 18.10 Gjerstad | 1980-dd. | 237 | 29.3 | 25.1 | 0.62 | Summer |
| 18.11 Tjellingtjernbekk | 1981-dd. | 2.16 | 28.7 | 24.0 | 0.00 | Summer |
| **19.73 Kilåi bru** | **1968-dd.** | **64.4** | **27.5** | **28.5** | **0.50** | **Summer** |
| **19.76 Tovsliøytjønn** | **1969-2002** | **115** | **32.0** | **32.8** | **2.67** | **Summer** |
| 19.78 Grytå | 1977-dd. | 18.7 | 21.4 | 24.2 | 1.76 | Summer |
| **19.79 Gravå** | **1970-dd.** | **6.31** | **20.1** | **22.1** | **0.32** | **Summer** |
| **19.80 Stigvassåni** | **1972-dd.** | **14** | **26.3** | **27.4** | **0.43** | **Summer** |
| **19.82 Rauåna** | **1972-dd.** | **8.93** | **22.1** | **23.9** | **0.34** | **Summer** |
| **19.89 Skornetten** | **1973-2002** | **2.62** | **25.2** | **27.3** | **0.00** | **Summer** |
| **19.91 Åbogtjønn** | **1973-2002** | **1.15** | **27.0** | **31.0** | **0.00** | **Summer** |
| **19.96 Storgama** | **1974-dd.** | **0.52** | **32.7** | **39.0** | **0.00** | **Summer** |
| **20.11 Tveitdalen** | **1972-dd.** | **0.44** | **34.1** | **34.7** | **0.00** | **Summer** |
| **21.47 Lislefjødd** | **1972-1995** | **19** | **35.8** | **35.8** | **1.32** | **Winter** |
| 22.5 Austerhus | 1922-1957 | 413 | 43.6 | 43.5 | 3.49 | Summer |
| **22.16 Myglevatn** | **1951-dd.** | **182** | **44.9** | **44.8** | **0.81** | **Summer** |
| **22.22 Søgne** | **1973-dd.** | **210** | **29.9** | **29.9** | **1.45** | **Summer** |
| **24.8 Møska** | **1978-dd.** | **121** | **50.2** | **50.2** | **2.50** | **Summer** |
| **24.9 Tingvatn** | **1922-dd.** | **272** | **61.0** | **61.2** | **2.47** | **Summer** |
| **25.24 Gjuvvatn** | **1971-dd.** | **97** | **65.4** | **65.4** | **7.41** | **Winter** |

Table 2.continues.

| | | | | | | |
|---|---|---|---|---|---|---|
| 26.4 Fidjedalsvatn | 1919-1969 | 506 | 85.6 | 80.7 | 5.42 | Winter |
| 26.5 Dorgefoss | 1913-1969 | 808 | 76.4 | 76.4 | 4.53 | Winter |
| **26.6 Lindeland** | **1913-1969** | **963** | **74.2** | **74.2** | **5.42** | **Winter** |
| 26.7 Sirdalsvatn | 1894-1964 | 1528 | 70.1 | 70.1 | 7.31 | Winter |
| 26.8 Lundevatn | 1897-1964 | 1899 | 68.2 | 68.2 | 9.14 | Winter |
| 26.10 Liland | 1933-1970 | 72.7 | 64.2 | 64.2 | 2.31 | Winter |
| **26.20 Årdal** | **1970-dd.** | **77.3** | **67.0** | **68.1** | **5.02** | **Summer** |
| 26.21 Sandvatn | 1970-dd. | 27.5 | 61.8 | 62.1 | 4.76 | Summer |
| **26.26 Jogla** | **1973-dd.** | **31.1** | **70.4** | **70.5** | **2.73** | **Winter** |
| **27.15 Austrumdal** | **1980-dd.** | **60.5** | **95.8** | **95.8** | **11.42** | **Winter** |
| **27.20 Gya** | **1933-dd.** | **60.7** | **97.1** | **97.1** | **4.51** | **Summer** |
| 27.24 Helleland | 1896-dd. | 186 | 79.6 | 79.5 | 9.85 | Summer |
| **27.26 Hetland** | **1970-dd.** | **69.5** | **58.0** | **58.5** | **3.15** | **Summer** |
| **28.7 Haugland** | **1918-dd.** | **140** | **49.8** | **49.8** | **3.31** | **Summer** |
| **31.2 Lysedalen** | **1953-1984** | **47.2** | **90.5** | **90.5** | **11.33** | **Winter** |
| 33.2 Tveid | 1896-1956 | 513 | 88.9 | 88.9 | 10.69 | Winter |
| **35.2 Hauge bru** | **1905-1980** | **394** | **85.5** | **87.0** | **5.88** | **Winter** |
| **35.16 Djupadalsvatn** | **1990-dd.** | **45.4** | **65.0** | **70.48** | **5.59** | **Winter** |
| **35.9 Osali** | **1982-dd.** | **22.6** | **86.6** | **86.6** | **4.78** | **Winter** |
| 36.11 Stråpa | 1904-1964 | 1307 | 73.5 | 73.5 | 5.33 | Winter |
| 36.14 Røldalsvatn | 1913-1964 | 496 | 72.2 | 73.2 | 3.22 | Winter |
| **36.32 Lauvastøl** | **1985-dd.** | **20.5** | **105.1** | **105.1** | **4.88** | **Winter** |
| **48.5 Reinsnosvatn** | **1918-2004** | **121** | **73.72** | **76.50** | **4.34** | **Winter** |
| **50.1 Hølen** | **1923-2004** | **232** | **51.45** | **53.22** | **2.72** | **Winter** |

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

Table 2. Classification of data quality

| Width of credibility intervals | Quality class |
|---|---|
| 0-20% | Very good |
| 20-40% | Good |
| 40-60% | Satisfactory |
| 60-80% | Bad |
| >80% | Very bad |

