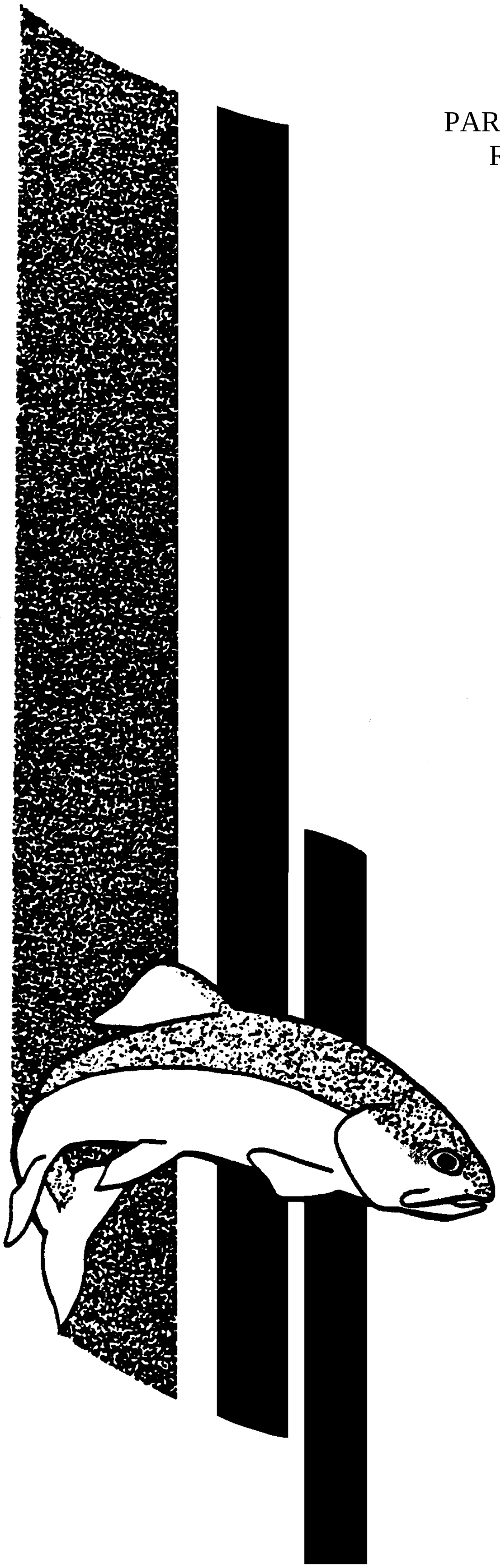


PARTIAL VALIDATION OF STREAMFLOW METRICS AND RUNOFF PARAMETERS, AND IDENTIFICATION OF PHYSICAL VARIABLES TO HELP PREDICT STREAMFLOW SENSITIVITY USING SELECTED ECOPROVINCES WITHIN BRITISH COLUMBIA

by

J.-S. Moore
Département de Biologie
Université Laval
1030 Avenue de la Médecine
Québec, QC, G1V 0A6

and

Jordan Rosenfeld
Conservation Science Section
Province of British Columbia
2202 Main Mall
Vancouver, BC, V6T 1Z4

Fisheries Project Report No. 142
2013

Province of British Columbia
Ministry of Environment

Partial validation of streamflow metrics and runoff parameters, and identification of physical variables to help predict streamflow sensitivity using selected EcoProvinces within British Columbia

by

Jean-Sébastien Moore
Département de Biologie
Université Laval
1030 Avenue de la Médecine
Québec, QC, G1V 0A6
581-888-1868

and

Jordan Rosenfeld
Aquatic Conservation Science Section
Province of British Columbia
2202 Main Mall
Vancouver, BC, V6T 1Z4
Phone: (604) 222-6762
Fax: (604) 660-1849

BC Ministry of Environment
Fisheries Project Report RD142
2013

Cite as: Moore, J.-S., and Rosenfeld, J.S. 2013. Partial validation of streamflow metrics and runoff parameters, and identification of physical variables to help predict streamflow sensitivity using selected EcoProvinces within British Columbia. BC Ministry of Environment, Fisheries Research Report RD142, Victoria, B.C.

ABSTRACT

Moore, J.-S., and Rosenfeld, J.S. 2013. Partial validation of streamflow metrics and runoff parameters, and identification of physical variables to help predict streamflow sensitivity using selected EcoProvinces within British Columbia. BC Ministry of Environment, Fisheries Research Report RD142, Victoria, B.C.

Prediction of flow in ungauged streams can be approximated using data from nearby fully gauged water stations, provided that variation between adjacent basins is small relative to variation at larger spatial scales (e.g., Oderkoff 2003), or if this variation can be controlled using physical correlates that influence the quantity of runoff (e.g. elevation). Using this approach, Ron Ptolemy, BC Ministry of Environment, has developed a database (HydroMaster) for calculating areal runoff, Mean Annual Discharge, and summer and winter monthly minimum flows by Ecoprovince and EcoRegion based on historic flow records published by the Water Survey of Canada. However, this database includes stations with different periods of record, different lengths of record, and stations with seasonally incomplete records (temporal gaps). The goals of this report are *i)* to extract flow metrics from Water Survey of Canada data for a subset of 3 EcoProvinces with a common period of record and a minimum of 10 years of data to compare to the same metrics for the same EcoProvinces estimated using the HydroMaster database, which includes a larger set of flow stations with variable periods of record; and *ii)* to determine whether including watershed attributes other than EcoSection (e.g., lake area, presence of glaciers, mean elevation, relief, regulation, i.e., regulated or natural) improve predictions of streamflow.

Although estimates of MAD in HydroMaster were positively biased relative to MAD estimates from the subset of streams with a fixed common period of record and minimum 10 years of annual flow data, the bias was relatively small ($1.2\ l \cdot s^{-1} \cdot km^{-2}$), may be attributed at least in part to the common period of record being relatively dry, and is unlikely to be of significance for screening or management purposes. Classifications of summer or winter low flows as flow sensitive based on Tennant criteria (lower than 20% MAD) were generally very similar between HydroMaster (with a variable period of record) and the analysis with a common period of record, indicating that classification of streams as flow sensitive is fairly robust to variation in the period of record, and that data from HydroMaster should provide a reasonable representation of regional flow sensitivity based on Tennant criteria. Some of the physical variables examined in the current analysis appear to be useful predictors of flow (in particular presence of glaciers), and may be useful as covariates to refine predictions beyond the basic EcoSection or EcoProvince classifications.

INTRODUCTION

Water management in British Columbia (BC) has relied on a variety of water supply information such as 7-day low flows and use of proxy gauged streams to infer flow limitations on ungauged streams (Obedkoff 2000; Obedkoff 2003; “proxy-gauged streams” in this context refers to using gauged Water Survey of Canada streams to estimate flows in nearby un-gauged basins). Simple evaluations of flow effects on fish habitat relies largely on the Tennant or Montana method (Tennant 1976), based on flows inferred from this historical instreamflow approach (Jowett 1997), i.e. by inferring critical flow thresholds on the basis of a fixed percentage of historical mean annual discharge (MAD). The method is therefore simple to apply to streams for which historical data are available, i.e., streams where the Water Survey of Canada (WSC) maintains or maintained a gauging station. More challenging are situations where no historical data is available, which is the case for most streams in BC.

Although useful syntheses, the analysis of areal runoff in different hydrologic zones throughout B.C. by Obdekoff (2000, 2003) were based on a limited subset of gauged streams relative to the total number of gauged sites in the province. While providing estimates of areal runoff (from which MAD can be estimated), the analysis by Obdekoff did not collate summer or winter low flow metrics, which may be critical limiting periods for fish populations. The HydroMaster dataset was developed to help with flow assessments on ungauged streams (Ptolemy and Lewis 2002), and builds on the Obdekoff analysis by incorporating data from a larger number of gauged streams, while additionally extracting low flow (winter and summer) metrics for each station. The HydroMaster dataset collates data for most BC streams for which historical flow data is available through the WSC online database, as well as local Ministry of Environment water management files. This dataset then served as the basis for an analysis of flow sensitivity (based on Tennant criteria) whose goal was the identification of relatively hydrologically and ecologically homogeneous landscapes (EcoSections, EcoRegions and EcoProvinces) that are sensitive to water withdrawals in the winter low-flow period, summer low-flow period, or both, based on regional mean low flows and how they relate to Tennant thresholds (R. Ptolemy, BC Ministry of Environment, personal communication Dec. 2012). Some of the main products derived from this analysis are maps of flow sensitive EcoSections that can be used as primary screening tools to inform the water license granting process.

This report includes an analysis of the HydroMaster database to satisfy two main goals. First, this analysis considers some concerns recently identified by independent reviews of the HydroMaster dataset. Second, more detailed analysis of flow data were explored that might allow predictions of regional flow sensitivity to be made with more precision and spatial resolution. More specifically, it would be useful to determine if variables other than EcoSection can be useful in predicting MAD or summer/winter low flows in ungauged streams. This could potentially improve the accuracy of predictions, as well as provide insight into the drivers of flow variation within and across landscapes. We should also clarify that reference to “flow sensitivity” in this report refers to whether summer or winter flows fall below 20% MAD, a recognized Tennant threshold; the degree to which flows below this threshold result in impairment to fish populations or other biological attributes is a separate question that is beyond the scope of this study.

Ecofish research recently reviewed the analysis of winter flow sensitivity based on the Hydromaster dataset as part of the BC Ministry of the Environment (MOE) Winter Flows Project (Memorandum to Jennifer Turner, 29 February 2012, file 1144). In addition to other reviewers, they identified several concerns with the method that the present analysis attempts to evaluate in more detail to better understand the magnitude of any biases. More specifically, the present analysis addresses the following concerns previously raised with respect to clarification of methods used to populate the HydroMaster database:

1. The HydroMaster analysis of regional mean flows does not account for inter-year variation in relation to periodic climate fluctuations like the Pacific Decadal Oscillation (PDO). While the present analysis does not specifically test for the effects of PDO on hydrological parameters, it uses a **common period of record** (POR) that should minimize climate-induced variation that might otherwise obscure regional relationships.

2. Many of the WSC gauge stations have **less than 10 years of records**. It was suggested that a minimum of 20 years of data should be used to appropriately represent the mean flow attributes of a stream. Because this would severely limit the number of streams in the analysis, we did not follow this suggestion, but only included streams that have more than 10 years of data within the common period of record. This allows comparing predictions between a database with a fixed period of record and the more diverse HydroMaster data set and to assess the magnitude of any potential bias associated with the (often) shorter period of record for many streams included in the HydroMaster data set.

