REQUIRED KNOWLEDGE CREATION

WHAT DELIVERABLES ARE NEEDED?

- Develop specific understanding of how different hydrologic regimes within BC Hydro's geophysical portfolio have responded in the past and might respond in the future
- Daily or monthly hydrographs needed for some long-term planning studies: require watershed-specific inflows as input to reservoir operation or planning models
- Reliable quantitative assessment of uncertainty: both total uncertainty in future hydroclimatic conditions, and the sources of that uncertainty with their relative importance
- Must also think about some issues that are ~somewhat peculiar to BC
    - massive glacier change and consequent requirement for glacier dynamical modelling?
    - climate-influenced pest infestations (e.g., mountain pine beetle) & forest hydrology change?

BC hydro
FOR GENERATIONS

hydrology & technical services

STEPS TAKEN

THREE PATHS TO ENLIGHTENMENT (WE HOPE)

- Internal research work
- Alliance with Pacific Climate Impacts Consortium (PCIC)
- Partnership with Western Canadian Cryospheric Network (WC^{2}N)

BC hydro
FOR GENERATIONS

hydrology & technical services

INTERNAL RESEARCH

TECHNICAL ANALYSIS PERFORMED IN-HOUSE BY H&TS

- BC Hydro has been paying attention to the issue since early 1990s
- Examples of recent work:
    - UBC Watershed Model assessments of climate change impacts to Arrow, Revelstoke, Mica, and Peace – delta downscaling method – used an assumed glacier cover change – averaged anomalies from 15 GCMs
    - Ongoing assessment of historical changes in annual inflow volumes for ~21 reservoirs with quality-controlled data – linear modelling w/ S:N analysis – nonparametric (Spearman rank) correlation using bracketing analysis w/ and w/o deserialization – possibly low-pass filtering

BC hydro
FOR GENERATIONS

hydrology & technical services

PCIC

ALLIANCE WITH PACIFIC CLIMATE IMPACTS CONSORTIUM

- Applied research consortium based at University of Victoria
- BC Hydro support began in late 2006
- Pledge $200K/yr
- In between a “hands-off” pure research grant and a “hands-on” consulting contract
- We have a measure of direct influence through membership in PCIC’s Program Advisory Committee (PAC) and the project’s Technical Advisory Committee (TAC)

PCIC

PROJECTS COMPLETED OR UNDERWAY

- Climate Overview Report completed 2007: thorough synthesis of existing & some new research on historical climatic/hydroclimatic variability & change in British Columbia – setting the context
- Several modelling studies underway to make projections of future conditions
- Fully distributed semi-physical VIC model or modified VIC (glaciers + autocalibration algorithms):
    - Williston (northeast)
    - Canadian Columbia (southeast)
    - Upper Campbell (southwest)
- Fully distributed heavily physical WaSIM-ETH model: Columbia at Donald
- PCIC has completed much modelling & other work w/ other partners too, e.g., community outreach, Fraser River studies for MoE, (partial) climate downscaling subcontractor in WC^2N Mica study

WC^{2}N

ALLIANCE WITH WESTERN CANADIAN CRYOSPHERIC NETWORK

- Fundamental & applied research consortium, nominally based at University of Northern British Columbia but with active members across western Canada & US
- Focus on snow & glacier science with heavy emphasis on climate variability & change
- Launched in 2005 with BC Hydro support from the start: in-kind support through $Ms' worth of data, initial cash contribution of $10K, separately contracted Mica study for $127K
- Somewhat hands-off academic research relationship but have some direct influence on direction through participation on Board of Directors (BoD)
- Separate Mica study (technically, contracted to UBC) is more conventional consulting relationship in which participating WC^{2}N members report (and are accountable) directly to us

WC^{2}N

PROJECTS COMPLETED OR UNDERWAY

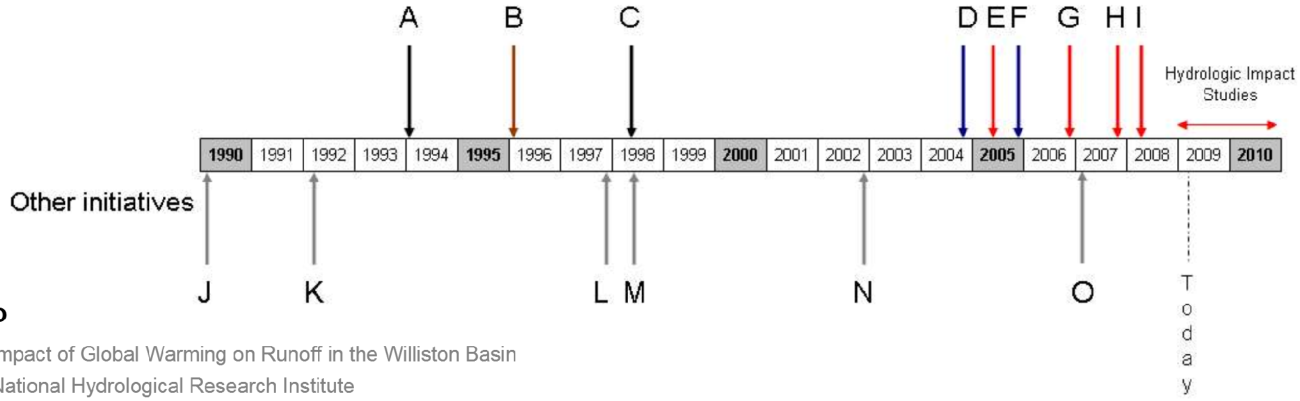
- Streamflow impact of climate change & associated glacier recession at Bridge: published as Stahl et al., Water Resour. Res., 2008. Used HBV, with empirical glacier recession model applied over a DEM. Even under uncertainty, streamflow losses will be significant, even if climate stabilizes.
- Streamflow impact of climate change & glacier recession at Mica: underway. Similar to Bridge work, but advanced glacier dynamical modelling employed, and reservoir inflows directly modelled. Important supporting project to work underway by PCIC and others on the Columbia basin.
- Additional WC^{2}N-related projects not directly linked to climate change work:
  - Comprehensive glacier mapping throughout western Canada: provided direct quantitative input to current operational forecast model recalibration, and to forecast model intercomparison study currently contracted to Conestoga-Rovers & Associates (“trickle-down” benefits)
  - Through BC Hydro-UBC Geography Grant Agreement: WC^{2}N personnel assessing MODIS satellite imagery for operational forecasting-related purposes.

BC hydro
FOR GENERATIONS

hydrology & technical services

BC HYDRO-SUPPORTED EXTERNAL WORK

CLIMATE CHANGE IMPACT ASSESSMENT TIMELINE



BC Hydro

A.	Impact of Global Warming on Runoff in the Williston Basin
B.	National Hydrological Research Institute
C.	Glacier Runoff Component Model and Application
D.	Impacts assessment by Corporate Environment
E.	Western Canadian Cryospheric Network (WC2N)- Letter of Interest
F.	Pacific Climate Impacts Consortium (PCIC) – Launched
G.	WC2N – Launched
H.	PCIC - Letter agreement for Climate Overview
I.	PCIC – Climate Overview
J.	PCIC - BCH Grant Contribution

Other

J.	Canada Green Plan
K.	UN Earth Summit
L.	Kyoto Protocol
M.	Canada signs Kyoto
N.	Canada ratifies the Kyoto protocol
O.	Western Climate Initiative

BC hydro
FOR GENERATIONS

hydrology & technical services

Climate Change studies update

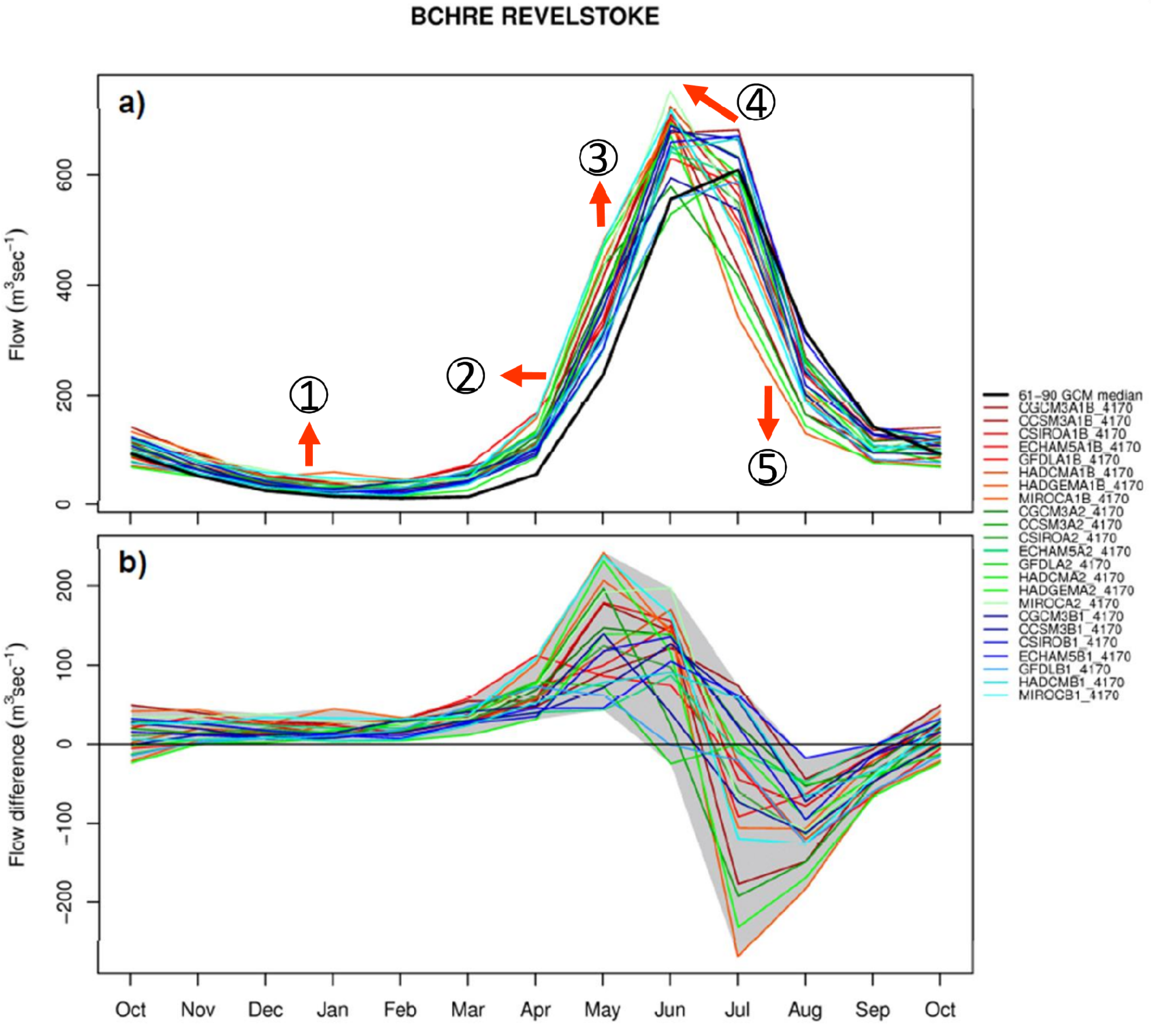
- BCH provided summary of climate change hydrologic impact studies for Canadian Columbia to Canadian and US CRT 2014 staff
- Studies
  - River Management Joint Operating Committee – Joint U.S. studies of entire Columbia and Pacific Northwest
  - Pacific Climate Impacts Consortium – Columbia above Birchbank
  - Mica basin glacier study