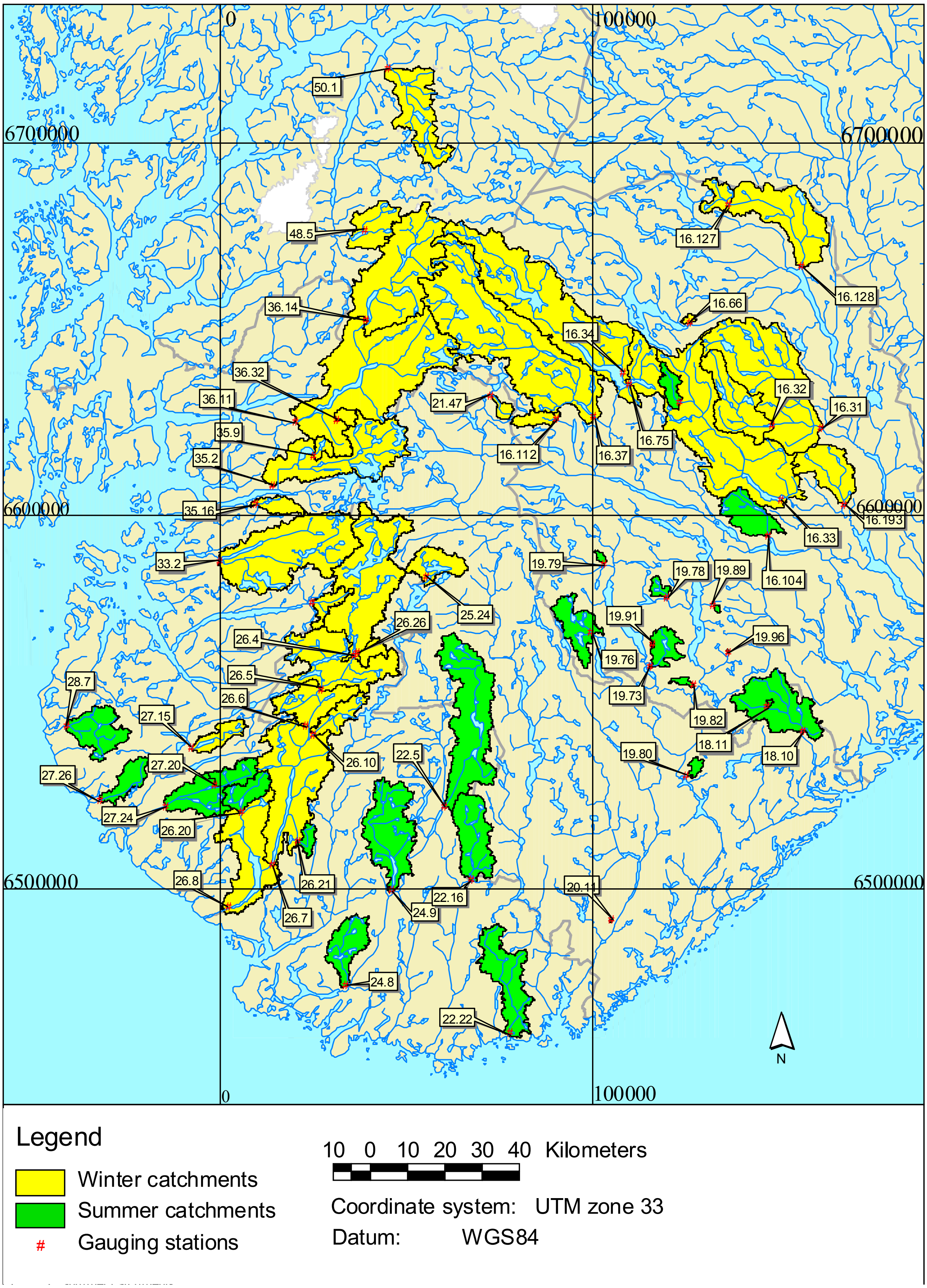

Figure 1. Catchments and corresponding streamflow stations used in this study.

## 2.3 Catchment characteristics

The physiographic catchments descriptors were obtain from the GIS system of NVE. Table 3 lists the descriptors together with climatic descriptors. All the land cover percentages are based on the N50 maps (Scale 1:50 000). The gradients are based on a digital elevation model with resolution 100x100 m. A digital river network was used to calculate the river gradients. The average precipitation $P_A$, $P_S$, and $P_W$ and temperature $T_A$, $T_S$, and $T_W$, was provided by the Norwegian meteorological institute. They were given as average values for the period 1961-1990 on a regular grid with resolution 1x1km and aggregated to catchment averages.

Table 3. The catchment characteristics included in the regression analysis:

| **Symbol** | **Description** |
|---|---|
| A | Catchment area (km$^2$) |
| $Q_M$ | Mean annual runoff (l/s km2) obtained from the runoff map of Norway (ref) |
| $R_L$ | Length of main river (km) from the outlet to the most distant river string. |
| $R_G$ | River gradient (m/km) |
| $G_{1085}$ | River gradient excluding the 10 % lowest parts and the 15% highets parts1085 (m/km) |
| $C_L$ | Catchment length (km) from outlet to the mots distant point at the water divide |
| $C_W$ | Catchment width (km) |
| $C_G$ | Catchment gradient (m/km) |
| $H_{max}$ | Maximum elevation (masl) |
| $H_{min}$ | Minimum elevation (masl) |
| $D_H$ | Elevation gradient (m) |
| $U_{\%}$ | Urbanised areas (%) |
| $A_{\%}$ | Agricultural areas (%) |
| $F_{\%}$ | Forested area (%) |
| $B_{\%}$ | Bogs (%) |
| $M_{\%}$ | Mountainious areas (%) |
| $L_{\%}$ | Lake percentage (%) |
| $L_{eff}$ | Effective lake percentage (%) |
| $T_A$ | Average annual temperature ($^o$C) |
| Ts | Average summer temperature ($^o$C) |
| Tw | Average winter temperature ($^o$C) |
| $P_A$ | Annual precipitation (mm) |
| Ps | Summer precipitation (mm) |
| Pw | Winter precipitation (mm) |

# 3 Regional regression analysis

The regional regression analysis was performed in two steps. The first step was to divide the data into groups that can be regarded as homogeneous with respect to their low flow behaviour. In Norway, the two important low flow classes are summer- and winter low flow. For estimation of clf in ungauged catchments, it is necessary to classify them as summer or winter catchment using climate and geographical data. The classification of the gauged catchments was therefore compared to average temperature and precipitation data available for the whole of Norway.

In the second step multiple linear regression was used to obtain relationships between clf and catchment characteristics for the winter and summer regions separately. In total, 24 catchment characteristics (Table 3) were potential candidates for the regression equation, and a stepwise procedure was used to select the most important characteristics that explain the low flow. The Akaikes information criterion (AIC) was used for the selection. AIC is calculated as: $\mathrm{AIC} = 2\mathrm{NLL}(\theta_i) + 2i$, where NNL is negative log likelihood and $\theta_i$ is the parameter vector containing $i$ elements. The first part describes the model fit to the data whereas the second part penalize for model complexity. The model with the smallest AIC is preferred. It was also required that the regression coefficients should be significantly different from zero on a 5% level.