3. There are some discrepancies between data in HydroMaster and the WSC data (e.g. drainage area at some sites).

4. Some data sources or parameter calculations are unclear in the HydroMaster data set (e.g. some sites are only gauged during summer months, and it is unclear how mean annual discharge was calculated for these sites).

5. Potential effects of watershed elevation on stream discharge and areal runoff were not specifically considered in HydroMaster, although they are indirectly included because of the limited range of physical attributes associated with streams in a particular Ecosection. The present analysis takes elevation (and other landscape variables) into consideration, to evaluate their potential as additional predictors of regional flow metrics.

It should be noted that many of the analyses below were inspired by the regional analysis of flood flows in BC by Eaton et al. (2002), which should be consulted for a more detailed description of the rationale behind some of the analyses that follow.

METHODS

Selection of stations for analysis

Choice of regions – We first selected three EcoProvinces (Demarchi 2011) as the focus of the current analysis (Table 1). The goal was to extract a representative subset from the larger HydroMaster database to make progress possible within a limited timeframe, while at the same time limiting the amount of unaccounted for variation due to differences in region. We chose the EcoProvinces non-randomly so as to maximize the potential differences in climate and physiography (and therefore areal runoff; Figure 1); however, they should be representative in other respects, e.g. in terms of data quality, errors, or biases. For each of the three EcoProvinces, three EcoSections were selected. In each of the EcoSections, all streams that had data for more than 10 years over the common period of record were included. Raw Water Survey of Canada data was extracted from the HydroMaster database for all analysis. Although we used untransformed Water Survey of Canada data from the HydroMaster database rather than downloading data independently from the Water Survey of Canada website, we downloaded data from a minimum of 15 streams to spot-check against HydroMaster without finding any discrepancies in the raw flow data.

Common period of record – Climate has a large influence on flows, and year-to-year variation in climate (particularly long term temporal trends) can be important. In British Columbia, the decadal-scale climatic oscillations associated with the Pacific decadal oscillation (PDO) have been shown to influence flows

(Whitfield et al., 2010). Data from time-series that occur in different regimes of the PDO can therefore be non-comparable. Our goal was therefore to reduce such temporal noise and estimate flow metrics for a common period of record. We then compared these metrics with estimates of long-term MAD from HydroMaster that did not use a common period of record to determine whether they lead to comparable insights regarding low-flow sensitivity (i.e., to determine whether variable periods of record in HydroMaster significantly bias results). Following Eaton et al. (2002), we chose as a common period of record (POR) the years 1976-1995 because their analysis showed that these years were hydroclimatologically similar. Only stations that had data for at least 10 years during this time period were included in the analyses presented here. Note that previous reports suggest that a minimum of 20 years of data should be used (King et al. 2003). However, such a criterion would have severely limited our sample size, i.e., a tradeoff between increasing confidence associated with a long flow record vs. adequate replication among sites is inherent to this sort of analysis. In addition, the use of a common period of record should control for some of the variation that might require a 20 year POR.

Table 1. Subset of BC streams from the HydroMaster dataset used in the present analysis.

EcoProvince	EcoSection	No. of stations	No. of stations with natural flow
Coast and Mountains (CAM)	Eastern Pacific Range	6	5
	Southern Pacific Range	8	4
	Nass Mountains	5	5
	TOTAL	**19**	**14**
Southern Interior (SI)	Northern Okanagan Highlands	15	5
	Shuswap Basin	1	0
	Western Okanagan Uplands	8	7
	TOTAL	**24**	**12**
Southern Interior Mountains (SIM)	Central Columbia Mountains	6	6
	Southern Park Ranges	4	3
	Selkirk Foothills	6	4
	TOTAL	**16**	**13**

Regulated vs. natural flows – Unlike Eaton et al. (2002), we chose to include regulated flows to see if the effects of flow regulation could be predicted along with natural flows. While we include them, they are often analyzed separately to ensure that the results are not biased by their inclusion (i.e., regional annual runoff should only be estimated using unregulated streams). Regulation is defined as alterations of the flow regime, including storage (behind mainstem dams or off-channel ponds) that change the hydrograph but may not alter MAD, and diversions into or out of a stream that may not alter storage but may alter MAD or low flows. Watersheds without storage or diversion were considered natural, although we recognize that land-use impacts affecting surface runoff or soil permeability may also alter hydrographs. We recognize that many diversions or water withdrawals may be unrecorded or poorly documented, so that classification of many streams as "natural" may be in error. The degree of hydrologic alteration would most accurately be treated as a continuous covariate rather than a binary variable, but time and analytic constraints prevented this higher level of analysis.

Limit on drainage size – Again unlike Eaton et al. (2002) we chose not to limit drainage size. This was done because we are interested in predicting flows for small as well as large basins. Note that 10 of the 59 stations used in this analysis have a drainage area smaller than the lower size threshold of 60km^2 used in Eaton et al. (2002). None of the stations used had a drainage basin larger than the maximum threshold of

10,000km^2 of Eaton *et al*. (2002). It should also be noted that when the drainage areas reported by the WSC and those from the HydroMaster dataset differed, we used values from the HydroMaster dataset because they are considered more accurate (i.e., R. Ptolemy measured DA using the iMapBC mapping tool and used this corrected measure in the HydroMaster dataset when he uncovered obvious gross errors in reported drainage areas in the WSC database).

Stream-flow metrics used

For evaluating HydroMaster, we used streamflow metrics that were easily available in HydroMaster and that were used in previous analyses. Note that it would be relatively easy to repeat the analyses described below for additional metrics if necessary (r code is available from the first author).

Mean annual discharge (MAD)

MAD was calculated as the mean of the mean monthly discharge values for the common period of record (in the original HydroMaster data set MAD was extracted directly from the WSC published accounts). Sometimes referred to below as MAD_POR to emphasize that it is from a common *P*eriod *O*f *R*ecord.

Median annual discharge (MedAD)

MedAD is the median of the mean monthly discharge values for the common period of record. This variable was included as an alternative to MAD, since some reviewers suggested that it may be a more representative measure of average flow. Although the median of all daily flows throughout the year is the ideal median statistic, time constraints prevented downloading WSC daily flows, and we used the median of mean monthly flows as a substitute.

Summer critical period streamflow (sumCP)

Yearly summer critical period low flows are the lowest monthly discharge measured between the months of July and October. SumCP therefore represents a measure of the severity of low flow habitat limitation for fish. To obtain sumCP, we estimated the geometric average of the minimum monthly values over the common period of record. The geometric mean has properties similar to the median in that it gives less weight to outliers (i.e., high flows) in skewed distributions (see http://en.wikipedia.org/wiki/Geometric_mean). Note that in HydroMaster, sumCP is expressed in terms of percentage of MAD. Here we use the actual numerical values because they are more appropriate for statistical analysis.

Winter critical period streamflow (winCP)

Same as sumCP but for the months of January to April or November to March, depending on the consistent timing of icing. The same time periods were used as in HydroMaster for each stream.

Effects of using a common period of record

To investigate whether failing to use more than a minimum number of years of data or a common period of record biases estimates of MAD in HydroMaster, we compared our estimates of MAD based on a minimum of 10 years of data from a common period of record to the estimates of long-term MAD (LTMAD) from the larger HydroMaster data set (variable period of record, with many records less than 10 years duration). This was done only on streams with natural flows, because some LTMAD in the HydroMaster data set are adjusted for regulated flows, which would lead to LTMAD being systematically higher than MAD_POR. We first investigated the extent to which the two indices were correlated using Pearson’s correlation. We also used a paired t-test to determine whether there was a significant difference between LTMAD and MAD_POR. A significant difference would indicate a systematic bias between the two indices.

Scaling effects

Because drainage area (DA) has such a large influence on discharge, all analyses are conducted on specific runoff data, i.e. discharge divided by drainage area (i.e., $l \cdot s^{-1} \cdot km^{-2}$). Despite this correction, however, some flow metrics may still be influenced by DA (e.g., Eaton et al. 2002). We therefore controlled for potential scaling effects of DA on flow metrics by using it as a covariate. We started by running ANCOVAs with the area-specific ($l \cdot s^{-1} \cdot km^{-2}$) flow metrics as the dependent variable and DA (the covariate) and EcoProvince as the independent variables to see if there was a common slope for the scaling effects among EcoProvinces (if so, we would expect a non-significant interaction between DA and EcoProvince). Because significant interaction terms were frequent (see Results section), we also analyzed each EcoProvince separately using simple linear regressions with only DA as an independent variable.

Physical variables tested as predictive correlates of flow

We used iMapBC to collect data on a series of physical variables associated with each WSC station and upstream basin that earlier discussions identified as potentially relevant predictors of low flow. These included:

Presence of glaciers in the watershed

For each stream in the dataset, the entire watershed was examined for the presence of glaciers. We adopted the simple approach of coding presence of glaciers as a binary variable (i.e. '1' whenever a glacier was present in the watershed, '0' when absent), although we recognize that a continuous variable representing the percentage of a drainage basin covered in glacier would likely provide a more sensitive index of glacial influence.

Lake area

If a lake larger than 1 km^2 was present in the watershed, its area was calculated and included in the dataset. When multiple lakes were present, we summed their surface areas. If multiple lakes smaller than 1 km^2 were present, but their surface area summed to more than 1 km^2, they were also included. We also included the ratio of lake area to DA as an additional variable to account for the fact that the hydrological effects of lakes may scale with drainage size.

Elevation

Using the contour layers in iMapBC, approximate measures of basin elevation were included in the dataset. Specifically, elevation of the station (minimum elevation) and maximum elevation of the drainage were measured, and mean elevation (i.e. max+min/2) and relief (i.e. max-min) were also calculated.

Multivariate analysis

We explored the relationship among variables using multivariate analyses. Multivariate analyses allow the ordination of categorical data (e.g., streams in different EcoProvinces) in a multi-dimensional space defined by multiple correlated variables. In addition to allowing a simple visualization of the data set along a reduced number of axes, this analysis allowed the identification of variables that are strongly correlated with each other and thus redundant. Such variables should be removed from subsequent analyses to avoid multi-colinearity and model over-fitting. A Principal Component Analysis would generally be used for this purpose, but the presence of both categorical and continuous variables in the dataset made application of this method problematic. Instead, we used Factor Analysis (FA), which is a related method that allows jointly analyzing categorical and continuous variables. The method is implemented in the R package 'FactoMineR'.