Results from Canadian Studies

- Inflow trends over 1984-2007
    1. Most of the historical trends are quite subtle
    2. No evidence for declining annual total water supply - some evidence for a modest historical increase
    3. Clear evidence for some changes in the seasonality of inflows
- Projected Annual trends in inflow volumes to 2050:
    1. Annual flow is projected to increase at the majority of the project sites in Canada
    2. Differences between emission scenarios are smaller than the combined inter-annual and inter-GCM differences

PROJECTED TRENDS IN MONTHLY INFLOWS

REVELSTOKE



PROJECTED TRENDS IN ANNUAL INFLOW VOLUME

1. Annual flow is projected to increase at the majority of the project sites
2. Differences between scenarios are smaller than the combined inter-annual and inter-GCM differences

	median annual precipitation* (%)	**median annual temperature* (ºC)**	**median annual flow* (%)**				
SRES	**Columbia region**		**MCA**	**REV**	**ARD**	**DDM**	**KLK**
B1	**+10**	**+1.8**	**+16**	**+13**	**+12**	**+13**	**+12**
A1B	**+7**	**+2.7**	**+22**	**+17**	**+16**	**+18**	**+13**
A2	**+5**	**+2.3**	**+17**	**+10**	**+8**	**+12**	**+7**
mean	**+7.3**	**+2.3**	**+18**	**+13**	**+12**	**+14**	**+11**

* relative to 1961-1990 baseline

BChydro

TECHNICAL REPORT

Climate change signal detection in BC Hydro reservoir inflows

15 December 2010

Sean W. Fleming PhD, PPhys, ACM, PGeo
Senior Hydrologic Modeller
Hydrology & Technical Services Group
Generation Resource Management

File: J:\PS\PSO\Hydrology\Climate Change\Climate Change Signal Detection\Write Up\Climatic Trend Analysis Report v12.doc

hydrology & technical services

Executive Summary

BC Hydro's ability to deliver abundant, low-cost energy in a socially, environmentally, and economically sustainable way is in large part the consequence of British Columbia's prodigious hydroelectric generation potential: rivers are the fuel that keeps the lights on in the province. However, river flows are sensitive to climatic variability and change. Rainfall, winter snowpack, and summer glacial melt are key drivers of BC Hydro's reservoir inflows. Thus, climatic drift might not only affect demand (by changing per capita winter heating or summer cooling loads, for instance), but also the supply of hydroelectric power available to British Columbians. As such, careful analysis of the possible impacts of climate change to our reservoir inflows is prudent, and indeed has been undertaken through a variety of projects in GRM.

Empirical trend analysis of historically observed data is a key component of any attempt to come to grips with the implications of climatic changes. The results of such analyses provide the closest thing available to a "ground truth" about the nature of hydroclimatic changes, and against which projections developed using general circulation models (GCMs) of global climate may be assessed. Although data scarcity in the province places limits upon the certitude and completeness of inferences that can be drawn from historical records, considerable prior work on climate trend analysis in BC has successfully outlined the general shifts that have occurred in temperature and precipitation regimes over the decades. Smaller bodies of work have similarly been developed for river flow or other hydrologically salient variables. However, no study to date has systematically assessed long-term trends in the inflows to BC Hydro reservoirs. This is a notable limitation, as various meteorological driving forces combine with local-scale terrestrial hydrologic properties in complex ways to generate net hydroclimatic signals. Consequently, the only way to know with confidence the effects of climate change in a given watershed is to directly assess changes in runoff data for that river.

This important knowledge gap was addressed by performing an in-house study of long-term trends in the inflows to 20 BC Hydro reservoirs, in addition to the combined usable inflow, a very approximate but useful summary measure of system-wide hydroelectric power potentially available for generation. The study consisted of two sets of analysis:

- In the primary analysis, a relatively simple linear technique was combined with direct estimation of signal-to-noise ratios (*S:N*) and applied to monthly and annual mean inflow data over a common period of record for all basins. The result traces out changes in both seasonal runoff patterns and annual net water supply, in a manner conducive to direct comparisons between basins.

- An auxiliary set of analyses was then performed to investigate the sensitivity of these results to data selection, processing, and analysis choices. To manage project scope, we focussed on total annual inflow volumes for eight indicator basins, as well as combined usable inflow. The data were analyzed using different data selection and processing schemes: common period of record across all indicator basins; full record for each basin; residual after Pacific Decadal Oscillation (PDO) signal removal; and data from Water Survey of Canada (WSC) streamgauges on catchments within or near each indicator basin. Several distinct trend analysis techniques were also employed: linear analysis with *S:N* estimation, as in the primary analysis; nonparametric Spearman rank correlation with adjustment for serial dependence; change point detection using Mann-Whitney sliding window analysis; and visual assessment of time series that were low-pass filtered using polynomial fits. While we focussed mainly on the question of long-term changes in average hydroclimatic conditions, the auxiliary analyses also briefly considered the question of changes in year-to-year water supply variability. This was accomplished using a split-sample approach, implemented using a standard F-test; a nonparametric Ansari-Bradley test; and a novel application of information theory involving the use of a Monte Carlo bootstrap method to assess potential changes in Shannon entropy.

Although not every possible analysis technique was used, this study may constitute the most methodologically diverse and complete trend analysis to date of hydroclimatic variables anywhere, and certainly for hydroelectric reservoir inflows in British Columbia.

Key outcomes were as follows:

- There is no significant evidence for a historical decline in annual total water supply in any basin considered, and there is some evidence in some basins for a modest historical increase.
- There is clear evidence for some changes in the seasonality of inflow. In particular, fall-winter inflows have shown an increase at all almost all locations considered. There is additionally some weaker evidence for other seasonal shifts, such as a possible modest decline in late-summer flows for those basins driven primarily by melt of glacial ice and/or seasonal snowpack.
- The strength and robustness of historical inflow changes are variable. Overall, however, most of the historical trends are quite subtle, as quantified for instance using the signal-to-noise ratio. Possible reasons for this result include a low-amplitude climate change signal; comparatively high-amplitude year-to-year fluctuations, due (for instance) to organized modes of climatic variability such as El Niño-Southern Oscillation; and measurement noise in the reservoir inflow estimates.
- Overall conclusions with respect to long-term trends in water supply were found to be largely insensitive to methodological (data selection, processing, and analysis) choices.

- Historical trends identified in the seasonal and annual inflow datasets were found to be readily interpretable in terms of past observed and future projected climate changes. Conversely, however, several hydrologic trends that might be conceptually inferred from these climate changes were not reliably identified in the data. Possible explanations are that some of the practical hydrologic impacts of these climatic changes are too subtle to make themselves readily apparent in net water supply measurements obtained over a relatively short period of record, particularly if these data are subject to significant measurement error; or that the conceptual inferences are incorrect.
- Similarly, the historically identified inflow trends identified are consistent with results to date from GCM-driven hydrologic modelling studies of current and future climate, but certain hydrologic signals suggested by these models were not readily apparent in the historical inflow datasets. Potential reasons might include subtlety of these gradual hydrologic changes over time relative to other dataset components; measurement noise in the inflow data; conceptual or technical problems with some portion of the modelling chain employed in model-driven hydroclimatic studies; and a possible artificial inflation of signal clarity in these modelling results. This result may have implications for the practical meaningfulness of model studies.
- No evidence was found for historical changes in the severity of year-to-year fluctuations in annual reservoir inflow volumes. This result implies that the reliability and predictability of annual water supply have not changed appreciably.

This report provides detailed discussion of technical aspects of the analysis procedures, results of the investigation along with detailed interpretations and assessments, and recommendations for additional work.

Table of Contents

1 Motivation

Two considerations motivate an analysis of historical natural inflows to BC Hydro reservoirs for long-term trends that may be associated with climate change.

The first is that it is a scientifically necessary complement to the large, modelling-based, future-looking simulation studies commissioned by BC Hydro and supervised by Hydrology & Technical Services. Although highly sophisticated, and of course valuable for planning purposes, the projections of future conditions being developed in the Pacific Climate Impacts Consortium (PCIC) and Western Canadian Cryospheric Network (WC^2N) projects are based on a variety of environmental models and greenhouse gas (GHG) emissions scenarios, all of which are in turn based on a broad array of assumptions, and are subject to much uncertainty. Indeed, the major goal of modelling studies such as those could be viewed as an effort to characterize and quantify future uncertainty in reservoir inflows under climate change. As such, empirical, data-driven analysis of observational records from the recent past – that is, taking a close, evidence-based look at what has actually happened under historical climatic changes – can add much value by providing solid grounding and context.

The second is that it is important to take stock, both at the operational level of GRM, and perhaps at what might be called an organizational or corporate level, of the changes that have occurred in reservoir inflows – the ultimate source of revenues for a heavily hydroelectric-based power utility like BC Hydro. For instance, in the British Columbia Utilities Commission (BCUC) Decision on the 2006 Integrated Electricity Plan (IEP), the Commission Panel concluded that (M. Rucker, BC Hydro, pers. comm., 2009):

> "BC Hydro should continue to assess the potential effects of climate change on its hydroelectric resources and that in addition to the activities it is currently involved in, BC Hydro should conduct statistical analyses of snow pack, annual precipitation and stream flows, freshet timing and other relevant variables and survey the relevant literature on an ongoing basis for relevant regional trends, with a view to assessing the impact on stream flows and on its major reservoirs. The Commission Panel directs BC Hydro to file a report with the Commission in its next IEP, identifying significant trends in the literature and summarizing the results of its statistical analyses of historical streamflows."