An important part of a regression analysis is to check whether the necessary requirements are fulfilled in order to perform a statistical inference. Requirements to be controlled are:

1) Distribution: are the residuals (prediction - observation) normally distributed?
2) Homoscedasity: does the variance of the residuals depend on the predicted value?
3) Bias: are the residuals biased or does the bias depend on the predicted value or some covariate?
4) Constant variance: does the variance depend on some covariate?
5) Linearity: is the relationship between the dependent and independent variables linear or non-linear?

For many hydrological applications it is necessary to adjust for 1) and 2), commonly by applying a Box-Cox transformation of which the square root- and log-transformations are special cases. If non-linear relationships are found between dependent and independent variables, 5), the independent variables can be transformed in order to obtain a linear equation. 3) and 4) are indicators of problems with the model structure and if possible, alternative models should be investigated.

In order to evaluate the predictive capability of the model, cross validation testes were carried out. Each observation was successively left out in the estimation of the regression parameters. The clf was then predicted at the independent site. The explained variance for the prediction was then calculated.

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

The cross validation test was carried out for the summer and winter regions separately.

## 3.1 Classification

The catchments with observations were divided into two groups according to their dominating low flow period as describe in section 2.2. In order to decide whether an ungauged catchment has low flow during winter or summer, annual average and average seasonal variation in temperature and precipitation for the period 1960-1990 were used as indicators. The temperature and precipitation averages were available as grid maps with resolution 1x1 $km^2$. The average July temperature was found to be the best indicator to reproduce the initial classification. If this temperature is higher than 10.4 $^oC$ the catchment has summer low flow. Using this classification only two stations with observations are not classified according to a manual classification based on the lowest three month average streamflow (Fig. 2). The station 16.122 Grovåi was based on the manual procedure, classified as a summer station, but is now classified as a winter station. Inspection of the hydrograph shows that this station has a mixed regime with low flow periods during both summer and winter. The error is therefore not big if it is included in the winter group. Station 16.193 Hørte was originally classified as a winter stations, but is now classified as a summer station. Further inspection of the summer stations, indicates that this station has the smallest ratio of winter precipitation divided by summer precipitation (Fig. 2), i.e. summer precipitation is very high compared to winter precipitation. Inspection of the hydrograph indicates that this station has a mixed regime with the winter as the most pronounced low flow period.

The July temperature is strongly correlated to the other monthly averages, so the July temperature is therefore also and indicator for how important the snow cover is in the catchment. The July temperature is in addition a good indicator for the importance of evapotranspiration. A high July temperature indicates that the snow covered period is short and that the evapotranspiration is high. Such catchments will therefore have summer low flow. Low July temperatures indicate low evapotranspiration and a long period with precipitation being stored as snow. Thereby the winter is the dominating low flow period. The results presented in this report use classification based on July temperatures and the winter- summer precipitation ratio. This classification is only valid for this region, and it might be necessary to use another classification strategy for other regions in Norway.

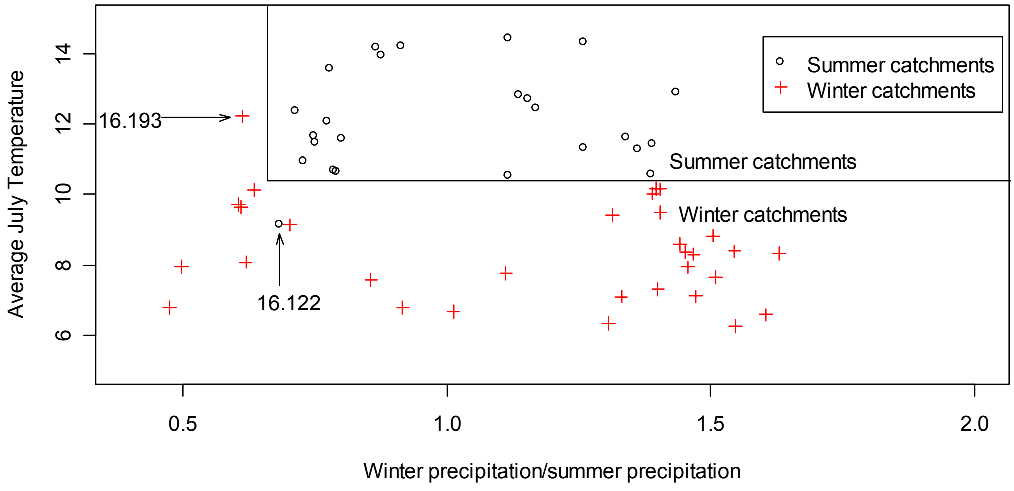

Figure 2. Classification of summer- and winter catchments. The circles and crosses indicate the summer- and winter catchments respectively according to the initial classification, whereas the lines indicate the limits according to the classification based on climatic conditions.

## 3.2 Regression models

Several alternative models were tested. They are all listed in Table 4. The estimated regression coefficients for the different regression models are shown in Table 5. $Q_c$ denotes the clf. Table 6 lists the results of the cross validation test using (1). In all models, except M6, separate regression equations were established for summer- and winter catchments. Note that the best model is the one with the highest value for $R_{CV}^2$ .

Figs. 3–12 show some diagnostics for the fit of the regression equations and indications of how well the regression requirements 1)-5) are fulfilled. In each figure there are four plots for the summer catchments and four plots for the winter catchments. In the upper plots clf is untransformed, whereas the lower plots show the results as log-transformed clf. Note that identical regression coefficients from Table 5 were applied to estimate both the untransformed and the log-transformed clf. In each case the two first plots are for the summer region. The first plot shows predicted versus observed clf. For a good model fit the points should be close to the 1:1 line. The second plot is a qq-plot for the residuals. For normally distributed residuals, the points should be on the 1:1 line. The third and fourth plots show the same results, but for the winter region.

Different transformations of the dependent variable, clf, were tested in order to obtain homoscedasdic residuals. The first model M1 is for untransformed variables. From Fig. 3 we see that the residuals are heteroscedasdic. Several transformations of the clf were tested in order to obtain homoscedasdic residuals, and the log-transformation was found to be the best alternative.