Multiple regression analysis to identify environmental correlates of flow metrics

Using the reduced set of variables selected in the Factor Analysis, we ran multiple regression analyses for each EcoProvince separately, and for each of the three response metrics of streamflow (MAD, sumCP, winCP). The independent predictor variables included were lake area, presence of glaciers, mean elevation, relief, regulation (i.e., regulated or natural), and EcoSection. The categorical variables with two states (glacier and regulation) were coded as dummy variables (i.e., as 1 or 0). EcoSection was coded in R as a categorical factor, except in the SI EcoProvince where one of the three EcoSections only had one stream. Because a factor needs to have more than one observation to be included in a linear model, we removed this stream from the analysis. This meant that only two regions were present for this EcoProvince, and we thus coded it as a dummy variable. Also for the SI EcoProvince, none of the streams have glaciers in the headwaters, and this variable was removed from this analyses.

To select the best models, we used a stepwise regression procedure using the algorithm implemented in the 'step' function in R (R Development Core Team 2012). This function starts from the full model and finds the simplified model that has the lowest Akaike Information Criterion (AIC) value. This information theoretic approach penalizes models with greater numbers of variables, such that variation explained is maximized with as few explanatory variables as possible. The selected models at the end of this analysis are thus the most parsimonious. Two versions of the algorithm were tested: one where variables are sequentially removed (i.e., direction = backward), and one where variables are either removed or added at each step (i.e., direction = both). Because both algorithms lead to identical results in all cases, we report the results only once. Note that this type of analysis does not allow interaction terms to be included, and therefore assumes that there are no significant interactions among the variables. Once the most parsimonious model is selected, the value and significance ($\alpha = 0.05$) of all coefficients and the intercept of the linear models are generated.

All analyses were performed in R (R Development Core Team 2012), and the annotated code, along with the data file, are available on request.

RESULTS

Effects of using a common period of record

Our estimates of MAD relying on a common period or record and a minimum number of years of data and the long term MAD estimates from the HydroMaster dataset were highly correlated (Pearson's r = 0.99; Fig. 2a). The estimates of MAD from the HydroMaster dataset were significantly greater than the estimates based on a common period of record (paired t-test, t = -5.14, df =37, p < 0.005; Fig 2b; Table 2). This suggests a systematic bias such that LTMAD from HydromMaster would overestimate MAD compared to an estimate based on the common period of record we chose. Note, however, that the magnitude of this bias is relatively small (mean of LTMAD = 35.08 $l\cdot s^{-1}\cdot km^{-2}$; mean of MAD_POR = 33.90 $l\cdot s^{-1}\cdot km^{-2}$), and with an average of 1.18 $l\cdot s^{-1}\cdot km^{-2}$ can be considered effectively negligible over most of the reported range of areal runoff (~5-100 $l\cdot s^{-1}\cdot km^{-2}$).

Estimates of summer and winter critical period low flows (as a % of MAD) for the common period of record were similar to those estimated using HydroMaster (Fig. 3); the slope of the line for HydroMaster estimates of summer low flow plotted against estimates from the common POR was close to unity (1.04), although the slope for winter low flows was considerably less than one (0.75), largely because of large outliers among streams with higher winter low flows (Fig. 3b). However, variability was generally low for the subset of streams that had summer or winter low flows less than 20% MAD (Fig. 3b).

Treating the summer and winter low flow estimates from the common POR as benchmark ("true") values for comparative purposes, the "correct classification" for HydroMaster estimates (in terms of whether HydroMaster correctly assigned streams to above or below 20% MAD in the summer or winter) were quite high (Fig. 3). For both summer or winter flows all streams that had flows less than 20% MAD for the common period of record were correctly classified as having flows less than 20% MAD by HydroMaster.

Six streams had summer low flows incorrectly classified as below 20% (Fig. 3a), and five streams had winter low flows incorrectly classified as below 20% (Fig. 3b). These misclassifications were all near the threshold of 20% MAD (range of 20-27% MAD) as would be expected. Although HydroMaster appears to somewhat overestimate low flows relative to the common period of record, this comparison indicates that classification of streams as flow sensitive (less than 20% MAD) is fairly robust to a variable period of record as represented by the HydroMaster data.

Table 2. Comparison among EcoProvinces in means (95% confidence intervals in parentheses) of the three flow metrics used based on the calculation using the selected streams and for the common period of record between 1976 and 1995. Long term MAD as calculated for the same streams (selected) and for all available streams from the same three EcoSection (all) based on the HydroMaster dataset are also shown. Note that regulated streams are included, and that they are corrected for in the HydroMaster dataset (details can be found in the discussion). All values are in units of $l \cdot s^{-1} \cdot km^{-2}$. Although we present SumCP and WinCP in units of $l \cdot s^{-1} \cdot km^{-2}$, they are converted to percent of MAD in the HydroMaster analysis. EcoProvince metrics were calculated as the means of metric values calculated on individual streams.

	HydroMaster		**Common period of record**		
EcoProvince	**MAD (selected)**	**MAD (all)**	**MAD**	**SumCP**	**WinCP**
CAM	77.60	79.5	62.15 (49.73 - 74.58)	26.64 (16.37 - 36.91)	25.13 (16.55 - 33.71)
SIM	23.20	19.6	22.43 (17.64 - 27.22)	7.24 (4.72 - 9.76)	3.37 (2.30 - 4.44)
SI	7.47	7.83	6.40 (5.17 - 7.62)	1.57 (1.11 - 2.03)	0.78 (0.63 - 0.93)

Table 3. Comparison of MAD for selected EcoSections calculated from the HydroMaster dataset (including all streams from each section), and the dataset using a common period of record used in the present study. For the HydroMaster-derived values, samples sizes are given in parentheses (the sample sizes for the common period of record values can be obtained from Table 1. For the values calculated with a common period of record, the 95% confidence intervals are given in parentheses. All values are in units of l·s^{-1}·km^{-2}.

		HydroMaster			**Common period of record**		
EcoProvince	**EcoSection**	**MAD (natural)**	**MAD (regulated)**	**MAD (all)**	**MAD (natural)**	**MAD (regulated)**	**MAD (all)**
CAM	Eastern Pacific Range	60.2 (n=21)	58.0 (n=2)	60.0 (n=23)	66.6 (37.4 -95.9)	NA	60.0 (32.0 - 87.9)
	Southern Pacific Range	100.1 (n=25)	103.5 (n=10)	101.1 (n=35)	62.4 (11.9 - 112.9)	62.4 (7.2 - 117.6)	62.4 (36.7 - 88.2)
	Nass Mountains	51.5 (n=11)	NA	51.5 (n=11)	64.3 (38.5 - 90.1)	NA	64.3 (38.5 - 90.1)
SI	Northern Okanagan Highlands	10.7 (n=15)	7.8 (n=39)	8.6 (n=54)	10.6 (7.9 - 13.4)	4.7 (3.3 - 6.1)	6.7 (4.8 - 8.6)
	Shuswap Basin	10.0 (n=4)	8.1 (n=31)	8.3 (n=35)	NA	NA	4.8 (NA)
	Western Okanagan Uplands	6.6 (n=12)	4.9 (n=14)	5.7 (n=26)	5.7 (4.3 - 7.1)	NA	6.0 (4.6 - 7.4)
SIM	Central Columbia Mountains	30.1 (n=20)	16.0 (n=4)	27.8 (n=24)	32.2 (26.8 - 37.6)	NA	32.2 (26.8 - 37.6)
	Southern Park Ranges	18.1 (n=16)	11.0 (n=9)	15.5 (n=25)	18.7 (3.2 - 34.2)	NA	18.7 (10.7 - 27.0)
	Selkrik Foothills	17.9 (n=8)	14.0 (n=14)	15.4 (n=22)	15.1 (8.4 - 21.9)	14.9 (8.5 - 21.2)	15.0 (11.6 - 18.5)

Scaling

We found limited evidence for a common scaling effect among EcoProvinces (Table 4), but scaling effects were weak or non-significant within EcoProvince for all three flow metrics (Table 5). When all three EcoProvinces were analyzed together (Table 4), drainage area did have a significant effect on flow metrics. There was, however, a significant interaction between EcoProvince and DA for both winCP and sumCP (Table 4), suggesting that scaling effects are different among EcoProvinces for these metrics. The interaction term, however, was not significant for MAD. Because interaction terms are significant, EcoProvinces should be analyzed separately. Although the effect of DA is positive when all EcoProvinces are analyzed together (not shown), it is negative within EcoProvinces (Table 5); sumCP and winCP also have negative slopes within EcoProvinces, suggesting that the significant effect observed across all EcoProvinces is either an artifact, or indicates underlying scaling complexity that is difficult to interpret.

Table 4. Results of the ANCOVA used to determine if the slopes of drainage area-scaled flow metrics regressed over drainage area differed between EcoProvince (Flow metric = $\log_{10}$(DA) + EcoProvince + DA*EcoProvince).

	Drainage area (log-transformed)		**EcoProvince**		**Interaction**	
Flow metric	F-value	P	F-value	P	F-value	P
MAD	12.295	<0.005	106.727	<0.005	0.723	0.49
sumCP	37.036	<0.005	58.524	<0.005	4.156	0.021
winCP	15.539	<0.005	135.649	<0.005	3.206	0.048

Linear regressions of area-standardized flow metrics on drainage area conducted on each EcoProvince separately show that the only significant scaling effects observed are in the Coast and Mountains EcoProvince for the sumCP and winCP flow metrics (Table 5). We conclude from this analysis that if scaling is present it appears to be weak overall. A common scaling relationship is definitely absent, and when scaling is evident it appears to differ between EcoProvinces. Correcting for scaling would therefore require application of a different scaling correction for each EcoProvince separately, which would make comparison among EcoProvinces difficult in all subsequent analyses. We therefore chose to not correct for scaling, but to include drainage area as a covariate in all subsequent analyses.

Multivariate analysis

The factor analysis emphasized the importance of EcoProvince and EcoSection as explanatory variables, and permitted the identification of several redundant variables. The first four orthogonal principal components explain 28.58%, 15.18%, 10.46%, and 7.62% of the variation among WSC station/drainage attributes, respectively. This suggests that a large portion of the variation in flow metrics remains unaccounted. Plotting individual streams along the first two axes reveals strong clustering by EcoProvince (Fig. 4). This indicates that EcoProvinces have unique combinations of flow attributes, independent of the EcoProvince categorization. EcoProvince, EcoSection, MAD and MedAD all loaded heavily on the first axis, while EcoSection also loaded heavily on the second axis (Fig. 5a). Other variables appeared less important in this dataset. Clustering by EcoProvince is also strong if we repeat the ordination analysis excluding EcoProvince from the list of explanatory variables (not shown).