Although the implications of recent legislative changes to the relationship between BC Hydro, the no-longer-extant British Columbia Transmission Corporation (BCTC), and BCUC are such that the foregoing directive to BC Hydro may no longer be binding in a narrow sense, the underlying need for such a study nevertheless remains. Responsibility for that task has evidently fallen to Hydrology & Technical Services.

Trend analyses have previously been conducted for a variety of hydroclimatic variables over all or certain parts of the province (*e.g.*, Whitfield and Taylor, 1998; Whitfield and Cannon, 2000; Whitfield, 2001; Cunderlik and Burn, 2002; Zhang *et al*., 2000, 2001; Fleming and Clarke, 2003; Stahl and Moore, 2006; Fleming and Moore, 2008; Fleming, 2010). Of particular note for our current purposes is the report prepared by Rodenhuis *et al*. (2007). This study was conducted at PCIC with funding from BC Hydro, and perhaps importantly, its publication post-dates the BCUC commission determination. Rodenhuis *et al*. (2007) reviewed prior work and additionally assessed trends in a wide range of hydroclimatic metrics over British Columbia. These variables included:

- Annual minimum, mean, and maximum temperatures
- Winter, spring, summer, and fall minimum temperatures
- Winter, spring, summer, and fall maximum temperatures
- Total annual precipitation
- Winter, spring, summer, and fall precipitation
- April 1 snow water equivalent (SWE)
- Annual minimum, mean, and maximum streamflows.

As such, the PCIC study provided a good picture (to the extent that one can be developed given limited data resources available in BC) of historical long-term hydroclimatic changes, and clearly satisfies part of the requirement set forth by the BCUC determination.

Nevertheless, existing studies have two limitations. First, examinations of changes to annual stream hydrographs across the province have not been comprehensive. As an example, trend analyses for freshet timing have only been conducted in a subset of studies and were not explicitly included in the PCIC report. Second, no previous study has specifically assessed trends in BC Hydro reservoir inflow volumes. Unlike most meteorological variables, streamflow is not a continuous spatial field. Rather, it is determined by both broad-scale climatological forcing and very local-scale terrestrial factors, such as vegetation type, glaciation, and so forth within the watershed. The practical consequence can be considerable: for instance, profoundly different (indeed, even opposite) long-term water resource responses to uniform changes in climatic forcing have been observationally detected between adjacent basins due to differences in watershed ice cover (Fleming and Clarke, 2003; Stahl and Moore, 2006; Hodgkins, 2009). The implication, then, is that if one needs to know how a given river (or inflow to a given hydroelectric reservoir) has been responding to historical climatic shifts, ideally one should specifically study that basin. It is, in general, imprudent to blindly assume that results from nearby basins will provide an accurate picture.

Given the foregoing considerations, this study focuses on trend analysis of a selected set of hydrologic metrics derived from reservoir inflow data for BC Hydro hydroelectric projects. To the extent practicable within the scope of this project, trend analysis results are interpreted in terms of their spatial patterns, relationships to local basin characteristics and broader climatic changes, and signal strength.

2 Data

The inflow data used as the basis for this study are a more heavily quality-controlled version of those routinely used for operational purposes at BC Hydro. These local reservoir inflows are estimated operationally by a data group within Hydrology & Technical Services using the FLOCAL system. This method involves a water balance based on reservoir storage (estimated from measured reservoir elevation using a storage curve), reservoir releases as turbine flow (estimated, using rating curves, from data on hourly power generation, on unit status, and on unit head as determined from measured forebay and tailwater elevations), non-power releases through gates (as determined from head on each device and recorded gate positions), and for plants downstream of another dam, upstream discharge (L. Brownell and S. Matthews, BC Hydro, pers. comm., 2010). For headwater projects, these data give the inflow volume from the entire upstream basin area. For non-headwater projects, these data give the local inflows between the dam and the next hydroelectric project upstream.

The reservoir inflow data are quality-controlled on a daily basis, with further corrections made on an ongoing basis as the needs arises (L. Brownell, BC Hydro, pers. comm., 2010). The FLOCAL data were re-quality-controlled previously during the course of the UBC Watershed Model recalibration project, which involved (among things) updating the process-oriented watersheds models employed in BC Hydro's operational river forecast system (RFS) (see Fleming *et al*., 2010). It was these data which were employed in the trend analysis. Inflow data were sourced from the station inventory spreadsheet (also developed during the course of the recalibration project) for the auxiliary analyses, and from files generated during the course of the hydroclimate project (see Gobena, 2010; Gobena *et al*., 2010) for the primary analysis (see Section 4 for methodological details).

Inflow data for most of the roughly two dozen forecast points included in the UBC Watershed Model recalibration project, as well as some system-wide total inflows, were analyzed: Alouette, Coquitlam, Stave, Wahleach, Cheakamus, Clowhom, Carpenter (*i.e.*, Bridge wet + dry), La Joie, Ash, Comox, Strathcona, Jordan, Arrow, Duncan, Kootenay Lake, Mica (total), Revelstoke, Sugar, Whatshan, and Williston (total). Combined usable inflow was also considered; this is a rough but useful measure of total inflows across the BC Hydro system potentially available for hydroelectric generation, expressed as equivalent GWh under some simplifying assumptions, and was obtained from the monthly-updated GRM spreadsheet, BCHINFLO_TSM.xls. This selection of measures (see also Table 1, Figure 1) spans most of the BC Hydro hydroelectric system, and virtually all the locations for which quality-controlled reservoir inflow data are available, providing the basis for a good overview of the impacts of possible long-term climate trends upon reservoir inflows.

Table 1 Locations considered in the primary analysis. Upstream areas are GIS-derived (S. Weston, BC Hydro, pers. comm.., 2009). Inflows are 1984-2007 average hydrologic-year (October-to-September) annual volumes. One component of the auxiliary analyses considered some additional WSC hydrometric datasets; these are discussed in Section 3.2.1 and Table 2.

Location	Water Course	Dominant Regime[3]	Inflow (m^3)	Area (km^2)
South Coastal				
Alouette	Alouette River	pluvial-nival	6.7×10^8	202
Coquitlam	Coquitlam River	pluvial-nival	6.5×10^8	188
Stave	Stave River	pluvial-nival	3.5×10^9	956
Wahleach	Wahleach Creek	pluvial-nival	1.9×10^8	88
Cheakamus	Cheakamus River	pluvial-nival-glacial	1.6×10^9	721
Clowhom	Clowhom River	pluvial-nival	1.2×10^9	381
Ash	Ash River	pluvial-nival	6.4×10^8	238
Comox	Puntledge River	pluvial-nival	1.0×10^9	464
Strathcona	Campbell River	pluvial-nival	2.5×10^9	1,193
Jordan	Jordan River	pluvial	3.8×10^8	143
Bridge				
Bridge	Bridge River	nival-glacial	1.6×10^9	2,719
La Joie	Bridge River	nival-glacial	1.4×10^9	988
Columbia				
Arrow	Columbia River	nival-glacial	1.0×10^{10}	9,879
Mica	Columbia River[1]	nival-glacial	2.3×10^{10}	21,134
Revelstoke	Columbia River	nival-glacial	7.0×10^9	5,253
Sugar	Shuswap River	nival	1.2×10^9	1,127
Whatshan	Whatshan River	nival	2.7×10^8	393
Kootenays				
Duncan	Duncan River	nival-glacial	3.2×10^9	2,426
Kootenay Lake	Kootenay River	nival	1.1×10^{10}	20,701
Peace				
Williston	Peace[2]	pluvial-nival	3.6×10^{10}	72,078
System Total				
Combined usable inflow	n/a	n/a	4.3×10^4 GWh	n/a

[1] Mica is the total inflow to Mica dam. Note that this includes both measured flows at the WSC gauge on the Columbia River at Donald and all other contributions upstream of the dam; these components are split within the RFS.

[2] Williston refers to total inflow to Williston Reservoir. Note that this consists of the sum of the Finlay, Parsnip, Nation-Pack, and Williston (Local) sub-basins as defined within the RFS.

[3] Regime descriptions given here are general. In detail, regimes may be more complicated. For instance, the Peace basin is likely more strongly groundwater-influenced than other basins considered here, and also has some locally important glacial melt contributions though these are likely a negligible contribution to Williston total inflows.

For the primary analysis (see Section 3.1 below), we used inflow data consisting of annual time series of hydrologic-year mean flow, and monthly mean flows for each of the 12 months of the year. Each of the resulting 13 time series were independently analyzed. Annual mean flow captures overall changes in hydroelectric generation potential in dammed rivers, particularly larger interior projects with considerable storage. Note that mean flows were calculated over the water year, spanning October of one calendar year through September of the next, and were assigned to the second calendar year. Averages for individual months trace out possible changes in the seasonal distribution and timing of runoff, with implications for reservoir management, aquatic habitat, and flood potential, as well as hydroelectric generation potential for smaller projects with lesser storage. The use of means (monthly or annual) also has an advantage insofar as the additive process whereby an average is calculated from the daily observations affords some noise suppression (provided that the noise is random). Such noise suppression is strongly desirable given the substantial measurement error that may remain in the FLOCAL data even after several QC passes. By the same token, it seems doubtful that the FLOCAL data provide accurate representations of annual high or low extremes in daily flows. Note also that extreme low and high flows were considered in previous studies, including that of Rodenhuis *et al.* (2007), which employed better-quality WSC data. Additionally, some relatively common summary hydrograph metrics, though potentially locally useful, may not be generally meaningful across the BC Hydro network. For example, the centroid of the annual hydrograph, which has been used as a measure of annual freshet timing in snow- or glacier-dominated watersheds (*e.g.*, Zhang *et al.*, 2001; Fleming and Clarke, 2005), is of dubious applicability to the hybrid regimes typical of BC Hydro's coastal projects. Therefore, streamflow metrics were limited to the 13 volume-oriented measures described above. It is believed that these data should be sufficient to characterize substantial changes in flow amount and timing across the year to the extent that they occur.

Several other datasets were additionally employed in the auxiliary analyses described in Section 3.2. FLOCAL-derived annual total reservoir inflow volume, very similar in implications to annual mean flow but perhaps a more intuitive metric of total available water supply, was considered. A number of Water Survey of Canada (WSC) hydrometric records were also employed and were downloaded directly from the WSC website (WSC, 2010a). The Pacific Decadal Oscillation (PDO) index was additionally used for some of the auxiliary analyses. Monthly PDO index values were downloaded from the website of the University of Washington's Joint Institute for the Study of the Atmosphere and Ocean (JISAO, 2010) and were subsequently processed to generate an annual time series of November-December-January (NDJ) averages, assigned to the J year.