For the models M2-M5 different transformations of the independent variables were tested. M2 was based on the regression equation developed by Væringstad and Hisdal (2005). The same independent variables were used, but new regression coefficients were calculated since a slightly different dataset was used. All variables were log-transformed, and for the land cover variables, 0.1 was added prior to the log-transformation since they could have the value 0. For the temperatures 10 was added before log-transformation to avoid negative values.

To test the effect of dividing the data into summer- and winter catchments, a regression equation was developed for all data in one group (M6) using the same stepwise procedure as for M5.

Table 4. List of the different regression models that were tested.

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

Table 5. The different regression models.

| **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** | $\ln(Q_c) = -0.570 + 0.770 \ln(Q_M) - 0.202 \ln(H_{\min})$ |
| **M2-Summer** | $\ln(Q_c) = -2.080 + 1.166 \cdot \ln(Q_M) - 0.534 \ln(R_G)$<br>$- 0.368 \ln(B_{\%} + 0.1) + 0.153 \ln(M_{\%} + 0.1)$ |
| **M3-Winter** | $\ln(Q_c) = -6.387 + 0.835 \ln(Q_M) - 0.391 \ln(M_{\%} + 0.1)$<br>$- 0.175 \ln(F_{\%} + 0.1) + 0.274 \ln(C_L) + 2.350 \ln(T_A + 10)$ |
| **M3-Summer** | $\ln(Q_c) = -4.288 + 1.282 \ln(Q_M) + 0.379 \ln(L_{\%} + 0.1)$<br>$- 0.272 \ln(B_{\%} + 0.1)$ |
| **M4-Winter** | $\ln(Q_c) = -0.00758 + 0.00767 C_L + 0.0330 L_{\%}$<br>$+ 0.0204 Q_M - 0.00609 M_{\%} + 0.0894 T_A$ |
| **M4-Summer** | $\ln(Q_c) = -2.983 + 0.00285 P_S + 0.0976 L_{\%}$<br>$+ 0.0116 M_{\%} + 0.0150 C_L$ |
| **M5-Vinter** | $\ln(Q_c) = -3.3325 + 0.0102 C_L + 0.03298 \ln(L_{eff}) + 0.026485 Q_M$<br>$+ 1.601 \ln(T_S) - 0.215 \ln(F_{\%} + 0.1) - 0.0173 M_{\%}$ |
| **M5-Summer** | $\ln(Q_c) = -4.734 + 1.301 \ln(Q_M) - 0.448 \ln(B_{\%} + 0.1)$<br>$+ 0.102 L_{\%} + 0.0130 C_L$ |
| **M6-All** | $\ln(Q_c) = -1.355 + 0.0238 (Q_M) + 0.380 \ln(C_W) + 0.243 \ln(L_{\%} + 0.1)$ |

Table 6 The cross validated $R^2_{CV}$ for log-transformed and for non-transformed values.

| **Model** | $R^2_{CV}$ **Summer Transformed/ untransformed** | $R^2_{CV}$ **Vinter Transformed/ untransformed** | $R^2_{CV}$ **All catchments Transformed/ untransformed** |
|---|---|---|---|
| M1 | -/0.447 | -/0.5656 | -/0.587 |
| M2 | 0.689/0.467 | 0.561/0.480 | 0.712/0.537 |
| M3 | 0.792/0.659 | 0.695/0.667 | 0.803/0.703 |
| M4 | 0.845/0.676 | 0.696/0.607 | 0.829/0.670 |
| M5 | 0.820/0.757 | 0.816/0.711 | 0.855/0.755 |
| M6 | | | 0.6915/0.5203 |

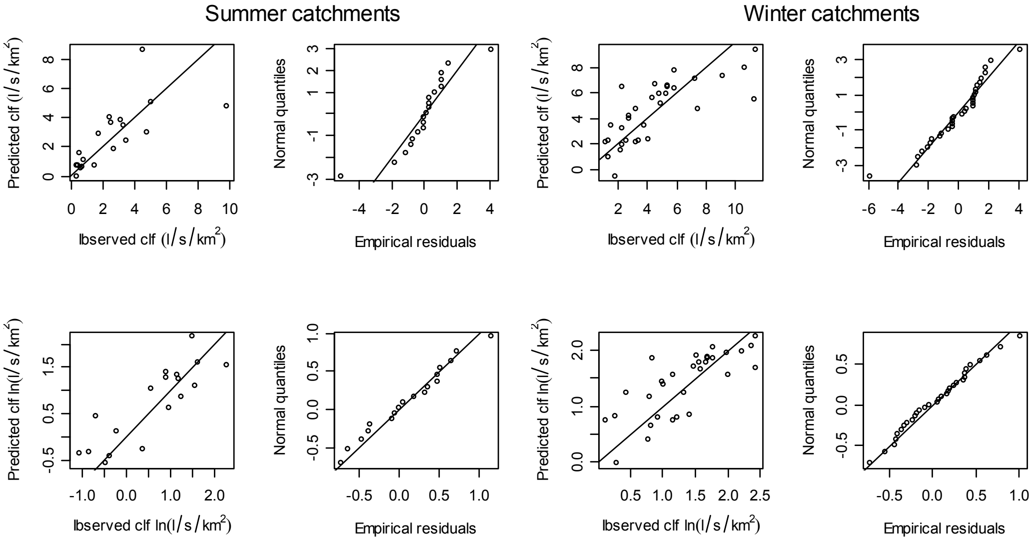

Figure 3. Cross validation of Model 1. The upper plots show the results as untransformed clf, whereas the lower plots show the results with clf log-transformed.