Table 5. Results of linear regression analyses of area-standardized flow metrics on drainage area, conducted on each EcoProvince separately for each flow metric.

EcoProvince	Flow metric	Slope	r^2	F-value	P
CAM	MAD	-0.0389	0.0129	0.2215	0.6440
	sumCP	0.3582	0.4031	13.160	0.0021*
	winCP	-0.2461	0.3394	8.7330	0.0089*
SIM	MAD	0.0159	0.0021	0.0288	0.8678
	sumCP	0.1306	0.0342	0.4953	0.4931
	winCP	0.0093	0.0002	0.0030	0.9572
SI	MAD	-0.1222	0.1568	4.0910	0.0555
	sumCP	-0.0409	0.0107	0.2367	0.6314
	winCP	0.0645	0.0280	0.6339	0.4344

The correlation circle for the continuous variables (Fig. 5b) revealed strong correlations between MAD and MedAD (Pearson's r = 0.99), indicating that the two variables are redundant in our analysis. All subsequent analyses were thus conducted on MAD only. Lake area and lake area divided by drainage area were also highly correlated, and since dividing by DA did not appear to change the interpretation of this variable, we only retained lake area in all further analyses. Relief and DA were also highly correlated (Fig. 5b), which is expected since large drainage basins would also be expected to span a greater range of elevations. Since the previous analysis showed that scaling with DA was not important, and since relief seem to explain more variation than DA alone (i.e. the arrow is longer), we chose to remove DA from all further analyses.

Multiple regression analysis

The results of the stepwise multiple regression analysis suggest that different environmental variables influence flow in different ways in different EcoProvinces (Table 6 and 7). Surprisingly, however, EcoSection was not generally a good predictor of flow metrics, suggesting that regional variation between EcoProvinces was much more consistent than variation within them. The exception was the Southern Interior Mountains, where it was selected in the final models for all three flow metrics (some of the coefficients also tended to be significant). In the Southern Interior Mountains, presence of glaciers was also selected in two models (Table 6) and the coefficients were significant. Presence of glaciers was an important predictor of flow in the Coast and Mountains EcoProvince, but only for winter and summer critical period flows (Table 5). The coefficients were relatively large and significant (Table 7), but the effect was opposite in the two seasons (positive in summer but negative in winter). Elevation, relief, lake area and state (regulated or unregulated) were sometimes included in the final models, but their coefficients were small and generally non-significant (Table 6). In order to help visualize the relationship between flow metrics and the different variables, we also plotted all three flow metrics against all the variables that were retained in the models for each EcoProvince separately (Fig. 6).

Table 6. Summary of the stepwise multiple regression procedure showing the models selected and their AIC values.

EcoProvince	Flow metric	Model selected (backwards)	AIC
CAM	MAD	Intercept only	184.29
	SumCP	Glaciers+Mean_elevation+relief+state	157.36
	WinCP	Glaciers	157.23
SIM	MAD	Glaciers+EcoSection	94.57
	SumCP	LakeArea+Glaciers+EcoSection	70.92
	WinCP	LakeArea+Relief+EcoSection	51.52
SI	MAD	Mean_elevation+lakeArea+EcoSection+state	99.15
	SumCP	Intercept only	72.737
	WinCP	LakeArea	20.517

Table 7. Details of the models selected using the stepwise multiple regression procedure. Values of the linear regression coefficients are provided for the selected variables from the final models.

EcoProv.	Flow metric	Sec1	Sec2	Sec3	Glacier (present)	Elev.	Lake Area	Relief	State (Reg)	Intercept
CAM	MAD									59.98**
	SumCP[1]				36.72**	-0.077		0.023	-11.79	38.37**
	WinCP				-23.08**					37.27**
SIM	MAD	1(CCM)	-14.15** (SFH)	-11.81** (SPR)	5.97*					29.19**
	SumCP	1 (CCM)	-6.96** (SFH)	-2.17 (SPR)	3.22*		0.130			9.37**
	WinCP	1 (CCM)	-3.17** (SFH)	-2.50 (SPR)			0.089	0.001		3.56*
SI	MAD	1 (NOH)	-2.48* (WOU)	NA		0.007*	-0.09		-1.58	0.006
	SumCP									1.6**
	WinCP						-0.02			0.81**

Significance codes: ** < 0.001; * < 0.05

1 sumCP = 36.72 (if glacier present) -0.077(Elevation) + 0.023(relief) – 11.79 (if regulated) + 38.37
0 (if glacier absent) 0 (if regulated)

DISCUSSION

Discrepancies between HydroMaster and WSC data

We did not find serious discrepancies between HydroMaster and WSC data for the subset of HydroMaster data that we analyzed (59 of 1024 stations, which we consider to be broadly representative of the rest of the database). When these did occur, they were due to two main causes. First, some of the estimates of drainage area differed. When they did, it is because drainage areas were re-calculated using iMapBC when DA values reported by WSC were clearly in error (R. Ptolemy, BC Ministry of Environment, personal communication Dec. 2012). It is unclear whether this was done in a systematic way (i.e. whether the DA values reported by WSC were always checked or how frequently they are in error). We did not check all DA values used in the current analysis because of time limitations, but we did verify a few that were corrected and the HydroMaster estimates were always more accurate than the ones reported by WSC (assuming the contour lines indicating watershed boundaries in iMapBC are accurate).

A second source of discrepancy was observed with the flow metrics of a majority of regulated streams. This discrepancy occurs because the regulated flows reported by WSC were naturalized utilizing available historical information on the regulation regime of those streams (R. Ptolemy, BC Ministry of Environment, personal communication Dec. 2012). For example, if records show that a stream became regulated in year 'X', and WSC data is available prior to year 'X', only the data from before regulation occurred was used to estimate MAD. In other cases where this was not possible, naturalized flows were reconstructed using stations upstream of the regulation site (if available). MAD at the station downstream of the regulation was then estimated by extrapolating the discharge of the upstream station to the DA of the downstream station to estimate MAD. Yet in other cases, such as run of river diversions that only affect flows in specific times of year, MAD was not adjusted because run of river operations do not change discharge, but only its timing. It should be noted, however, that in those cases, seasonal flows (e.g., sumCP, winCP) might be affected. In other cases, where no external information was available, the estimates of regulated MAD from WSC were used 'as is'. These transformations were applied in a systematic way to every regulated stream included in the HydroMaster analysis (R. Ptolemy, BC Ministry of Environment, personal communication Dec. 2012). Such adjustments most likely explain some of the differences in mean flow metrics observed when regulated streams are included (i.e. Tables 2 and 3), but not when they are excluded (Fig. 2).

The approach for calculating natural flows in regulated streams outlined above has a number of potential problems that should be pointed out. First, each regulated stream received the flag 'regulated' in the analysis regardless of the method of adjustment (or lack thereof). This could be problematic because some rivers will have very precise naturalized flow based on good data (e.g. Cheakamus River) while data limitations mean that others may not have been naturalized with much confidence. Sub-categorization of regulated streams according to the quality and/or method of flow naturalization, and documentation of the method of naturalization for each regulated stream, would therefore seem valuable. Second, in cases where flows are extrapolated from upstream stations, it may be argued that those data points represent pseudo-replication. While this is definitely true if both stations are used as independent data points in a statistical analysis, it is nevertheless a reasonable approximation if one were interested in the MAD value at a specific site. It may also be unwarranted to include such stations in analyses such as ours where the effects of physical variables on flow are examined. It should be noted, however, that this would only be problematic in cases where the section of streams between the two stations differ dramatically in one or more physical variables (e.g., a glacier-fed tributary entering between the two stations, large relief change occurring at a greater gradient than in the rest of the watershed, presence of a large storage lake between the stations, etc.). In most cases this effect should be fairly minor. Finally, the reconstructed flows are in many cases not associated with a specific time period, and they therefore cannot be included in an analysis conducted on a common period of record such as the one presented here.

Despite those potential problems, the adjustments made to the MAD estimates appear valuable for at least certain types of analyses. Most importantly, in certain parts of the province, few rivers are unregulated, and large enough sample sizes for statistical analyses of flow sensitivity are impossible if regulated rivers are not included. It thus appears like a very reasonable course of action in such cases to use all the available information to reconstruct flows, although better documentation of the specific adjustments that were made to naturalize individual streams would improve clarity of the database.

Data interpolations in HydroMaster

The main data extrapolation in HydroMaster appears to be interpolating missing winter flow at stations where winter flows were not recorded (i.e., for WSC gauges where only summer data is collected). Data from missing months was estimated based on the ratio of monthly area-specific flows to MAD for fully gauged stations in the same EcoSection or EcoProvince, in order to allow calculation of MAD and associated statistics at the station with missing monthly flows. While this seems like a reasonable thing to do because it allows inclusion of data from stations that would otherwise be discarded (and WSC gauge coverage in some regions is a serious problem), it would also constitute pseudo-replication if regional means were calculated with this data (i.e., confidence intervals around means would be tighter than expected), although the mean should not be distorted. Note that none of the streams included in the current analysis suffered from this problem because we only used streams for which there was 10 years of data for all months in the common period of record.

Potential for systematic bias in MAD using HydroMaster

Our analysis has revealed a systematic bias between the estimate of MAD from the larger HydroMaster dataset and the estimate from a common period of record with a minimum 10 years of reported data for the subset of sites analyzed across a range of EcoSections. However, the period of record we used (1976-1995) was a period of relatively low flow for the 20th century (Eaton et al. 2002), and it is therefore not surprising that MAD estimated from this period would be lower than the long term MAD estimated in the HydroMaster dataset. As noted earlier, however, the size of the bias is also fairly small ($1.18\ \mathrm{l \cdot s^{-1} \cdot km^{-2}}$ compared to a mean difference in runoff between EcoProvinces of $37.17\ \mathrm{l \cdot s^{-1} \cdot km^{-2}}$). In short, it seems unlikely that management decisions made on the basis of the MAD estimates from the HydroMaster dataset would be biased in any meaningful way. The analysis, however, makes it evident that the period of record does influence estimates of LTMAD.