3 Method

3.1 Primary Analysis

Linear regression was used to obtain straightforward estimates of average yearly rates of change in mean inflow rates (m^3/s/yr), and associated signal-to-noise ratios were estimated as per Fleming and Moore (2008) and Fleming (2010). Though it is assumed in such an analysis that a linear estimate is serviceable, the lack of formal statistical inference is such that few assumptions are otherwise required (statistical significance tests were not performed in this portion of the study, but see Section 3.2). Additionally, the results are highly intuitive and readily interpreted and presented. Linear regression is almost certainly the most broadly common approach to quantifying long-term trends in hydrology, climate, and other fields. This technique may be viewed as the classic engineer's back-of-the-envelope approach, to which we add an explicit estimate of the signal-to-noise ratio.

This linear assessment technique was applied independently to each of the 13 annual hydrograph metrics (hydrologic-year average and monthly average inflow rates) for each of the 20 locations discussed above plus combined usable inflow, giving a total of 273 analyses. All datasets were truncated to their common period of record (1984-2007), facilitating apples-to-apples comparisons of hydrologic trends across basins and regions. Note also that the basis for generating the inflow data changed at the start of this period (the hourly FLOCAL model was introduced in January 1984; S. Matthews, BC Hydro, pers. comm., 2010). Thus, truncating the inflow datasets (when longer) to that date might conceivably better ensure homogenous data records.

3.2 Auxiliary Analyses

A number of further analyses were performed to assess the possible implications of a number of factors upon the hydrologic trend analyses. These auxiliary analyses are an important project component as they may provide important caveats to (or alternatively, boost confidence in) the results of the full primary analysis.

In the interest of managing project scope, however, these additional analyses were completed using a subset of the data – specifically, basins which are particularly important to BC Hydro's overall generation capacity and/or which serve as (at least approximate) indicators for a given region or hydrologic regime type. These basins were Williston Total, Mica Total, Kootenay Lake, Bridge Total, Cheakamus, Strathcona, Jordan, and Stave; only total annual flow volume was analyzed. Annual combined usable inflow was also studied.

Two general types of issues were considered in the set of auxiliary analyses: data sensitivity, and methodological sensitivity.

3.2.1 Data sensitivity assessment

The primary analysis uses data which were truncated to a common period of record but were otherwise unprocessed (apart from quality control measures). Three alternatives were investigated in the auxiliary analysis.

The first was to employ the full period of re-quality-controlled record available for each basin in the auxiliary analysis. This approach diminishes the inter-basin comparability of results, but facilitates the best possible picture available of climate change trends for a given basin. It may also provide insight into trend robustness, given that the use of different dataset lengths can yield different trend analysis results (see recommendations provided by Khaliq *et al.*, 2009). The resulting record lengths ranged from 35 years (Mica, spanning the water years 1973 through 2007) to 47 years (Cheakamus and Stave, both spanning the 1961 through 2007 water years), with mean and median lengths of 42 and 43 years respectively (see Figures 3 to 11). Limitations on the start of the record either result from lack of carefully quality-controlled data for earlier years at the older hydroelectric plants, or simply correspond to the start of operation at the newer plants.

The second was to linearly remove the Pacific Decadal Oscillation (PDO) signal from the hydrologic time series considered. A number of studies have shown that river flows in British Columbia are influenced by the PDO (*e.g.*, Mantua *et al.*, 1997; Fleming *et al.*, 2007; Gobena, 2010). Although a number of such coherent modes of oceanic and/or atmospheric circulation are known to leave their mark on river flows in BC, the PDO is specifically problematic from a trend analysis perspective due to its long period. In particular, for the relatively short record lengths typically available in BC, in many cases only one clear PDO regime shift may be sampled, and this can in turn lead to ambiguities in the empirical detection of climate variability *vs*. climate change signals (Chen and Grasby, 2009; Whitfield *et al.*, 2010). This problem might be circumvented by regressing the target time series upon the PDO index and then performing trend analysis upon the residual (Whitfield *et al.*, 2010). Gobena (2010) showed that the relationship between the PDO and annual flow volumes for BC Hydro reservoirs was statistically strongest when a winter PDO index was considered; thus, the November-December-January (NDJ) mean PDO index is used here as the predictor in the regression. The full period of record available for each basin was used. Note that such an approach is not without strong caveats: for instance, it is not yet clear that the PDO and regional climate change signals are independent, which is effectively assumed in this procedure (see Whitfield *et al.*, 2010). Nevertheless, performing trend analysis upon series which have had their PDO influences removed, at least to the extent possible using a pragmatic linear approach, may prove an informative exploratory exercise – specifically, with respect to identifying trend robustness (or lack thereof) under presence/absence of decadal-scale oscillations.

Third, for each of the eight basins selected for the auxiliary analyses, the trend assessment was repeated for two WSC hydrometric station records thought likely to adequately represent short- and long-term catchment dynamics for that basin. The stations selected are described in Table 2. To the extent possible, the sites were selected from those in the Reference Hydrometric Basin Network (RHBN; Harvey *et al.*, 1999) as published by WSC (2010b). The RBHN is a list of WSC hydrometric station datasets felt to be appropriate for long-term studies of natural conditions and dynamics, including climate variability and change impacts. As RHBN spatial coverage is incomplete, however, four non-RBHN WSC records were additionally employed. In several cases, an appropriate WSC station was located upstream on the main stem of the river being studied: these WSC stations included the Columbia River at Donald, Cheakamus River above Millar Creek, Kootenay River at Kootenay Crossing, and Bridge River (south branch) below Bridge Glacier. In other cases, the hydrometric station was located on an upstream tributary to one of the eight basins considered: Omineca River above Osilinka River (Peace), Finlay River above Akie River (Peace), Canoe River below Kimmel Creek (Mica), and Duncan River below B.B. Creek (Kootenay). In still other cases, primarily in the smaller basins of southwest British Columbia, stations believed to be representative but lying outside the catchment of interest were used. A unique set of WSC stations was selected for each of the Columbia, Kootenays, and Peace regions. For the hydroclimatically very heterogeneous southwest portion of British Columbia, however, stations were instead selected to explicitly capture the region's hydrologic regime gradient, from pluvial, to pluvial-nival hybrid, nival, and glacial: Koksilah River at Cowichan Station, Sproat River near Alberni, Capilano River above intake, Cheakamus River above Millar Creek, Lillooet River near Pemberton, and Bridge River (south branch) below Bridge Glacier. Most of these southwest British Columbia WSC stations, which span BC Hydro's South Coastal and Bridge regions, were employed as index stations for more than one BC Hydro reservoir. The analyses were performed for the full period of record available for each hydrometric station. In some cases, one or more monthly values were missing; in that event, no annual total volume was calculated. Usable record lengths varied widely, from a minimum of 24 years (Cheakamus River, spanning the water years 1985 through 2008) to a maximum of 89 years (Capilano River, spanning the 1915 through 2008 water years with some short gaps), with mean and median lengths of 49 and 47 years respectively (see also Figures 3 to 11).

Table 2 List of additional WSC stations used in data sensitivity analysis. "BC Hydro basin" denotes the basins (within the subset selected for auxiliary analyses) for which the specified WSC gauge is used as an index station.

Station name	WSC ID	RHBN	BC Hydro basin
Columbia River at Donald	08NB005	Y	Mica
Canoe River below Kimmel Creek	08NC004	Y	Mica
Omineca River above Osilinka River	07EC002	Y	Williston
Finlay River above Akie River	07EA005	N	Williston
Kootenay River at Kootenay Crossing	08NF001	Y	Kootenay Lake
Duncan River below B.B. Creek	08NH119	N	Kootenay Lake
Koksilah River at Cowichan Station	08HA003	Y	Jordan
Sproat River near Alberni	08HB008	Y	Jordan, Stave, Strathcona
Capilano River above intake	08GA010	Y	Stave, Strathcona, Cheakamus
Cheakamus River above Millar Creek	08GA072	N	Cheakamus
Lillooet River near Pemberton	08MG005	Y	Bridge
Bridge River below Bridge Glacier	08ME023	N	Bridge

3.2.2 Technique sensitivity assessment

Alternative analysis techniques were also explored. At first blush, it might seem that assessing a data series for monotonic trend – that is, whether the values are generally increasing, decreasing, or remaining about the same – is a simple matter, and in some contexts, it can be. Unfortunately, the high complexity and low signal strengths typical of geophysical and environmental time series are such that rigorous trend analyses can be challenging. For instance, genuine structural trends can be subtle, yet spurious chance trends can appear powerful; the assumptions underlying many traditional statistical methods are of dubious or inconsistent applicability to geophysical datasets; direct observational records are often short for the purpose of climate impact analysis; and hydroclimatic datasets can be subject to substantial measurement errors, such as those arising from station moves, local land use changes, shifts in rating curves, or issues around the FLOCAL estimation process. Further, there can be pragmatically relevant philosophical ambiguities around trend analysis – for instance, what constitutes a monotonic trend over one timescale of observation can be part of an oscillation over a longer timescale of observation, and indeed, over longer (*e.g.*, geologic) timescales there are few if any truly permanent, monotonic trends in geophysical series; the physical and/or interpretive meaningfulness or usefulness of some common methods in statistical climatology and hydroclimatology, like null-hypothesis significance testing, have been questioned; and long-memory processes like the Hurst effect and fractal dynamics seem to raise questions about the fundamental nature of hydroclimatic trends and the concept of stationarity. In part for the foregoing reasons, there is no single, standard, widely agreed-upon technical method for hydroclimatic trend analysis. It is therefore prudent to apply a number of techniques to the same datasets, and form overall conclusions about trend presence, directionality, strength and origin on the basis of the full pool of ensuing results, while noting limitations and assumptions as they arise. This general approach is consistent with recommendations made in the literature (*e.g.*, Khaliq *et al.*, 2009), although in practise it does not often seem to be applied.

Kundzewicz and Robson (2004) catalogued a dozen methods for change detection in hydrologic records, and even that list is significantly incomplete. Additionally, new methods, refinements of existing methods, critiques of existing methods, and comparative evaluations of methods identifying one technique or another as being "best" under some specific set of evaluation criteria, continue to be presented in the literature every year (*e.g.*, Franzke, 2009; Khaliq *et al.*, 2009). Practical constraints require that we choose only a limited sampling from this smorgasbord of techniques.