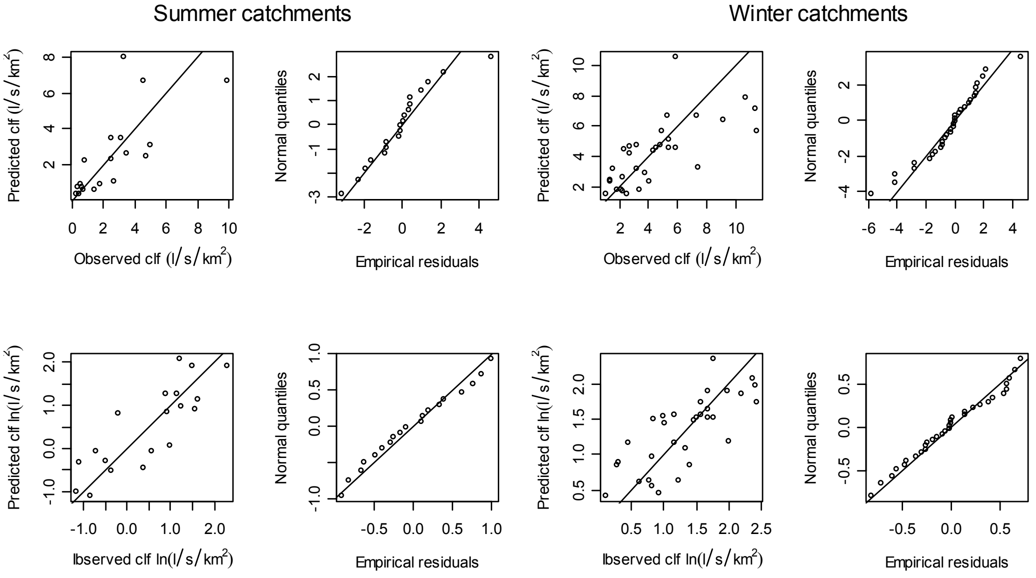

Figure 4. Cross validation of Model 2. The upper plots show the results as untransformed clf, whereas the lower plots show the results with clf log-transformed.

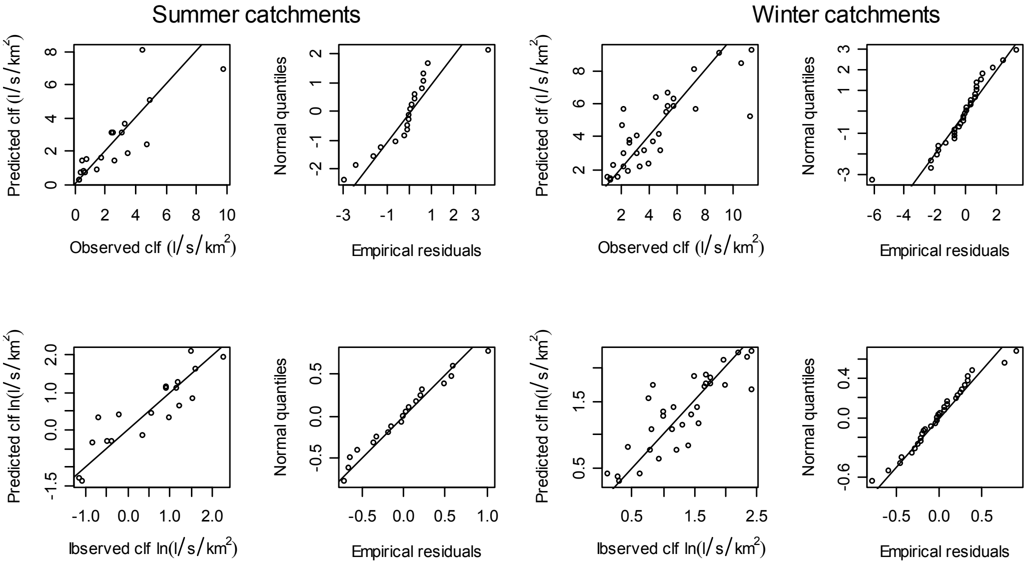

Figure 5. Cross validation of Model 3. The upper plots show the results as untransformed clf, whereas the lower plots show the results with clf log-transformed.

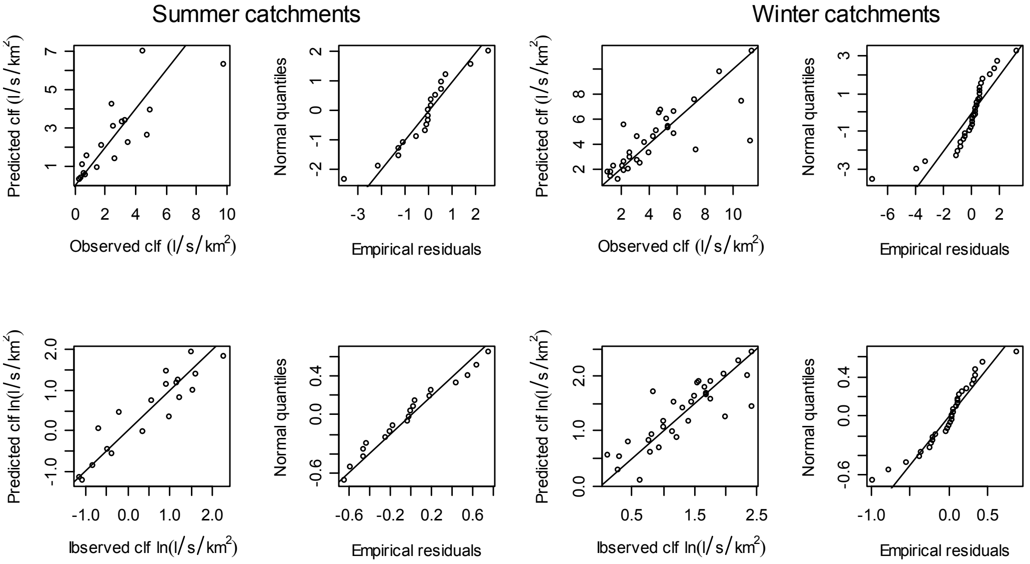

Figure 6. Cross-validation of Model 4. The upper plots show the results as untransformed clf, whereas the lower plots show the results with clf log-transformed.

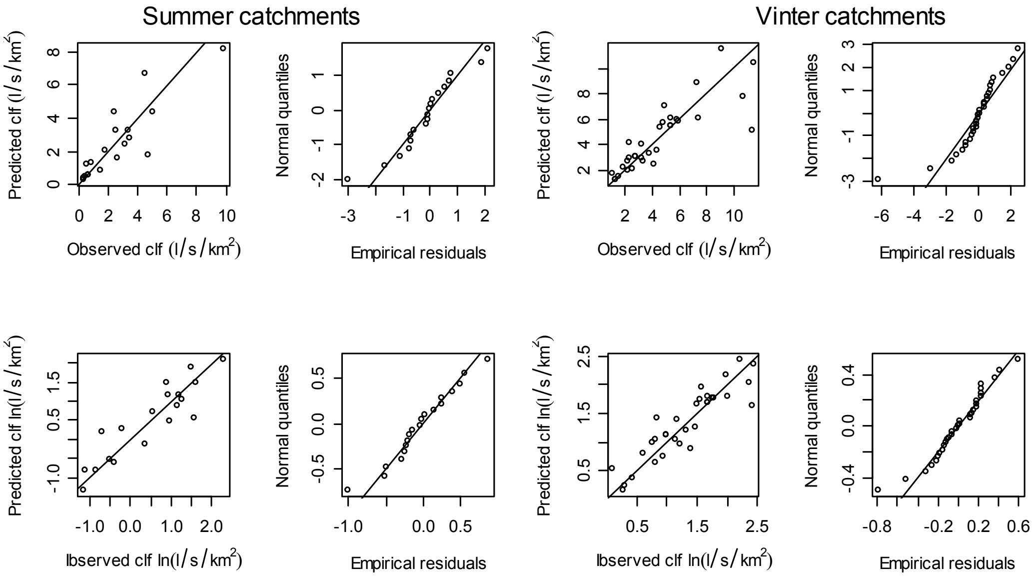

Figure 7. Cross-validation of Model 5. The upper plots show the results as untransformed clf, whereas the lower plots show the results with clf log-transformed.