HydroMaster also had a relatively high correct classification rate for identifying streams as flow sensitive based on Tennant criteria (i.e., streams with summer or winter flows below 20% MAD; Fig. 3). This indicates that the classification of streams as flow sensitive is fairly robust to a variable period of record, as used in HydroMaster. Streams that were likely to be misclassified were near the 20% MAD threshold (Fig. 3), as would be expected. While HydroMaster did tend to overestimate flows (in terms of percent MAD) relative to the common period of record, this may be related to the common period of record being relatively dry as noted above, which could increase relative low flow discharge estimates if MAD decreased more than low-flows. However, some of the HydroMaster estimates that showed large underestimates of low flows (i.e., data points that deviate well below the 1:1 line in Fig. 3) relative to the common POR may in fact be correct, while the common period of record is effectively in error because the flows have not been naturalized. MAD estimated for the common period of record post-dated many diversions or other hydrologic alterations, so that relative sumCP flows (as a percent of MAD) are inflated by a reduced estimate of MAD; in contrast, the longer period of record in HydroMaster allowed some MAD estimates for regulated streams to be based on pre-diversion values, resulting in a much higher MAD estimate and consequently a smaller relative low flow (as a percent of MAD). Insofar as pre-diversion flows are the appropriate reference condition for estimating low flows as a percent of MAD for many regulated streams, the HydroMaster low flow estimates are in fact more appropriate as benchmarks than estimates from the common period of record. In this sense sites considered “misclassified” by HydroMaster in this comparison may in fact be correctly identified as flow sensitive. Regardless, discrepancies between the two data sets with respect to the 20% MAD threshold are relatively small, and become largely irrelevant at flows well below 20% MAD.

Use of median annual discharge as an alternative to MAD

The use of mean annual discharge instead of median annual discharge has been criticized because MAD is heavily influenced by extreme values that tend to make it less conservative in most cases, i.e., in terms of Instream Flow Thresholds the mean will be biased above the median in proportion to the relative variation in the annual hydrograph (i.e., frequency of extreme high flows). For streams that experience greater flow variation, MAD will be proportionally higher than MedAD, so that application of Tennant criteria will apply higher thresholds when lower ones might be more appropriate. Ideally, the median annual flow metric should be calculated using daily flow estimates (Instream Flow Thresholds). Daily data, however, is not available from the HydroMaster dataset because it only includes mean monthly flows collated from the WSC online database. The flow metric based on the median flow that we used, therefore, was merely the mean of the median monthly discharges over the period of record. Perhaps not surprisingly, therefore, it did not differ in any meaningful way from MAD. This was evident from the factor analysis, which revealed a very tight correlation (Pearson's $r = 0.99$) between the two metrics. We conclude from this result that, while MedAD values calculated from daily flow data may be useful, the MedAD metric used in the present report (based on median of monthly mean flows) does not allow better inference of flow sensitivity than MAD. Irrespective of this, MAD also has the advantage over a median metric that fixed percentages are already associated (at least qualitatively, and in some case quantitatively) with habitat thresholds (Tennant 1976), while MedAD is not.

Ability of elevation and other variables to refine predictions

The lack of consideration of other physical predictor variables was identified as a potential limitation of the first estimation of regional mean flows based on the HydroMaster dataset. In the present analysis, we considered the effects of elevation and other physical variables such as presence of lakes and glaciers along with EcoProvince and EcoSection to determine whether these variables could account for additional variation in flow metrics. In brief, we found that inclusion of such variables could in some cases improve predictions of flow metrics, but only for some EcoProvinces.

Despite the fact that EcoSection appeared to be an important explanatory variable in the factor analysis, it was only selected as an important predictor in the final models in a few cases. In fact, it was mainly important in explaining flows in the Southern Interior Mountains EcoProvince. Moreover, the significant effect of EcoSection there was largely explained by the lower flows measured in the Selkirk Foothills EcoSection. This result points to a potential limitation of our analysis: if a single EcoSection differs considerably from other EcoSections and drives this patterns, it is possible that increasing the number of EcoSections included in the analysis would increase the likelihood of including an EcoSection of large effect. Thus, it should not be concluded from the present analysis that EcoSection is not a good predictor of flow in some cases. It is also possible that some EcoProvinces have more homogeneous EcoSections than others, e.g. Coast and Mountains. This is certainly possible and would not necessarily be reflected in the EcoSection designation.

Contrary to expectations, elevation (and relief) did not appear to be important explanatory variables in most cases within the subset of Ecoprovinces analyzed. Elevation was only retained in two of the final models, and the coefficient was only significant in the Southern Interior for MAD. Similarly, relief was retained in two of the final models, but in neither case was the coefficient significant. This either suggests that elevation is not very important in driving flows (which is unlikely to be the case), or that other variables like EcoProvince and EcoSection capture most of the meaningful variation in elevation. On the basis of our analysis, it is therefore not recommended that elevation be used as a predictor of flow, unless this data is already readily available with minimal investment of effort. However, this conclusion should also be tempered by the consideration that variables that are insignificant in this dataset might be significant predictors in other regions.

Presence of glaciers had a large influence on flows, but only in the Coast and Mountains EcoProvince. While presence of glaciers had no effect on MAD, it was the most important variable explaining variation in critical low flows in summer and winter. Not surprisingly, presence of glaciers had a positive effect on flows in the summer, with streams with a glacier in the catchment having flows on average 36.72 $l \cdot s^{-1} \cdot km^{-2}$ higher than streams without glaciers irrespective of the EcoSection. This represents a proportional difference of 396%, which is obviously substantial. Interestingly, this effect of glaciers was reversed in the winter, with streams with a glacier in the catchment having flows on average 23.08 $l \cdot s^{-1} \cdot km^{-2}$ lower than streams without glaciers irrespective of the EcoSection. Again, this represents a proportional difference of 262%. We hypothesize that this negative effect is related to the negative effect of elevation on winter low flows: glaciers are more likely to be present at higher elevation, and winter precipitation is also more likely to fall as snow at higher elevations. The reversed effect of glaciers on flow in winter and summer probably also explains why presence of glaciers has no influence on MAD, since their positive effect in one season is offset by their negative effect in the other. Glaciers also appeared to have a significant positive influence on flows in the Southern Interior Mountains, but it was proportionally much smaller than in the Coast and Mountains EcoProvince. Here, the effect was only significant on the summer low flows, which probably drove their significant effect on MAD.

Despite the fact that it was retained in some of the final models in the multiple regression analysis, lake area did not appear to correlate strongly with any of the chosen flow metrics. Indeed, the coefficients tend to be small and non-significant, perhaps because of the variable nature of lake effects depending on their location in a drainage. Lake area therefore does not appear useful as a generic predictor of flow in the analyzed data set, although it may likely be a useful predictor of flow attributes in individual streams where the attributes of the lake (location, size, volume) are more specifically accounted for.

Another unexpected result was that whether a stream was regulated or not did not appear to have a large influence on annual flow, even if we did not correct for it. Indeed, regulated vs. unregulated state was only selected in two of the final models, and in none of the cases were the coefficients significant (although in the Coast and Mountains EcoProvince regulated streams had summer flows on average 11.79 $l \cdot s^{-1} \cdot km^{-2}$ lower than natural streams, indicating that the expected footprint of water extraction during the critical low flow period was strong enough to be retained in the model in only one of 3 EcoProvinces). This would perhaps suggest that, on average and on large spatial scales, water withdrawals have relatively small effects on flows across the landscape of stream populations included in the analysis. Alternatively, regulation may be highly correlated with other variables in the statistical models, or streams classified as unregulated may in fact be experiencing some level of unaccounted hydrologic alteration, contributing to underestimation of the effects of flow regulation. This latter possibility is most probable, since designation of streams as "regulated" assumed perfect knowledge of water use, which is unrealistic. Similarly, classifying hydrologic alteration as a binary variable (present/absent) represents a relatively weak statistical test for what is really a gradient of hydrologic alteration, which should ideally be represented as a continuous covariate. Regardless, low statistical significance does not exclude the likelihood that water withdrawals have an important influence on the flow of particular streams or regions, or that withdrawals can lead to extreme low flows events of short duration that would not be captured in the present analysis. Another consideration is that regulated streams include streams with both diversions and augmentations of flows. This should not influence the conclusions of the present study, however, since only Oyama River in the Southern Interior EcoProvince was flagged as augmented.

The variables selected did comparatively poorly at explaining variation in flows in the Southern Interior EcoProvince. Most selected models in this EcoProvince included few variables, and when they did the coefficients showed small effect sizes and were not significant, or only marginally so. This could reflect the reduced environmental variation among streams in the Southern Interior EcoProvince in terms of both flow (e.g. Fig. 1) and predictor variables (e.g., none had glaciers or a large contrast in vertical relief). This is an unfortunate result since prediction of flows for dry regions like the Southern Interior is a greater priority for sustainable water management. On the other hand, it also suggests that predictions of MAD from only DA are more accurate in such regions than in others, and may be sufficient in most cases, especially since EcoSection itself was a significant predictor of runoff. The only other potentially useful variable was elevation, which explained a small but significant amount of variation. Predictions of flow for the Southern Interior EcoProvince could thus conceivably be improved by inclusion of this variable.

In summary, some of the physical variables examined in the current analysis appear to have potential as useful predictors of flow. Further refinement of regional runoff relationships with local predictor variables might therefore improve predictions based solely on EcoSection (except in the Southern Interior Mountains where EcoSection was the best predictor variable). It should be noted, however, that in most cases the predictor variables explained only a small portion of the variation (except for presence of glaciers in the Coast and Mountains EcoProvince). If such refined predictive tools are deemed useful, the present analysis also makes it clear that their implementation will need to be region and/or season-specific. In the meantime, this review indicates that EcoProvince or EcoSection means summarized in HydroMaster are a useful and credible starting point for generating estimates of areal runoff and low flow metrics in ungauged streams for screening or preliminary management purposes.

ACKNOWLEDGMENTS

Thanks to Robin Pike and Scott Babakaiff for providing helpful comments on an earlier version of this report, and to the UBC Biodiversity Research Centre for providing funding for a BRITE internship to J-S Moore. Special thanks to Ron Ptolemy for providing the data and support to carry out this work.

REFERENCES

Demarchi, D. (2011) Ecoregions of British Columbia, 3rd edition. (Ecosystem Information Section, M. O. E., ed.). pp. Victoria, B.C.

Eaton, B., Church, M. & Ham, D. 2002. Scaling and regionalization of flood flows in British Columbia, Canada. *Hydrological processes* **16**: 3245-3263.

Jowett, I. G. 1997. Instreamflow methods: A comparison of approaches. *Regulated Rivers: Research and Management* **13**: 115-127.

King, J., Brown, C. & Sabet, H. 2003. A scenario-based holistic approach to environmental flow assessments for rivers. *River Research and Applications* **19**: 619-639.