In addition to the linear scoping method employed in the primary analysis, the following four techniques were applied to the full period of record available for each member of the subset of inflow records selected for auxiliary analysis. Each of these methods has a reasonably solid track record in the detection of monotonic trend in geophysical or environmental systems, or is directly derived from a standard technique, and represents a different basic way of phrasing the trend analysis problem – and its solution.

3.2.2.1 Formal nonparametric statistical inference

Spearman rank correlation and its associated test for statistical significance was applied to formally assess the datasets for monotonic trend. This technique is robust to nonlinearity and non-Gaussian distributions. It is applied by correlating the target dataset against time in the context of a rank-based, essentially nonparametric paradigm for significance testing. Overall, this technique may be viewed as the classical approach in statistical hydroclimatological research, with the addition of a pragmatic approach to dealing with non-independent samples as discussed below. Note that other commonly used nonparametric null-hypothesis significance testing (NHST) methods, like the Mann-Kendall test, are based on very similar ideas and assumptions, and appear to deliver closely comparable performance (Yue *et al.*, 2002b).

A potential complication that has been identified in the climatic and hydroclimatic literature for all types of NHST-based assessment of monotonic trend, including but not limited to Spearman rank correlation, is the role of red noise. If serial correlation is present in a time series, and it arises from a stochastic memory process rather than a monotonic trend, the actual significance level of the trend test may diverge from its nominal value – in some cases, to the point that the test is rendered meaningless. Unfortunately, various methods that have been suggested as solutions to this potentially common problem may have important drawbacks – in some cases, similarly rendering the test meaningless. In a nutshell, the problem is that on the one hand, red-noise processes by definition tend to produce similar values from one time step to the next, and these may be expressed as clusters of values above or below the mean, or extended runs of increasing or decreasing values, which may appear as trends in a finite-length observational record; but on the other hand, a genuine structural trend gives an elevated value for the observed serial correlation coefficient, resulting in an overcompensation by techniques built to deal with the former issue. Some further insight into these issues is provided in the Appendix. Considered collectively and objectively, the literature on the subject seems to suggest that no clear and well-accepted solution has yet been identified (*c.f.* von Storch, 1995; Fleming and Clarke, 2002; Yue *et al.*, 2002a; Zhang and Zwiers, 2004; Yue and Wang, 2004; Kallache *et al.*, 2005; Bayazit and Önöz, 2007; Khaliq *et al.*, 2009). Thus, the pragmatic rule of thumb suggested by Fleming (2007a) was adopted here. If the observed autocorrelation coefficient of the target time series is 0.20 or less, serial dependence is ignored. If it is greater than 0.20, then the significance test will be performed both with and without the single-stage deserialization method of von Storch (1995), giving bracketing (liberal and conservative) trend assessments.

Tests for field significance (Livezey and Chen, 1983), an additional step sometimes employed to provide a statistically rigorous assessment of regional trends, were not performed here, for three reasons. First, in a hydrologic context, the various methods potentially used for this task appear to remain somewhat experimental (see for example Khaliq *et al.*, 2009). Second, although concerns over climate change impacts clearly motivate the project, our primary concern lies not with the broad fundamental scientific question of global climate change signal detection per se, but instead with assessing the long-term trends at each individual hydroelectric reservoir. Third, and perhaps most importantly, the concept of a field significance test was devised in the meteorological community for application to continuous meteorological fields. In contrast, as noted in Section 1, streamflow datasets from different watersheds are measures of hydroclimatic processes occurring over potentially adjacent but always discrete spatial domains, and substantially different or even opposite hydroclimatic responses may be observed from one basin to the next depending on local terrestrial hydrologic controls (see also, for example, Neal *et al.*, 2002; Chen and Kumar, 2002; Lafrenière and Sharp, 2003; Fleming *et al.*, 2006; Fleming *et al.*, 2007). Thus, it is far from clear how well the fundamental concept of field significance carries over to hydroclimatic trend assessment. This question may be of particular concern in strongly heterogeneous geophysical environments like British Columbia, and in particular, the southwest portion of the province – where glacial, nival, hybrid, and pluvial watersheds may all lie next to each other, with each potentially having different seasonal or annual responses to climatic variability and change.

It should be noted that there are a number of substantial problems or ambiguities associated with classical, frequentist, NHST-based scientific inference. Some general issues were discussed by McCloskey (1995) and Nicholls (2000), and include (for example) the arbitrariness of significance level selection. In addition, the strictly binary decision framework upon which NHST is based can be especially problematic for short and noisy datasets, and the assumptions underlying NHST may raise issues for environmental assessments vis-à-vis the precautionary principle (for some discussions, see respectively Fleming, 2008 and Adelman, 2004). In the context of hydroclimatic trend analysis, the repercussions of these two issues are that: (i) for a dataset of insufficient length or quality to confirm or rule out the possibility of monotonic trend, the NHST result (failure to reject the null hypothesis) is identical to that which would be obtained if a trend is truly absent, so NHST results can be misleading; and (ii) problem definition in NHST normally begins by effectively assuming that the usual null hypothesis of no trend is true and places the burden of proof upon the alternative hypothesis to demonstrate its superiority, meaning that in the absence of strong evidence to the contrary the NHST result will automatically indicate that no environmental change is occurring. Nevertheless, NHST-based trend detection methods are more-or-less the standard in academic statistical hydroclimatology, and they can serve as a sort of quasi-objective "gate-keeper" (*sensu* Adelman, 2004). It can therefore be expected that the exercise will yield some useful insights.

3.2.2.2 Graphical analysis using low-pass filtering (smoothing)

Graphical or visual analysis seems to be enjoying a resurgence in geophysical and environmental trend analysis. It has long been recognized that exploratory data analysis, including but not limited to a thorough visual inspection of the data for potential presence and form of trend, should be the starting point of any trend analysis project (Kundzewicz and Robson, 2004). In the field of hydroclimatic trend analysis, however, such visual inspection has in general been viewed as a strictly preliminary step (if it is performed at all), with the substance of the analysis being subsequently formed by some analytical method, often some NHST-based statistical test. Presumably due either to the myriad questions that have been emerged around the presumed rigour of such tests (some of which are discussed very briefly above in Section 3.2.2.1), or the technically advanced methods for facilitating visual analysis that have more recently been developed or diffused into climate science, graphical methods are increasingly being used as a primary analysis approach for trend analysis. The overall notion is to use a formal filtering technique to remove high-frequency components from the time series, and then apply visual inspection, physical reasoning, and professional judgement to a graphical representation of the resulting low-frequency signal in order to assess presence and importance of long-term trend. The movement toward applying low-pass filtering techniques specifically in a trend analysis context seems to be most widespread within the air quality community, but has included climatic trend analysis as well (*e.g.*, Zurbenko *et al.*, 1996; Rao *et al.*, 1997; Eskridge *et al.*, 1997; Vingarzan and Taylor, 2003; De Jongh *et al.*, 2006).

Part of the appeal of these methods seems to be the mélange of old and new. Such an approach holds the attraction of technical sophistication and rigour on the one hand, yet it also represents a fallback to the purely instinctual yet generally highly effective human visual/perceptual system (*e.g.*, Mahajan, 2010) on the other. Another important advantage is that few or no *a priori* assumptions are required. Disadvantages include a lack of clear objectivity in the trend assessment process, and also a lack of compactness: in general, the overall result for any given time series must be presented graphically rather than via some summary numerical metric. When other methods for trend assessment are considered carefully and in sharp detail, it becomes unclear whether the former issue is truly any worse for visually based approaches than for most others. As for the second drawback, it is trivial for assessments of the handful of time series considered in our auxiliary set of analyses, though it can indeed become burdensome or even completely prohibitive for large-scale analyses where very many datasets are analyzed.

Several low-pass filtering methods are available for the task. By far the simplest of these is the moving average (MA). Though useful for certain applications, the MA smoothing operator is well-known to have undesirable qualities when viewed formally as a filtering device. It also causes data points to be lost from the beginning and/or end (depending on whether the MA operator is centred, forward, or backward) of the time series, which we can ill-afford given the shortness of the available hydrometric records. The Kolmogorov-Zurbenko (KZ) filter involves an iterated MA, and has a number of desirable qualities. Although still straightforward to implement, its filter characteristics are far superior to those of the MA operator, and its cutoff frequency can be easily and explicitly defined by the user. Weighted moving averages, such as the binomial filter, offer some similar advantages. Unfortunately, while superior to the simple MA, both techniques are still based on moving averages, and both suffer the same problem of data loss at the beginning and end of the time series, which may be problematic for the relatively short records available to this study. The adaptive Kolmogorov-Zurbenko (KZA) filter involves a modified version of the KZ technique, with filter characteristics that self-adapt to the local-in-time properties of the time series. While it continues to be essentially a low-pass filter, the KZA method can also explicitly detect and localize changes in long-term state that occur via sudden (and therefore high-frequency) step changes. Again, however, it involves the application of an MA-based operator. Frequency-domain filtering methods, implemented as operations upon the Fourier transform of the time series, are the classical approach to filtering and likely the most powerful. High-pass/cut, low-pass/cut, band-pass/cut, and notch filters are all available, as are others. Unfortunately, Fourier transform-based approaches tend to perform poorly on very short records, and proper frequency-domain filter design is nontrivial. Another relatively common low-pass filtering approach is the LOWESS smoother. This method involves fitting low-order polynomials to local subsets of the time series to ultimately construct a smoothed dataset. As it does not remove the ends of the dataset, it seems likely to be more appropriate to the task at hand. Its main disadvantage is that, while the algorithm for LOWESS smoothing can obviously be precisely defined, any given application of the procedure seems somewhat non-explicit due to the mish-mash of functions fit to different parts of the dataset. The resulting curve, though clearly a smoothed version of the original dataset, thus seems challenging to directly interpret either physically or in terms of the mainstream body of signal processing knowledge.

For the visual analyses performed here, we instead fit a number of polynomials (orders one through five) to the data and display them all simultaneously. This technique yields a mathematically explicit formula for the smoothed curve, does not lead to end-point data losses, and is simple to implement. Illustrating several smoothing models (polynomials of varying order) additionally facilitates some understanding of the sensitivity or robustness of the result to model choice.

3.2.2.3 Change point detection

Typically, empirical assessments of historical records for the signature of long-term change are phrased in terms of presence or absence of monotonic trend – a gradual, progressive, possibly but not necessarily linear, upward or downward trend. It is conceivable, however, that the nonstationarities associated with long-term climatic change mechanisms may be realized instead as abrupt shifts in mean state (*e.g.*, see Mauget (2003) and references cited therein). Methods for assessing gradual monotonic changes may often prove capable of detecting changes in the mean associated with such sudden climatic shifts, but clearly are not ideal for that purpose, being unable (for instance) to explicitly locate the shift in time. Additionally, such change points may be of profound interest in their own right and thus might warrant a technical method specifically built for their detection.