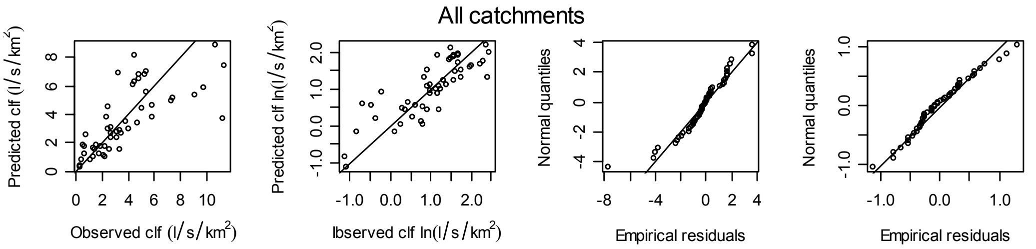

Figure 8. Cross validation of Model 6.

## 3.3 Model fit

The model that gives the best fit according to the $R^2_{CV}$ in case of recalculation to untransformed clf is Model 5 for both summer- and winter catchments. For this model the clf is log-transformed when the regression equation is established, and the stepwise procedure chooses between untransformed and log-transformed independent variables.

From Table 3 we see clearly that the $R^2_{CV}$ is much smaller for Model 6 than for Model 5 where we separate between summer- and winter catchments. We can therefore conclude that it increases the performance of the regression model to separate between winter and summer low flow regions.

The qq-plots of the residuals indicate that residuals from log-transformed clf are close to normally distributed. For untransformed residuals, the normal distribution does not fit so well.

The bias of the residuals is centred on zero, but for many of the models we can see a negative bias for the highest values meaning that the model under-estimates the highest values. For a few of the models the lowest values are over-estimated.

### 3.3.1 Important independent variables

The different models prefer different independent variables, but some common features are seen. In all models average runoff $Q_M$ is included, except in Model 4 for summer catchments where summer precipitation $P_S$ is includes instead. The clf increases with increasing average runoff.

Another important independent variable is lake percentage. It is included in 7 of the 11 equations, and for Model 5 it is included both for summer and winter catchments. Increasing lake percentage gives increasing clf. Bogs have the opposite effect. Increasing bog percentage gives decreasing clf. Bogs seem to act as swamps in the landscape. This variable is included in 4 of 11 equations, for M5 it is included for the summer region.

For the winter region the temperature is selected in 4 of 5 equations. The clf is increasing with increasing temperatures. This is reasonable for winter catchments where snow cover formation and snow melt, highly influenced by temperature, controls the magnitude of the low flow.

Catchment geometry is also important. In 5 of 11 equations either catchment length $C_L$ or catchment width $C_w$ is included. The clf increases with these variables. This is related to the catchment areas, indicating that larger catchments have larger clf.

The results show that a better fit is obtained for summer catchments than for winter catchments. The reason is that the low flow data in the winter period are more uncertain than the summer low flow. During winter the instrument might freeze, and the low flow might actually be estimated flow based on an ice reduction procedure.

### 3.3.2 Limitations

The results are based on a statistical regression procedure and the results are therefore limited to the selected region and to the ranges in the values of the catchment characteristics used to estimate the regression coefficients. The quality of the estimates might decrease if the equations are applied in an extrapolation mode. If the procedure is to be applied for other regions in Norway, new regression equations have to be developed. It should also be noted that since a regression model is used, the estimated low flow is not an exact value, a 95 % confidence interval can be calculated around the estimate. New or more precise data in the future might give different estimates and smaller confidence intervals.

Sometimes it is necessary to be careful with which independent variables to include in the regression equation. In M5 both forest- and mountain percentages are included in the regression equation. The clf is decreasing with both increasing mountain percentage and increasing forest percentage. It seems reasonable that for winter catchment clf is the smallest for catchments with a large mountain percentage, but it is difficult to explain why clf is decreasing with increasing forest cover. One important reason for non-intuitive regression coefficients is correlation between the independent variables. Table 7 list the correlation coefficients between the independent variables in M5. We see high intercorrelations between $T_S$, $F_\%$ and $M_\%$, that might give regression coefficients that are difficult to interpret and they should be used with care.

Table 7, correlation between the independent variables in M5.

| | **$C_L$** | **$L_{eff}$** | **$Q_n$** | **$T_S$** | **$F_\%$** | **$M_\%$** |
|---|---|---|---|---|---|---|
| **$C_L$** | 1.00 | 0.027 | 0.061 | 0.193 | 0.216 | 0.103 |
| **$L_{eff}$** | 0.0271 | 1.000 | 0.199 | -0.0441 | -0.198 | 0.025 |
| **$Q_n$** | 0.061 | 0.199 | 1.00 | -0.019 | -0.323 | 0.734 |
| **$T_S$** | 0.193 | -0.044 | -0.019 | 1.000 | 0.504 | -0.416 |
| **$F_\%$** | -0.216 | -0.198 | -0.324 | 0.504 | 1.000 | -0.696 |
| **$M_\%$** | 0.103 | 0.025 | 0.734 | -0.416 | -0.696 | 1.00 |

The regression procedure does not account for strongly correlated clf along the river network. This means that if the clf is to be calculated for a partly gauged catchments (e.g. if streamflow data are available further up- or downstream) application of these measurements can give better results than the regression equations.