Obedkoff, W. 2000. Interior Community Watershed Streamflow Inventory. Resources Inventory Branch, B.C. Ministry of Environment,

Obedkoff, W. 2003. Streamflow in the Lower Mainland and Vancouver Island. Aquatic Information Branch, B.C. Ministry of Environment,

Ptolemy, R. & Lewis, A. (2002) Rationale for multiple British Columbia instreamflow standards to maintain ecosystem function and biodiversity. pp. Draft for agency review prepared for Ministry of Water Land and Air Protection and Ministry of Sustainable Resource Management.

R Development Core Team (2012) R: A language and environment for statistical computing. R Foundation for Statistical Computing Vienna Austria.

Tennant, D. L. 1976. Instreamflow regimens for fish, wildlife, recreation and related environmental resources. *Fisheries* **1**: 6-10.

Whitfield, P. H., Moore, R. D., Fleming, S. W. & Zawadzki, A. 2010. Pacific Decadal Oscillation and the Hydroclimatology of Western Canada-Review and Prospects. *Canadian Water Resources Journal* **35**: 1-28.

Figure 1. Box plots comparing median values of the three flow metrics among the three EcoProvinces.

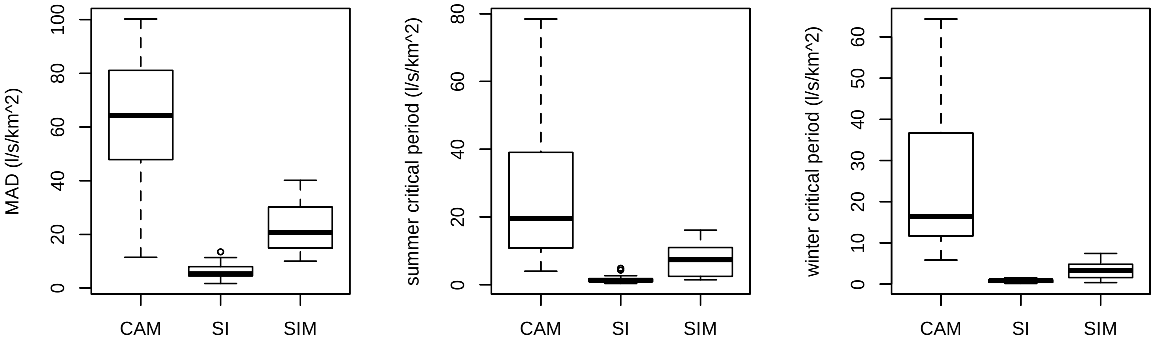


Figure 2. Comparison of long term MAD from the HydroMaster dataset to that used in the present analysis, which is based on a common period of record and a minimum of ten years of data. (a) Correlation between MAD from HydroMaster (x-axis) and MAD from the present analysis. Note the outlier stream, Coquitlam, which was removed from statistical analyses. (b) Histogram of the results of subtractions of HydroMaster MAD from MAD_POR showing that there is a systematic bias such that MAD from the HydroMaster overestimates MAD.

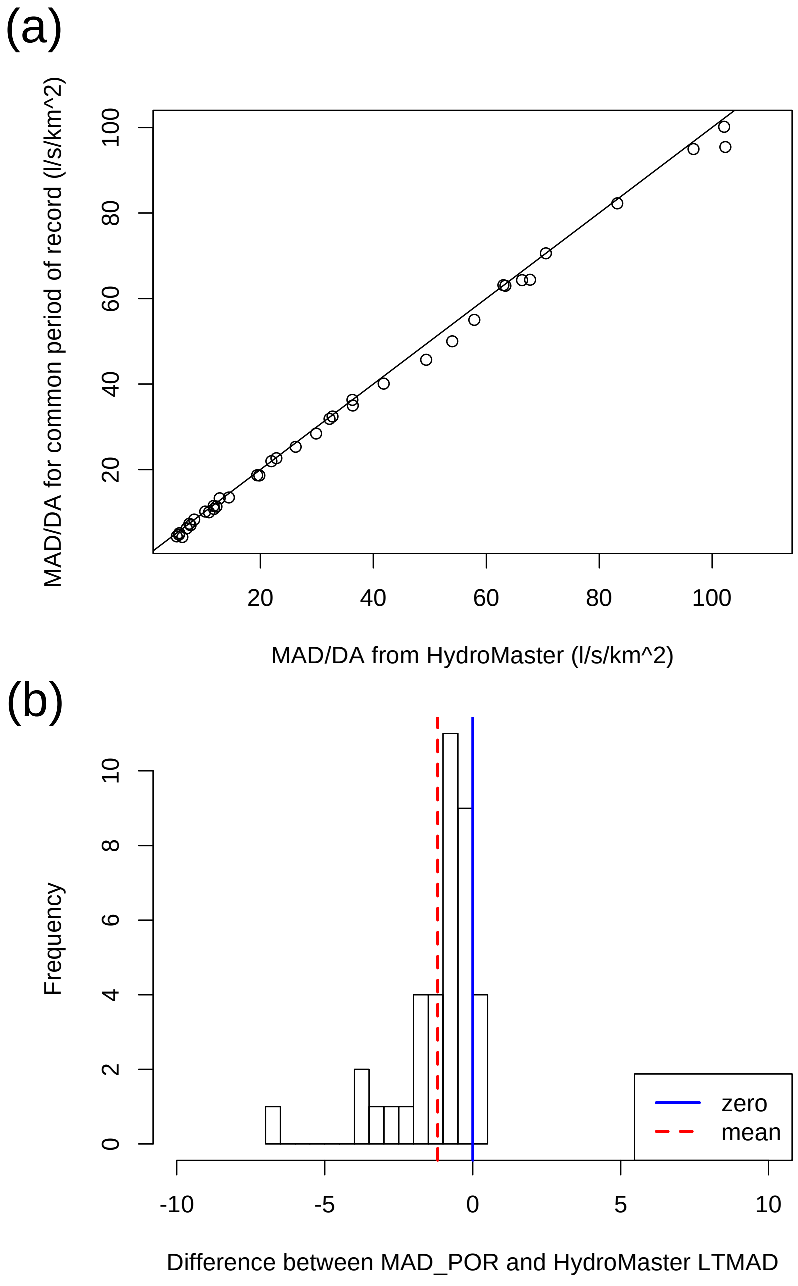


Figure 3. Plots of summer (upper panel, a) and winter (lower panel, b) critical period low flow (% MAD) estimates from HydroMaster (with a highly variable period of record) against low flow estimates for a common period of record (the x-axis). Circles represent streams that were correctly classified by HydroMaster as having low flows below 20% MAD, relative to estimates from the common period of record; squares were misclassified sites, and triangles are sites correctly predicted by HydroMaster to have low flows in excess of 20% MAD.

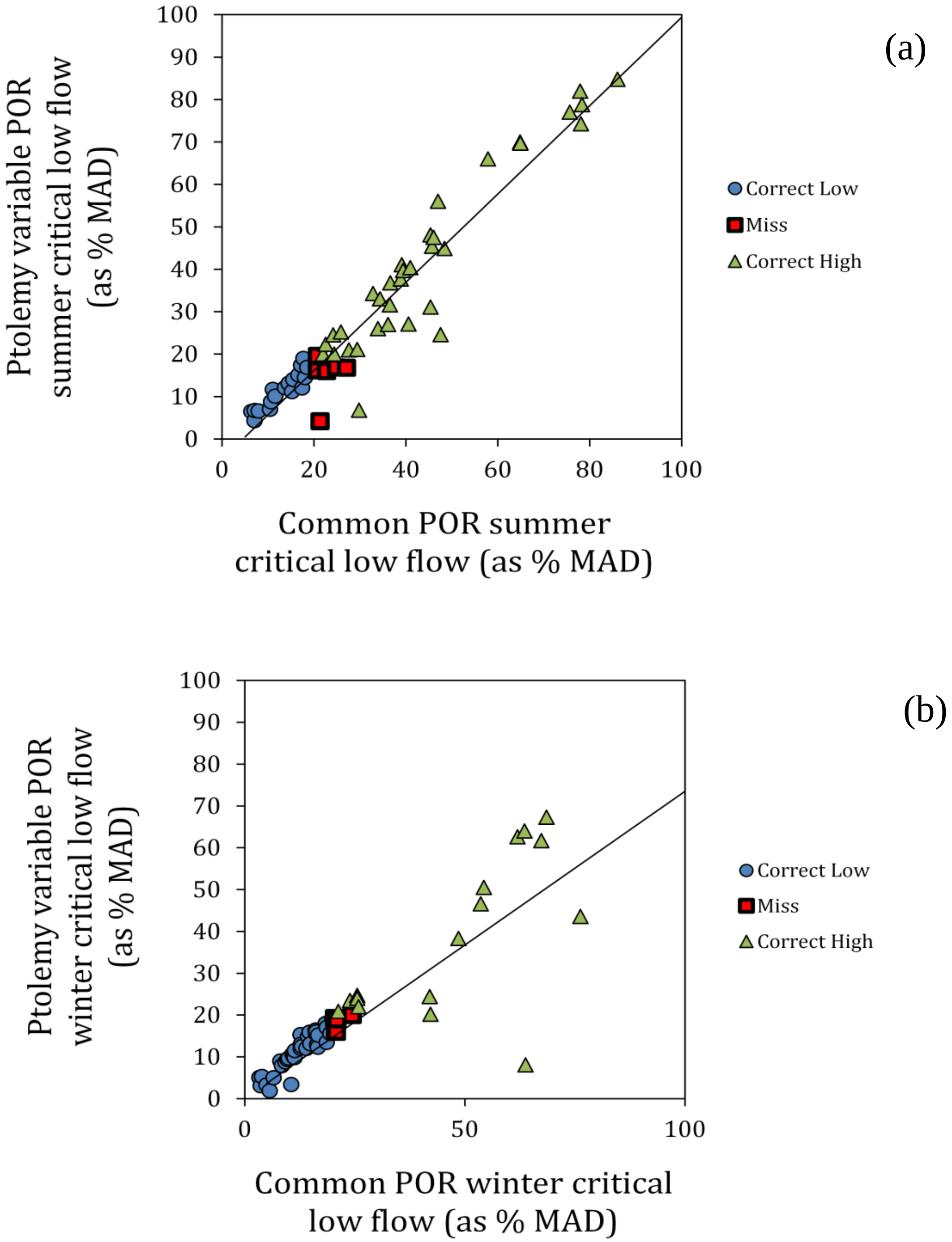
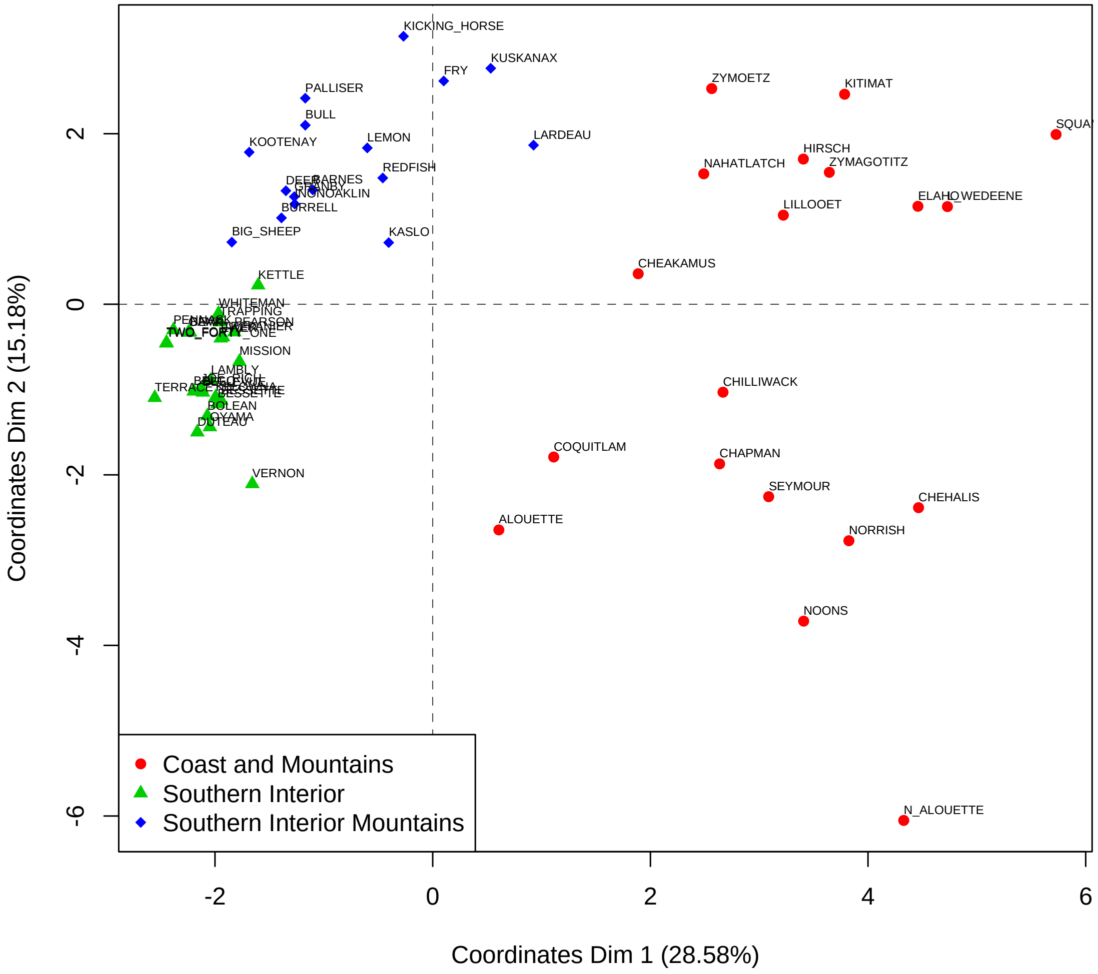