A wide range of change point detection techniques has been developed for a broad variety of applications in the physical, life, and social sciences (for some interesting examples, or intercomparisons under specific evaluation criteria, see Lee and Heghinian, 1977; Buishand, 1984; Easterling and Peterson, 1995; Lanzante, 1996; Ducré-Robitaille *et al*., 2003; Mauget, 2003; Rodionov, 2004; Kundzewicz and Robson, 2004). Two broad types of applications appear in the hydroclimatic literature: tests for the homogeneity of observational records, and assessments of climate variability and change. Although there is much carry-over of methods between the two applications, the performance criteria for each purpose will usually be different, and the suitability of each method will therefore generally be context-dependent. For instance, an ability to detect multiple changes in the mean occurring over short lengths of time is a useful asset for a technique intended to assess a temperature record for possible impacts from meteorological station moves or changes in instrumentation, of which there may be several, and which may not be spaced far apart in time; whereas such an ability to detect shifts in the mean occurring close together will likely be of lesser use to a study of long-term climatic changes.

In our application, we wish to be able to statistically detect an unknown (but likely small) number of changes, occurring at unknown times, in the context of a retrospective analysis of long-term climate variability and change impact assessment. A useful technique we propose here for this goal is Mann-Whitney sliding-window analysis, which we construct as a straightforward mélange of established techniques in change point detection (*c.f.*, running Mann-Whitney *Z* analysis of Mauget, 2003; ST procedure of Ducré-Robitaille *et al.*, 2003; various step change detection algorithms discussed by Kundzewicz and Robson, 2004). In this method, a window is moved through the time series. At each location in the time series, the median of the first half of the data window is compared against the median of the second half of the data window. The comparison is performed using the nonparametric Mann-Whitney test, which makes no distributional assumptions and is relatively robust to outliers. First-order nonstationarities and the approximate time of their occurrence may thus be directly detected.

It has been noted that such approaches require independent samples and thus are not robust to serial correlation, so that a modification to the method may be warranted (Mauget, 2003). The fundamental issues at play here seem to be similar to those discussed above around the question of deserialization for trend detection: a red-noise process tends to generate groups of values above and below the mean, which may appear to be climatic regimes separated by change points, but are not in fact associated with any "real" underlying structural changes. Taking a somewhat broader view, however, there may be problems with this perspective, and therefore with adjusting for such potential effects: (i) the statistically expressed persistence which we choose in such a case to interpret and model as a stochastic red-noise process in fact corresponds to very real deterministic system physics (*e.g.*, Fleming, 2007b), and the associated periods of elevated or depressed time series values are also entirely real and thus might in some sense be legitimately considered genuine regimes separated by genuine change points; (ii) the strictly binary distinction between runs of positive or negative anomalies generated by a red-noise process, *versus* bi-stability in an underlying deterministic system, may lose much of its significance and meaning if one moves to a fully nonlinear interpretive framework (*e.g.*, Fleming, 2009); and (iii) such attempts to "adjust" the statistical detection process amount to attempts at signal attribution, which is a far more ambitious and sophisticated goal than signal detection, and one which an empirical statistical hypothesis testing procedure seems unlikely to answer well, perhaps being better suited to process-based climate models or other investigative techniques. In light of these complications and subtleties, adjustments for serial correlation (to the extent that such serial correlation occurs within the target time series) were not made to the test procedure. It is important to appreciate that this is done with the explicit understanding that the results indicate change points that stand out against a white-noise null hypothesis but which might variously originate from a stochastic process, a deterministic process, or both.

Published results (*e.g.*, Mauget, 2003), and preliminary scoping applications performed here on synthetic time series and the PDO index, suggest that the choice of window length can substantially affect the detailed results, but also that the broader features of a dataset are picked up using any reasonable value. Thus, window length selection appears important but not critically so for our immediate purposes. A window half-length of 10 years was selected: that is, a total window length of 20 years was used, where the median value of the 10 years of data prior to a candidate change point is compared against the median over the following 10 years (with the candidate change point assigned as the first year in the second half-window). A sample size of 10 years has been demonstrated to work well, both in analysis of continental US annual rainfall (see Figure 1 of Mauget, 2003), and in the aforementioned scoping analyses giving (for instance) PDO index change points which coincide well with Pacific basin regime shifts identified in the literature. A 10 year window half-length also represents a reasonable compromise between plausible climate state representation (it seems inappropriate to interpret state variations with timescales less than a decade as long-term climate regimes and changes) and change point detection resolution (for the relatively short record lengths available, a total window length much longer than a couple decades would often permit only a few window positions across the entire observational time series, so that changes points could not be localized in time). Finally, using a window half-length of 10 years amounts to specifically tracing out decade-to-decade variations in hydroclimate, which seems an intuitive and managerially relevant exercise.

3.2.2.4 Split-sample tests for changes in degree of variation

The goal of the work described above is, in essence, to detect and (to some degree) characterize first-order nonstationarity. However, detecting second-order nonstationarity – that is, changes in the variance – can also be important. Specifically, an increase or decline, over the decades, of the magnitude of year-to-year variations in inflow volumes has implications for the reliability and predictability of water resources and hydroelectric generation potential.

As several years of data are required to define the variance of the annual inflow volume, and the available record lengths are in general only a few decades long, methodological options available to us in this case are considerably more limited. To assess long-term changes in the year-to-year variation of annual inflow volumes, split-sample tests were performed: for each of the full-length datasets identified for use in the auxiliary analyses, the record was split in half, and measures of the dispersion in the data distributions for the first and second halves were compared using statistical tests.

Three such tests were employed, to obtain some feeling for the robustness of the findings under different approaches or assumptions. Two of these are standard statistical tests. The F-test on the variances assumes that the two populations being compared are each drawn from Gaussian distributions, which may not be unreasonable given that total annual flow volume forms the basis of the comparison, and this quantity is derived from non-Gaussian daily streamflow data by an additive process, so that the Central Limit Theorem should apply. The Ansari-Bradley test was also employed. This technique is a nonparametric equivalent to the F-test, but considers dispersion generally rather than variance specifically, and makes no distributional assumptions. It does, however, assume that the two populations being compared have the same median, which in principle could be problematic if first-order stationarity of the type assessed in the remainder of the project is indeed present.

The third approach is an (apparently novel) application of information theoretic principles. Shannon entropy is a now-standard measure of the information content in a data stream, developed by Claude Shannon, a mathematician working on fundamental problems of communications theory at Bell Laboratories. The underlying concept is that uncertainty is equivalent to information, or more precisely, that the degree to which uncertainty is removed by monitoring a signal is a measure of the information content of that signal. When base-2 logarithms are used in its computation, the Shannon entropy provides the average minimum number of binary (yes/no) questions that have to be answered to determine in which of N finite states a system currently resides. As such, it may be expressed in binary digits (bits). For a system of given dimension, the need to ask a larger number of questions on average implies greater uncertainty as to what the current state might be, in turn indicating that more information is gained by measuring the signal relative to the case when few or no questions have to be asked. Good background material may be found in the original work of Shannon (1948) and Shannon and Weaver (1949), as well as an accessible introductory text by Pierce (1980). The use of Shannon entropy has extended far beyond its origins in telecommunications, and it has enjoyed a variety of applications to the environmental and geophysical sciences, including watershed model assessment (Amorocho and Espildora, 1972), a common biodiversity index (Dodds, 2002), assessment of the stability of hydrologic regimes under climatic variability (Krasovskaia, 1995), near-real-time identification of marine pollution events (Jeong *et al*., 2006), and characterization of mesoscale air quality patterns (Fleming, 2007c). Although continuous measures of information content have been developed, including a continuous version of Shannon entropy, the discrete version has much to recommend it, including both simplicity of application and a good correspondence to frameworks often employed for environmental state classification and science communication (Fleming, 2007c).

In particular, the familiar climatological notion of a five-state description of conditions – normal, above-normal, below-normal, strongly above-normal, and strongly below-normal – is ideally suited to entropy applications. That is, the information content (usually denoted H, but denoted S here instead to avoid ambiguity with other symbols employed) of a time series of annual inflow volumes, Q, may be defined as:

$$S(Q) = -\sum_{i=1}^{N} P(q_i)\ \log_2 P(q_i) \tag{1}$$

where $P(q_i)$ is the frequency of occurrence of streamflow state i, and there are a total of N = 5 states, corresponding to the five aforementioned climatic divisions. Note that $\lim_{P\to 0+} P \log P = 0$, so that if a given term in the summation of Eq. (1) contains $P(q_i) = 0$, one sets that term to nil. The annual flow volume for each year, t, of the full time series is classified using the following scheme:

Table 3 Annual flow volume classification scheme

State	Definition
strongly below normal	$q(t) < \mu - 2\sigma$
below normal	$\mu - 2\sigma < q(t) < \mu - \sigma$
normal	$\mu - \sigma < q(t) < \mu + \sigma$
above normal	$\mu + \sigma < q(t) < \mu + 2\sigma$
strongly above normal	$q(t) > \mu + 2\sigma$

In Table 3, μ and σ are the mean and standard deviation, respectively, of the entire available annual flow volume record. We then calculate two values of S, over the first and second parts of the split sample. The result indicates whether sufficiently large differences in the variability of the annual flow volumes have occurred such that the frequencies of basic climate states (normal, high, etc.) have changed. As such, it speaks to a more broadly relevant or intuitive notion of climate stability than, say, the variance.

To gain an appreciation for how notable these observed differences in Shannon entropy may or may not be, we performed a null-hypothesis significance test:

$$\begin{aligned} H_0 &: S_{early} = S_{late} \\ H_1 &: S_{early} \neq S_{late} \end{aligned} \tag{2}$$

As to our knowledge no such hypothesis test exists, we constructed one using a nonparametric, computationally brute-force Monte Carlo bootstrap approach. The method proceeds as follows. The original time series is resampled with replacement to create a synthetic time series of the same length as the original dataset and having (approximately) the same probability density function. The entropies of the first and second halves of the synthetic record are then calculated, and the absolute value of their difference is determined. This resampling and entropy change calculation is then repeated many (here, 10^3) times to form a sampling distribution of entropy change values. If the absolute value of the observed change in entropy is greater than the 95th percentile value of the distribution, then H_0 can be rejected in favour of H_1 at $p < 0.05$. Similarly, the percentile to which the observed entropy change would correspond if it was inserted into the sampling distribution permits an estimate of the p-value at which H_0 may be rejected.