# 4 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 (Roald, 2006) and in combination with ecological modelling (L'Abée-Lund et al., 2004). The HBV-model operates on daily time steps and was in this study applied as a regional model to the study region. The model calculated the water balance for grid-cells of 1x1km. For each grid-cell the percentage of lake and glacier was decided in addition to the proportion of the two dominant out of five land use classes (Table 7). 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. The precipitation is defined as snow or rain decided by a threshold temperature. The snow melt is calculated using a degree-day-factor. Interception on vegetation is defined by a maximum interception storage. Rainfall reaching the ground and snowmelt leaving the snow pack infiltrates and is divided between the soil moisture zone and the upper groundwater zone, depending on the soil moisture content. From the soil moisture zone the water can evaporate. The evaporation is assumed to be proportional to the temperature and has a seasonal profile. The actual evaporation is reduced when little water is available in the soil moisture zone. From the upper zone the water can percolate to the lower zone or flow out like a piecewise linear reservoir. From the lower zone the water can be drawn up to the soil moisture zone if it is dry, or flow out like a linear reservoir. The model also contains special units for lakes and glaciers. To calculate the catchment runoff in our study, water at each grid-cell was sent without delay to the catchment outlet.

Table 7. 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. |

Precipitation and temperature observations were provided by the Norwegian meteorological institute. Stations with at least 20 years of observations in the period 1961-1990 were used. Precipitation and temperature observations 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. Then the gradient was halved. 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. By experience, the modelling results are very sensitive to these gradients, and the best results are obtained when the gradients are calibrated to each catchment. In our case, gradients that can be used everywhere within a region are needed.

## 4.1 Calculation of common low flow

The estimated clf should be easy to derive for any user-defined catchment. This requires that as a part of the pre-processing, the HBV-model is used to calculate clf for the whole of Norway. At least two possibilities exist:

1. Each grid-cell represents an upstream area/catchment. The HBV-model is used to estimate clf for each grid-cell. The user selects a point/grid, and obtains the clf based on the whole upstream area. This procedure requires that a correct drainage direction grid is created for the whole country. Locally, problems might arise either due to the resolution of 1x1 km or due to small errors in the drainage direction grid.
2. Each grid-cell represents only itself. The HBV-model gives the clf for each grid-cell. The user has to define the catchment boundary, and the clf for the catchment is the average of the clfs in the grid-cells within the catchment. This procedure might be difficult to use for large catchments where all parts of the catchment do not simultaneously contribute to the low flow events, e.g. some parts of the catchment have summer low flow and other parts winter low flow.

For practical reasons and because the need for low flow indices is largest for small catchments, the last alternative was selected in this study.

## 4.2 Calibration and validation

To evaluate the predictability of clf in ungauged catchments using the HBV-model, a split sample test was applied. Daily streamflow observations from 36 stations were used for calibration (Table 1). Only stations with observations in the period 1961-1990 were selected. The software PEST (Doherty, 1998) was used for automatic calibration of the model. Originally the HBV-model was calibrated using the average root means square error for runoff values measured in mm for selected catchments all over Norway (Beldring et al, 2002). This original calibration will be called ‘first calibration’ in the result section. This calibration gives a high weight to high streamflow values. Therefore a new calibration was performed using average Nash-Sutcliffe coefficient for log-transformed streamflow as calibration criterion. This calibration will be referred to as ‘second calibration’ in the results section. This criterion was applied to obtain a best possible fit to the smallest flows. To reduce the number of parameters needing 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 clf was calculated both for the calibration and the validation catchments and compared to observed values. The explained variance $R^2$ and bias was calculated both for the calibration and the validation set.

# 5 Results

In order to compare the prediction of clf 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 36 catchments were used for estimating the coefficients in the best regression model, M5, in Table 5. The regression equations were established separately for summer and winter catchments. The estimated coefficients were then used to obtain the predicted low flow in the independent catchments. The explained variance $R^2$ and bias was calculated both for the calibration and the validation set.

Fig. 9 shows the observed and HBV-estimated clf for the calibration catchments, and Fig. 10 show the observed and simulated clf for the validation catchments. Both figures also show the results for the regression method.

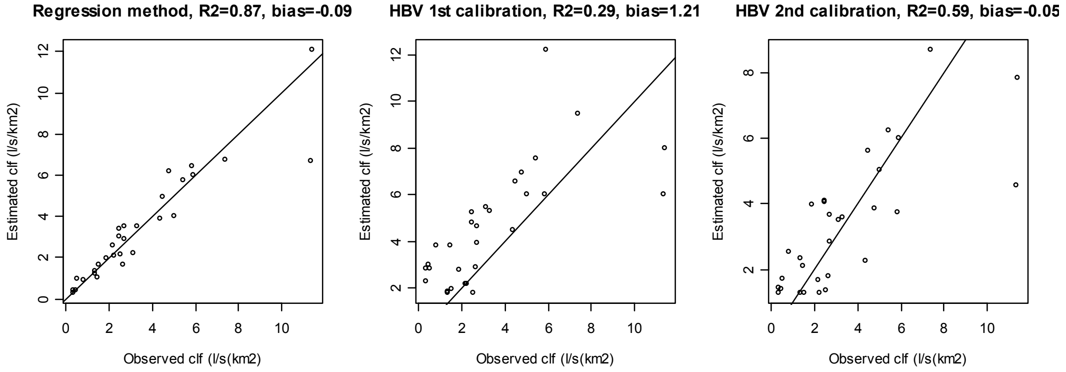

Figure 9. The observed and simulated clf for the calibration catchments for a) The regression method; b) HBV-model, first calibration; c) HBV-model, second calibration with high weights on low streamflow values.

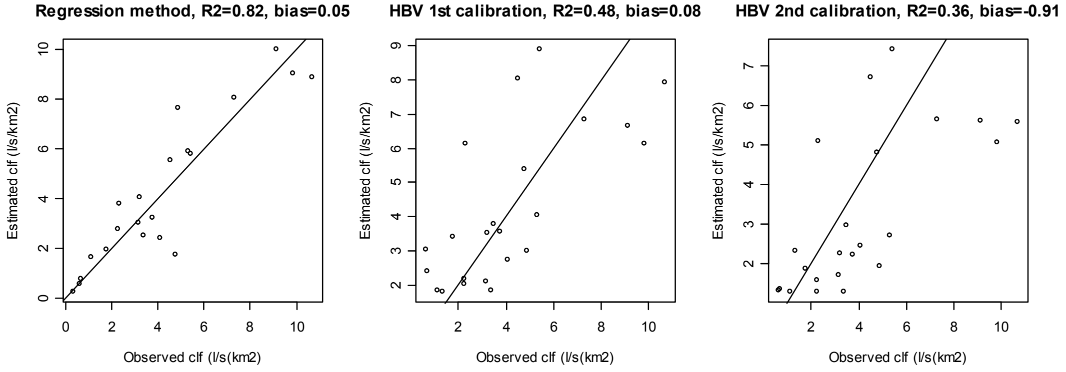

Figure 10. The observed and simulated clf for the validation catchments for a) The regression method; b) HBV-model, first calibration; c) HBV-model, second calibration with high weights on low streamflow values.