Figure 4. Results of the factor analysis. Each individual stream is plotted along the two orthogonal axes that explain most of the variation. The streams are color-coded according to their EcoProvince to make the importance of that variable apparent.



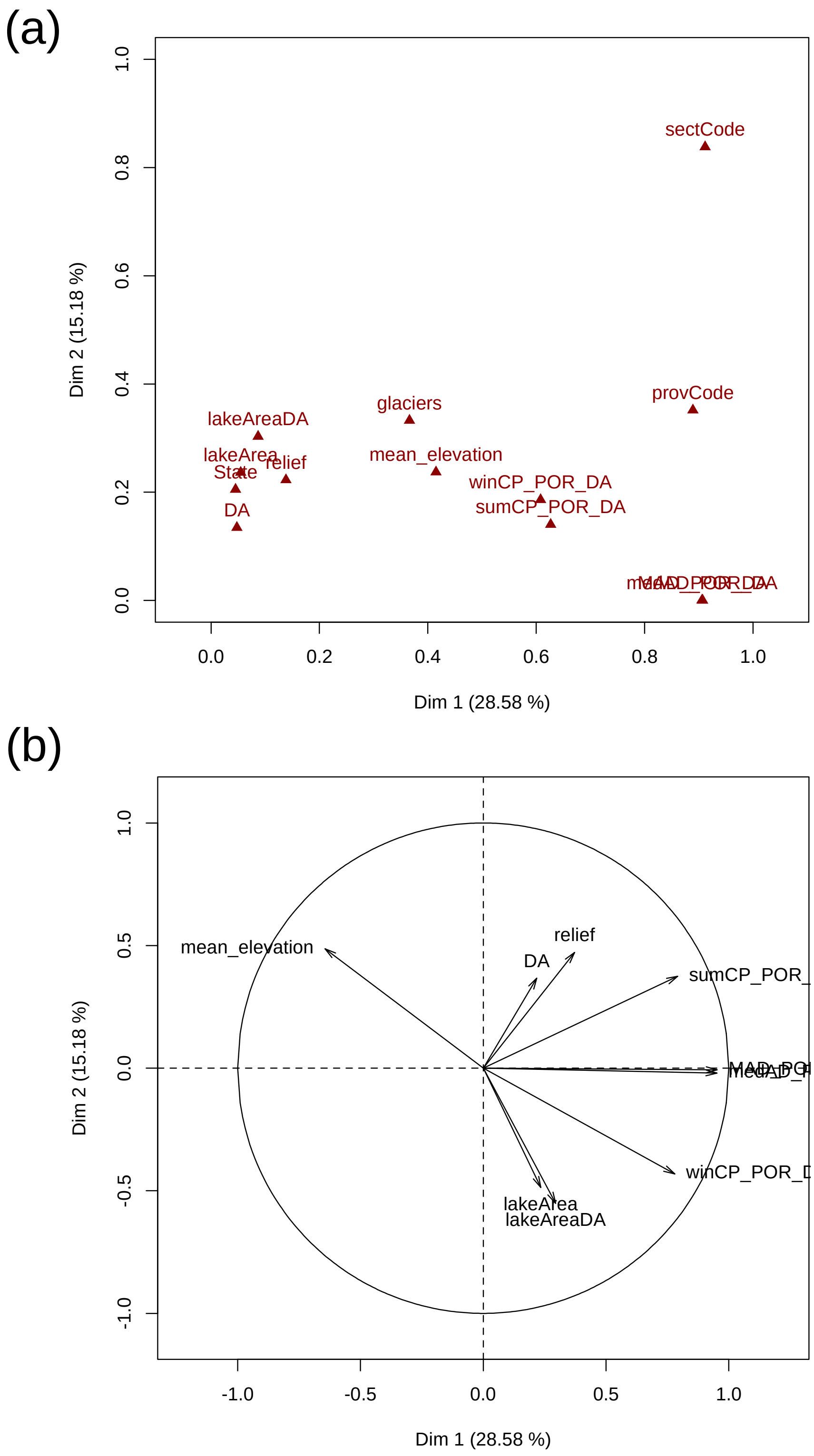
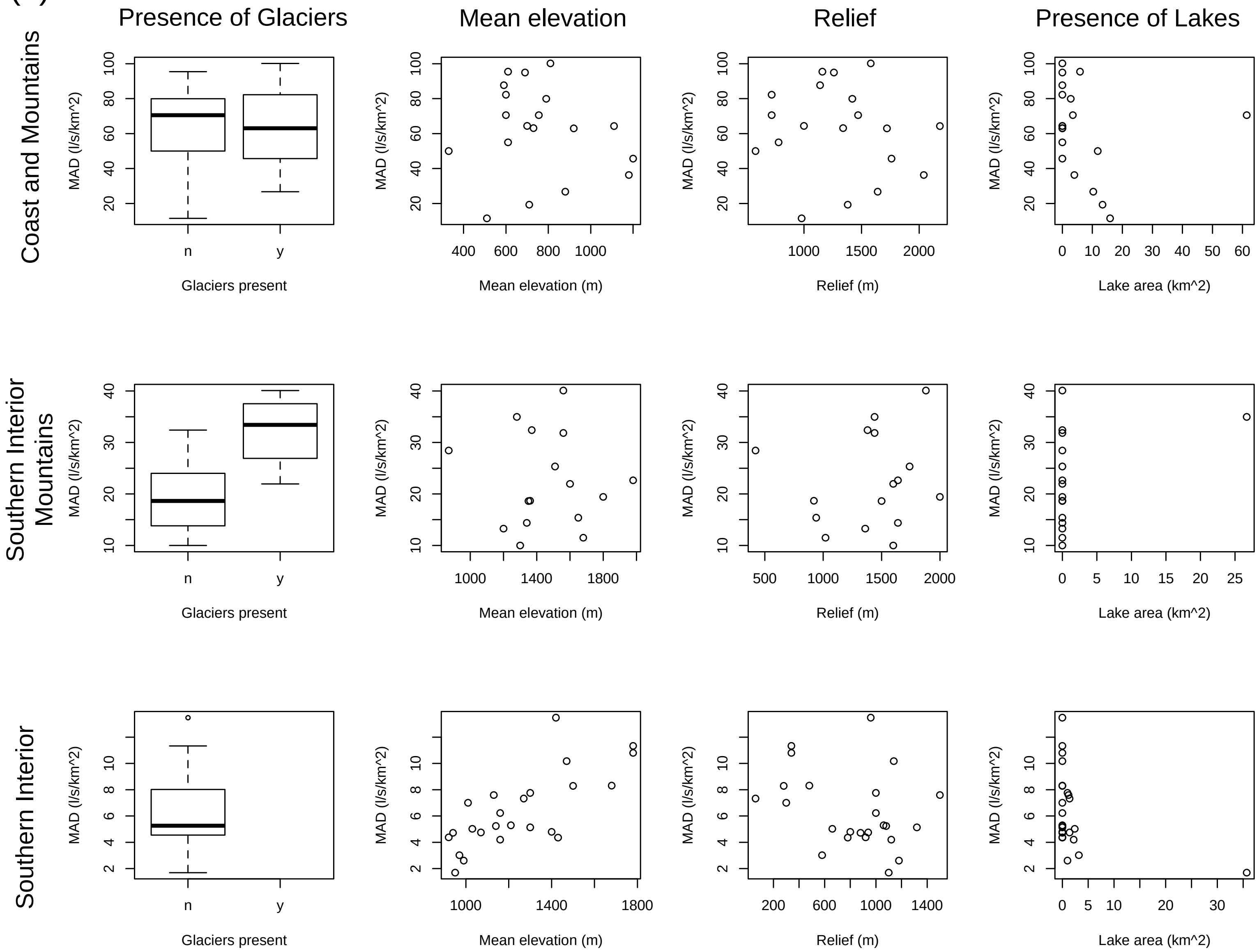