3.3 Implementation

A brief implementation summary is provided in the interest of transparency and reproducibility. The primary analysis was straightforwardly implemented in Microsoft® Excel. A Visual Basic for Applications (VBA) macro was written in Excel to perform some post-processing operations. Of the three approaches (classical, spectral, and statistical) to *S:N* estimation described by Fleming (2010), the statistical approach was deemed most convenient for application here. The auxiliary analysis was performed using a MathWorks MATLAB® script written for the purpose. The following MATLAB® functions were employed:

Table 4 Key technical functions used in MATLAB®

Name	Action	Notes
`regress`	Linear regression	Linear trend fitting and PDO signal removal
`corr`	Spearman rank correlation	Set parameter 'type' = 'Spearman'
`ranksum`	Mann-Whitney test	Equivalent to Wilcoxon rank sum test
`vartest2`	2-sample F-test	Parametric test for equal variances
`ansaribradley`	Ansari-Bradley test	Nonparametric test for equal dispersions
`polyfit`	Polynomial curve fit	Set parameter 'n' = 1, 2, 3, 4, 5
`polyval`	Polynomial evaluation	Smooth data using `polyfit` result

4 Results and Discussion

4.1 Primary Analysis

Results from the 273 analyses of monthly and annual mean flows and combined usable inflow are summarized in Figures 1 and 2. Increasing and decreasing water resource trends are denoted by blue and red shading, respectively. Trends in mean flow rates (expressed in $m^3/s/yr$) which are nil within one position after the decimal point are left unshaded. Trends associated with signal-to-noise ratios of 0.1 or greater are outlined, and those with $S{:}N \geq 0.20$ are additionally denoted in bold font. Detailed discussions of signal-to-noise ratios in the context of geophysical and environment signals in general, and application to trend assessment in particular, are given in Fleming (2010). In summary, the signal-to-noise ratio as applied here takes linear trend to be the signal and all other data series components – whether random, deterministic, or both – to be noise. It gives a relatively intuitive and straightforward measure of the strength and clarity of the trend signal. Experience suggests that $S{:}N \sim 0.30$ corresponds to a trend which visually dominates the dataset; $S{:}N \sim 0.20$ corresponds to a visually obvious though perhaps not dominant trend; and $S{:}N \sim 0.10$ corresponds to a trend which is considerably more subtle but still unambiguously present. No explicit attempt is made within the methodology itself to ascertain whether, for instance, the observed trend might be attributable to a chance occurrence in a random series or is instead structural. That is, using *S:N* to determine whether a trend is of sufficient substance to be of interest to us amounts to a test of one might loosely call physical, rather than statistical, significance. Recall, however, that several statistical significance tests were also conducted as part of the auxiliary analysis set.

4.1.1 Summary of past and future climatic changes

To interpret the inflow trend analysis results, a summary of past and projected future climatic trends over British Columbia is useful. There is much ongoing research in these fields, but the currently available body of work is likely sufficient to drawn a reasonably reliable picture of the overall types of climatic changes that have occurred, and will occur, in the province. The following summary of past and future changes in precipitation and temperature is drawn from Rodenhuis *et al*. (2007).

Historically, most measures of air temperature regime have shown increasing trends over the 1900-2004 period over most of the province. Specifically, annual minimum temperature, annual mean temperature, annual maximum temperature, winter minimum temperature, spring minimum temperature, summer minimum temperature, fall minimum temperature, winter maximum temperature, and spring maximum temperature have all shown increasing trends. Only summer maximum and fall maximum temperatures are partial exceptions: the former shows increasing trends in southeastern and northernmost BC, with cooling trends in a band from south coastal BC to the Peace basin, and the latter shows a strong positive trend in the southern third of BC and negative trends to the north. Rates of increase were locally as high as ~3.5°C per century, though averages are considerably lower and, as mentioned, some temperature metrics in some regions showed declining trends. Precipitation trends have shown a similarly coherent picture of positive long-term trends across most of the province. Specifically, annual, winter, spring, summer, and fall precipitation have all experienced increases. Regional and seasonal details vary somewhat, however, and the precipitation trend directions for the southwesternmost corner of British Columbia seem unclear. Locally, precipitation increases experienced by some areas in BC exceeded 50% over the 1900-2004 period, although such relative changes were in most cases far less.

GCM projections of future (ca. 2050) climatic changes for BC are for the most part closely consistent with those historically observed. This convergence may provide on the one hand a measure of confidence in the GCM results, or on the other hand, a measure of confidence in the notion of using historical data to validate GCM results. To the extent that discrepancies occur, they seem to include higher-order effects. For instance, historical observations suggest increased precipitation for almost all of the year, whereas the GCM results tend to call for similar winter and annual precipitation increases but a summer decrease. An interesting result from the GCM studies, although its hydrological implications are not entirely clear *a priori*, lies with the relative size and consistency of temperature *vs.* precipitation changes. Considering the projected changes the mean, the uncertainty in future climate states as captured by the range in projections obtained using 15 different GCMs with two different emissions scenarios, and the variability in historical climate as captured by the range in the historical record, temperature increases in British Columbia may be more noteworthy than those in precipitation. On the one hand, this result might suggest that hydrological changes associated with precipitation changes (*e.g.*, altered total annual runoff) may be secondary to those associated with temperature changes (*e.g.*, shifts in timing of the melt freshet). On the other hand, precipitation is the principal driving force behind runoff, so that even modest or subtle changes in precipitation may carry tremendous weight. In any event, both the historical record and GCM projections reveal a trend toward a considerably warmer, and probably wetter, climate in British Columbia.

4.1.2 Interpretation of results

The results of the primary analysis, when viewed on a regime-by-regime basis, appear for the most part physically interpretable in terms of the observed and modelled climatic changes discussed above in Section 4.1.1. There are, however, several exceptions. We begin with an assessment of overall patterns, and return later to an assessment of signal strength.

Jordan is the sole pluvial regime considered. On the basis of the foregoing temperature and precipitation changes, we might expect precipitation-driven increases in wintertime and annual mean flows, and a decrease in summer flows due to increased evapotranspiration (ET) losses. While Jordan is dominantly a pluvial basin, to the extent that snow does accumulate the higher temperatures are likely to yield decreased snowpack, by increasing the rain-to-snow ratio in this low-elevation coastal basin. This effect might be expressed as a further increase in cold-season flows (as more wintertime precipitation falls as rain to immediately contributed to runoff) and perhaps a slight decrease in spring flows (because of the lower snowpack). These expectations are generally borne out by the results (Figure 1), particularly with respect to winter flow increases. A notable exception is the lack of any hint of a negative late-summer baseflow trend over the period of record. This result may suggest that increased ET losses under climatic drift has not played a significant hydrologic role at the net watershed scale. However, it is to be noted that the relative (*e.g.*, percentage) error of the FLOCAL data has been found to generally increase with decreasing flows, and late summer inflows to Jordan (in particular, but also to other several other projects) are very low. As such, the lack of an apparent trend in late summer baseflows at Jordan might also conceivably reflect a data reliability issue.

Nival rivers are key to BC Hydro's power generation, including the Peace basin and a number of Columbia and Kootenay basins. The aforementioned temperature and precipitation trends could conceivably generate a number of effects, several of which may be in competition with each other. Warmer temperatures would likely shift the snowmelt freshet forward in time, giving increased early spring and decreased early summer flows. Further, the forward-shifted freshet implies a longer summertime dry season and lower late-summer baseflows. Because the Columbia-Kootenay region is snowmelt-dominated whereas summer rainfall may account for roughly half the year's runoff in the Peace (*e.g.*, Cunderlik *et al.,* 2010), this effect may be stronger in the former than in the latter. Both may additionally feel the effects of greater ET losses in late summer due to higher temperatures. Warmer temperatures could also increase the rain-to-snow ratio, at least during the shoulder seasons of late autumn/early winter or late winter/early spring, and/or at low elevations. This effect might give higher flows from late autumn through early spring. On the other hand, higher overall precipitation is expected, which could raise inflow through much of the year. At high elevations, inland locations, and/or northern latitudes, where temperatures under even a warmed future climate will likely remain consistently sub-freezing throughout the winter, higher precipitation may yield a greater winter snowpack and thus a larger freshet. Comparing these hypothesized impacts to the inferred trends for the nival basins, we see that they are generally borne out: as might be expected, annual means show (at least nominal) increases for all nival basins, as do flows across much of the year, except late summer and perhaps a period of time (varying from basin to basin) during spring, depending on the individual basin. These patterns seem clearer for the Columbia and Kootenay basins than for the Peace.

The south coastal region is dominated by hybrid watersheds. These basins contain two annual freshets: one in winter generated by lower-elevation rainfall, and another in spring generated by the melt of snowpack accumulated at higher elevations the previous winter. As such, we would expect the responses of hybrid watersheds to be a melange of those seen for pluvial and nival basins as discussed above. The results are partly consistent with this hypothesis. As would be expected, annual means increase under the heavier precipitation. Early winter values also increase, likely due to both heavier precipitation, and a higher proportion of precipitation falling as rain rather than snow under a warmed climate. Additionally, mid-to-late spring (April or May) values decrease for several hybrid watersheds, presumably reflecting a snowpack which is smaller (because in these coastal catchments, the rain-to-snow ratio will increase across much or all of the basin over much or all of the snow accumulation season); this flow trend may also reflect an earlier snow melt due to warmer spring temperatures, consistent with the observed tendency toward nominal increases in March or April. On the other hand, there is a subtle but very consistent (across basins) decrease in February runoff which seems difficult to explain. It is also unclear why, given the earlier and smaller snowmelt freshet and the presumably larger summertime ET losses, the late summer baseflow period does not show any decreasing trends for any hybrid catchment (except Cheakamus, discussed immediately below).