# 6 Discussion

## 6.1 The HBV results

The results of the first calibration show that the clf is over-estimated and the explained variance is only 0.29 (Fig. 10b). In the second calibration the bias is closer to zero and the explained variance increases to 0.59 (Fig. 10c). The re-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 both considering bias and explained variance.

The use of the HBV-model to calculate clf demands high performance of the simulated recession period. Since the estimates are to be used in ungauged catchments, it is necessary to use a regional parameter set where the parameters depend on landscape characteristics. Better results would be obtained if the model was calibrated to individual catchments. The regression method indicates that, in addition to climatic descriptors, lakes and bogs are important landscape characteristics that control the low flow. In this version of the HBV-model, lakes are accounted for within each grid element, as a percentage of the land cover. Individual lakes are not included as explicit elements in the model. The bogs are not included in the model parameterisation. Better results might be obtained in case of improved interpolation of precipitation, improved representation of lake elements and introduction of soil and land use classes that are important for the recession.

The construction of a mean annual runoff map for Norway contained two important steps (Beldring *et al.*, 2002) calculation of the mean annual runoff for each 1x1km landscape element and a bias correction procedure where the bias was calculated in catchments with observations and then interpolated in space using inverse distance weighting.

A similar procedure could have been applied to a map of clf. Such a procedure is based on the assumption that the bias is more similar for catchments separated by a short distance than for catchments separated by a long distance. A semivariogram for the bias will indicate whether this is the case. A bias correction procedure would also require a split sample test in order to evaluate its predictive performance.

## 6.2 Comparison to regression

The regression method in general gives better prediction of clf in ungauged catchments than the HBV-model. The regression method is especially superior to the HBV-model for the smallest clf-values. The predictive power for clf-values smaller than 2-3 $ls^{-1}km^{-2}$ for the HBV-model is rather limited.

The error in the regression estimated clf depends on the predicted level. The use of log-transformation to obtain heteroscedastic residuals indicates that the error is relative. This means that the absolute error is small for small predicted values of clf and larger for larger predicted values of clf. This is not the case when using the HBV-model. The absolute error seems to be independent of the magnitude of clf. This means that the precision of the lowest clf predictions is rather low.

The regression method will not be able to produce a static map of clf. The method will instead provide a set of regression equations and a system to extract the catchment characteristics. The HBV-model, however, would be able to provide a static map of clf. Also, the user could define catchment boundaries and the average clf of the grid-cells within the boundaries would be calculated. None of the methods account for the correlation in clf along rivers. If measurements are available in the same river, theses should be used to obtain estimates of clf.

# 7 Conclusions

A multiple regression procedure to estimate common low flow in ungauged catchments has been compared to the application of the HBV-model. Some conclusions can been drawn:

- The regression method gives better predictions of clf than the HBV-model.
- For the regression method, the best results are obtained when the clf is log-transformed. The independent variables should be either log-transformed or kept un-transformed.
- Important catchment characteristics are average runoff, lakes, bogs, catchment size and temperature.

For the development of a low flow map for Norway, several challenges are identified:

- Quality control of low flow data.
- Development of regression equations for several low flow indices for the whole of Norway.
- Development of a GIS-based program to extract clf and other low flow indices in ungauged catchments.

It should be noted that the regression equations provided in this report are valid only for the study region and that the equations should be regarded as preliminary. In the continuation of this work, new regression equations will be developed for other regions in Norway, and it might also be necessary to re-estimate the regression equations for the study region used in this report.

# 8 Refrerences

Beldring, S., Engeland, K., Roald, L.A., Sælthun, N.R. and Voksø, A. (2003) Estimation of parameters in a distributed precipitation-runoff model for Norway. *Hydrology and Earth System Sciences* **7**(3), 304-316.

Beldring, S., Roald, L.A. and 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.

Demuth, S. and Young, A.R. (2004) Regionalisation Procedures, In: L.M. Tallaksen & H.A.J. van Lanen (Eds), *Hydrological Drought Processes and Estimation Methods for Streamflow and Groundwater*. Developments in Water Sciences 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/documents/pestman.pdf - accessed January 6 2006)

Førland, E.J. (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. and Sauquet, E. (2004) Mapping runoff characteristics in a river basin. *Hydrol. Processes* (submitted).

Krokli, B. (1988) Analysis of low flows (in Norwegian), NVE Publikasjon No. 14, Norwegian Water Resources and Energy Directorate, Oslo, Norway.

Laaha, G. and Blöschl, G. (2006) A comparison of low flow regionalisation methods—catchment grouping. *Journal of Hydrology* (in press) Corrected proof available online 24 October 2005.

L'Abée-Lund, J.H., Vøllestad, L.A. and 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.

Lovdata (www) The water resources act (in Norwegian) http://www.lovdata.no/all/hl-20001124-082.html.

Petersen-Øverleir, A. (2005) Personal communication.

Roald, L.A., Beldring, S., Skaugen, T.E., Førland, E.J. and Benestad, R. (2006) Climate change impacts on streamflow in Norway. Consultancy report A no 1-2006, Norwegian Water Resources and Energy Directorate, Oslo, Norway.

Skaugen, T., Astrup, M., Mengistu, Z. and Krokli, B. (2002) Low flow – estimation and a basis for licences (in Norwegian), NVE rapport miljøbasert vannføring 1 2002, Norwegian Water Resources and Energy Directorate, Oslo, Norway.

Sælthun, N.-R. (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, Oslo, Norway.

Væringstad, T. and Hisdal, H. (2005) Estimation of common low flow in ungauged catchments (in Norwegian), Miljøbasert vannføring rapport 06-2005, Norwegian Water Resources and Energy Directorate, Oslo, Norway.

This series is published by Norwegian Water Resources and Energy Directorate (NVE)

## Published in the Report series 2006

No. 1 Kolbjørn Engeland, Hege Hisdal, Stein Beldring: A comparison of low flow estimates in ungauged catchments using regional regression and the HBV-model (34 pp.)