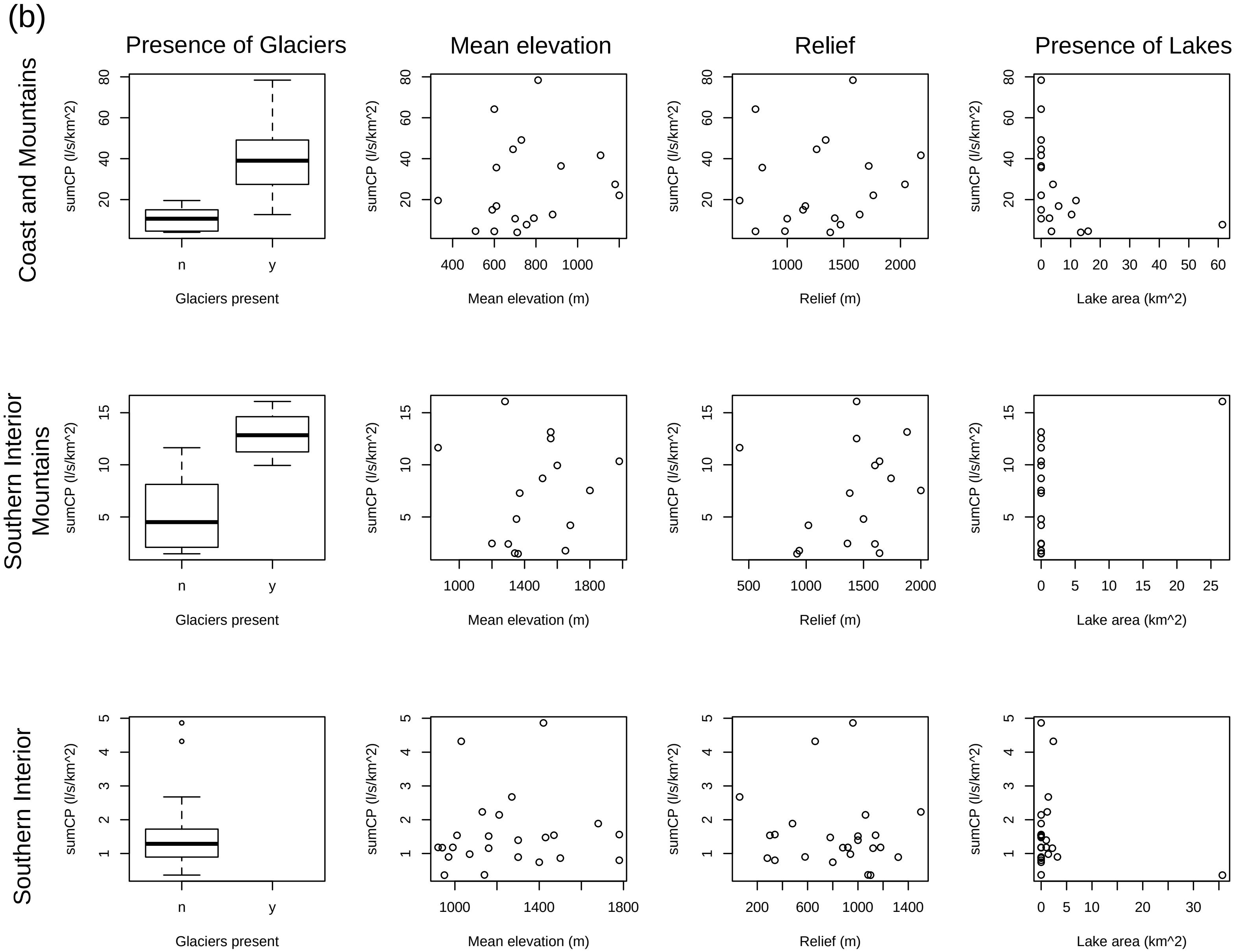
Figure 5. Results of the factor analysis showing loadings of each variable and correlation among continuous variables. (a) Loadings of all variables along the first two, most important axes. Variables ‘sectCode’ and ‘provCode’ are EcoSection and EcoProvince respectively, and both load heavily on both axes. (b) Correlation circle showing correlation among the continuous variables only. Arrows pointing in the same direction show variables that are highly correlated and thus redundant (e.g. lake area and drainage area-scaled lake area).

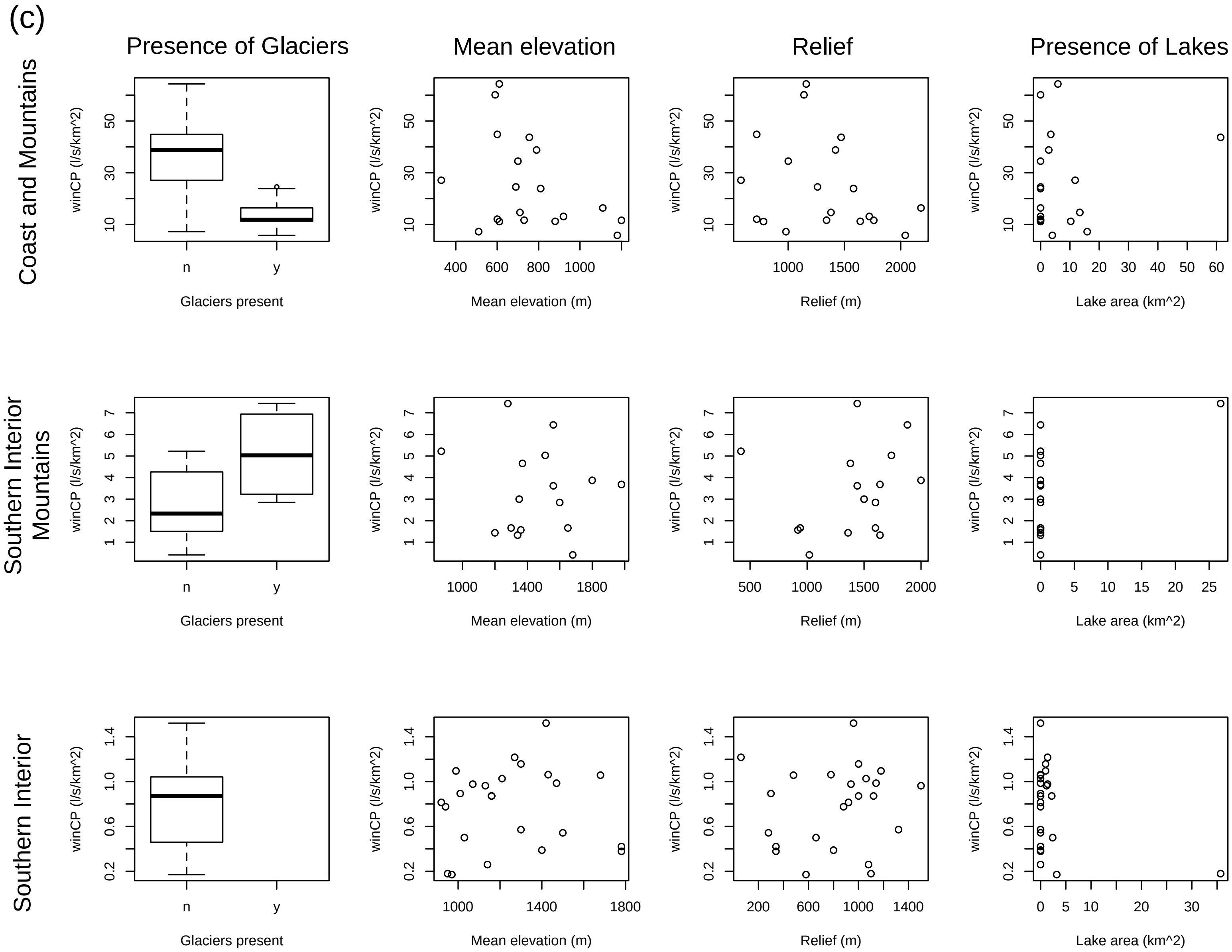
Figure 6. (next three pages) Plots showing the relationships between flow metrics and the four physical variables that were commonly selected in the stepwise multiple regression approach. In all cases, the columns are the four different physical variables tested and the rows are the three different EcoProvinces included in the analysis (because the analysis was run on each EcoProvince separately). Plots are presented for the three different flow metrics examined, where the vertical axes are: (a) MAD, (b) summer critical period low flow, and (c) winter critical period low flow.

(a)



Presence of Glaciers
Mean elevation
Relief
Presence of Lakes
Coast and Mountains
Southern Interior Mountains
Southern Interior
MAD (l/s/km^2)
Glaciers present
n
y
Mean elevation (m)
Relief (m)
Lake area (km^2)
20
40
60
80
100
400
600
800
1000
1500
2000
0
10
20
30
40
50
60
1400
1800
500
5
15
25
2
4
6
8
200


(b)
Presence of Glaciers
Mean elevation
Relief
Presence of Lakes
Coast and Mountains
Southern Interior Mountains
Southern Interior
sumCP (l/s/km^2)
20
40
60
80
5
10
15
1
2
3
4
5
n
y
Glaciers present
400
600
800
1000
1400
1800
Mean elevation (m)
200
500
600
1000
1500
2000
1400
Relief (m)
0
5
10
15
20
25
30
40
50
60
Lake area (km^2)


(c)
Presence of Glaciers
Mean elevation
Relief
Presence of Lakes
Coast and Mountains
Southern Interior Mountains
Southern Interior
winCP (l/s/km^2)
Glaciers present
Mean elevation (m)
Relief (m)
Lake area (km^2)
n
y
10
30
50
1
2
3
4
5
6
7
0.2
0.6
1.0
1.4
400
600
800
1000
1400
1800
500
1500
2000
200
0
5
10
15
20
25
30
40
50
60