The direction of glacionival basin responses to historical climatic changes appears to depend on regional context, and in particular, how far along the path of Holocene recession following the last major glaciation the area is positioned (see Fleming and Clarke, 2003; Stahl and Moore, 2006; Moore *et al*., 2009; Casassa *et al*., 2009). Glacial cover within BC Hydro watersheds occurs in the mountains of southern British Columbia, primarily in the Bridge region and a number of Columbia and Kootenay basins (see Table 1). Prior analyses of historical data (see in particular Stahl and Moore, 2006) suggest that in these areas, progressive loss of glacial cover over the decades – reflecting a contemporary warming climate, probably in addition to ongoing dynamical responses to past climate changes – has led to a loss of glacial melt contributions to streamflow. This loss is most acute during the late-summer baseflow period. During this hydrologic season, seasonal snowpack has melted, exposing glacial ice for melt. By the same token, other sources of runoff generation have generally waned by that point in the year, so that glacial melt becomes an important or dominant contributor to streamflow. The concomitant reduction in total water input to the river might also conceivably depress annual mean flows. In all other respects, the response of glacionival basins to climate change would likely be similar to that of nival watersheds. Overall, then, we would expect the glacionival watersheds considered in this study to show a response similar to that of otherwise comparable nival basins, except that downward trends in late-summer baseflow would be exacerbated; and perhaps, annual mean flows might show a lesser positive or potentially even negative trend. The former hypothesis is certainly borne out by the results of Figure 1: those basins having the most substantial glacial cover (Bridge, La Joie, Mica, Arrow, Revelstoke, and Duncan) exhibit very consistent downward trends in August mean flow, whereas nival and hybrid basins are less consistent in this respect. Interestingly, one hybrid basin – Cheakamus – includes significant contributions from not only winter precipitation and spring snowmelt, but also late-summer glacial melt; this basin is also the only hybrid catchment to show late-summer flow declines.

With a few exceptions, then, the overall patterns of nominal increases and decreases in seasonal and annual inflows are what might be expected given the historical changes in climate – but how strong are these signals? The *S:N* results suggest that the trends are for the most part very subtle. Indeed, if we screen the results for signal-to-noise ratios greater than 0.10, only the following three responses can be seen:

- Four of the ten south coastal catchments exhibit trend in annual water supply with signal-to-noise ratios exceeding 0.10. One of these (Comox) shows *S:N* > 0.20. No other regions considered in the analyses contained a basin with an annual mean inflow trend showing a significant signal-to-noise ratio. Conversely, none of the 20 basins considered (plus combined usable inflow) returned nominally negative trends in yearly mean inflow. Overall, it therefore seems reasonable to conclude on the basis of the primary analysis that no decreases, and possibly some modest increases, have occurred in total annual water supply across the BC Hydro system.
- Three basins in the middle portion of the Canadian Columbia-Kootenay region (Sugar, Whatshan, and Duncan) exhibit upward trends in October with *S:N* > 0.10. This result likely reflects heavier October rainfall, presumably arising from increases in the rain-to-snow ratio of precipitation during the fall shoulder season, possibly in addition to heavier total precipitation, under the observed historical climatic trends. Note, however, that most other basins in the study similarly show an upward (but weaker) trend for October. The result might have implications for the severity of fall storm events, and perhaps for salmonid spawning migrations.
- By far the strongest trend signal identified in the primary analysis is an upward trend in December and January inflows. 16 of the 20 basins, plus combined usable inflow, show a positive December or January inflow trend with a signal-to-noise ratio exceeding 0.10. In nine cases, plus combined usable inflow, this trend showed *S:N* > 0.20; and in three cases (Stave, Clowhom, and Mica) the signal-to-noise ratio is in excess of 0.30 (see Table 2), an indication that monotonic trend is the dominant feature of the dataset. This upward trend in December and/or January mean flows likely reflects similar factors to those controlling the upward October trend (see above). Given that November trends are also largely positive, albeit having *S:N* < 0.10 in all basins, it may be reasonable to conclude that flows over autumn through mid-winter are increasing across much of the BC Hydro system.

The results of any binary (yes/no) test for local trend must be interpreted with some caution, of course, irrespective if this "gate keeper" is a threshold magnitude of the signal-to-noise ratio or some other method, such as the outcome of a statistical hypothesis test. Some of the uncertainties are addressed through the auxiliary analyses below, but a point worth making explicitly here regards the disconnect between the significance of local trends *vs*. the significance of patterns. For instance, there are a few other isolated instances of $S:N > 0.10$, but some of these may be spurious trends. Conversely, the consistency of downward late-summer baseflow trends across many nival and all glacionival systems, and downward February mean flows trends across all south coastal systems, seem suggestive of physically real phenomena, although in these cases $S:N < 0.10$ almost without exception (and the effect if present is therefore very subtle over the historical period). Note that this is not quite the same question as the distinction between local and field significance, discussed at length in the climate science literature. Some attempts have been made to devise NHST tests for regional hydroclimatic trend significance explicitly based on pattern consistency (Fleming and Clarke, 2003; Fleming, 2007a), but these remain rudimentary.

4.1.3 Comparison against modelling studies

One of the principal goals of this analysis is to serve as a data-driven baseline or ground truth in assessing GCM-based projections of changes to future reservoir inflows. The notion is that, as anthropogenic climate change is believed to have occurred over much of the 20th and early 21st centuries, we would expect any observational trends in inflows to date to be generally consistent with those predicted by modelling studies of past and future hydroclimate – if those models are reliable. If this consistency is seen, it can be taken as a validation of the modelling process, and the greater the consistency, the stronger the validation. If on the other hand major discrepancies are encountered, we might question the results of the modelling studies. Substantial differences in trend direction would be particularly suggestive of major structural problems with the modelling approach. We have already seen that overall GCM-projected changes in temperature and precipitation are generally consistent with those observed over the historical record in BC (Section 4.1.1), and further, that the historically observed changes in hydroelectric reservoir inflows appear in turn mostly consistent with those broad shifts in climate (Section 4.1.2). But how do the inflow trend analysis results presented here compare to GCM-driven projections of hydrologic changes in BC Hydro basins?

No modelling studies have been completed or are currently planned to assess the impact of future climatic changes upon the Jordan River. However, Whitfield *et al*. (2002) included the Englishman River in their study of Georgia Basin responses to climate change using the CGCM1 climate model with the middle-of-the-road IS92a emissions scenario, in conjunction with statistical downscaling and the semi-distributed UBC Watershed Model. The

Englishman River is located on the southeast coast of Vancouver Island and is dominantly pluvial, albeit with some secondary snowmelt contributions in spring (*e.g.*, Whitfield *et al.*, 2002; Fleming *et al.*, 2007). As such, it should be at least loosely comparable to Jordan. Whitfield *et al.* (2002) found significant decreases in spring and summer discharge and increases in winter flow under future climates; note, however, that the summertime flow decreases were very small in absolute terms. An increase in annual total flow was also reported. Major elements of this projection match the general trend patterns observed for Jordan (Table 1); a difference is lack of a downward trend in summer flow in the historical record, discussed above in Section 4.1.2.

A number of modelling studies have been completed (see review by Rodenhuis *et al.*, 2007) or are currently being performed (Schnorbus *et al.*, 2010) for the climate change responses of various nival basins across the province. Preliminary results are currently available for one major BC Hydro reservoir, the Peace basin. Schnorbus *et al.* (2010) used over a dozen combinations of different GCMs and emissions scenarios (A2, A1B, and B1) to constrain uncertainties associated with future climate projections. The results were downscaled using a statistical method and were then coupled to the Variable Infiltration Capacity (VIC) distributed macroscale hydrologic model. Flow increases were projected for about November or December through June, and decreases over about July through August or September. That is, the summer freshet moves forward in time and increases in peak magnitude, late summer baseflows decline, and winter flows increase. The largest absolute increases were projected for May and June, and the largest decreases were projected for July. A nominal net increase is projected for annual mean discharge, although the quantitative uncertainty estimates are not suggestive of a strong and robust change. Many of these projected features are at least loosely consistent with trends identified in the observational record (Table 1). The largest positive absolute trend rate observed at Williston historically was for June, and the second-largest negative trend rate observed was for July, in general accordance with model projections. Additionally, nominal increases were observed for several of the winter months, and a small positive trend was observed for annual mean flow. However, there are several mismatches. Perhaps the most notable of these is a negative observed trend for May mean discharge, whereas the modelling study predicts the largest increase for that month. Further work would be required to determine the nature of those inconsistencies. On the one hand, there may be problems with the performance of the GCMs over this remote and data-poor northern basin, the statistical downscaling method, or the hydrologic model or its implementation. On the other hand, it is noteworthy that none of the observed trends for the Peace River exhibited $S:N > 0.10$, so that the observational trends against which we are comparing the model projections are very subtle indeed. It is perhaps possible that there is a problem with our empirical baseline, insofar as the historical trend signals might be sufficiently weak that a longer or better-quality observational record is required to robustly determine their direction and magnitude. Whatever the explanation, the mixture of consistencies and discrepancies imply that the

Peace modelling studies are likely on the right track, but also that this basin in particular may bear closer examination in future work.

Several modelling studies have considered hybrid-regime coastal basins. Two of these studies are of particular note here. The first is Whitfield *et al*. (2002), discussed briefly above with respect to Jordan, who also considered headwater locations in two BC Hydro basins (Coquitlam and Cheakamus), although only results for the Coquitlam River were shown. The second is Schnorbus *et al*. (2010), discussed briefly above with respect to the Peace basin, who also considered inflows to BC Hydro's Upper Campbell reservoir (Strathcona). The two sets of modelling results were remarkably similar. Both projected flow declines over May through September, and increases November through April; the springtime snowmelt freshet shifts forward and shrinks, and the winter rain freshet grows longer and larger. Whitfield *et al*. (2002) found that the Coquitlam River's hydrologic regime would progressively shift from hybrid to large pluvial by 2080; an analogous but less drastic shift was found by Schnorbus *et al*. (2010) for Strathcona by 2050. A slight difference between the two studies was that a mean annual flow increase was projected for Coquitlam, whereas the annual mean flow was projected to show essentially no change for Strathcona (although the quantitative prediction bounds were wide). There are both consistencies and discrepancies between these modelling studies and the empirical analysis results. Some decreasing trends were seen in the observational record of springtime flows, but these were subtle and perhaps somewhat inconsistent (Table 1), providing only partial support to the modelling results. Additionally, the observed nominal downward trend in February flows is not predicted in the modelling studies, and the downward trend in late-summer flows predicted in the modelling studies is not apparent in this set of historical trend assessments. In most other respects, the matches were better. The trend analyses (Table 1) showed upward trends in annual mean flow with $S{:}N > 0.10$ for both Coquitlam and Strathcona, generally consistent with the modelling studies. The historical analyses also generally pointed to an increase in wintertime flows (particularly in December and January, when $S{:}N$ exceeded 0.20 for these and several other south coast hybrid basins), again supporting the modelling studies.

Detailed models of climate change impacts upon streamflow in glacierized basins are relatively new, reflecting the greater associated technical demands – the study must additionally include a model of climate change impacts upon glacier area. However, the field is growing quickly, and a number of studies are available for comparison, all of which have considered either inflows to a BC Hydro reservoir or streamflows at an upstream headwater location feeding into a BC Hydro reservoir.

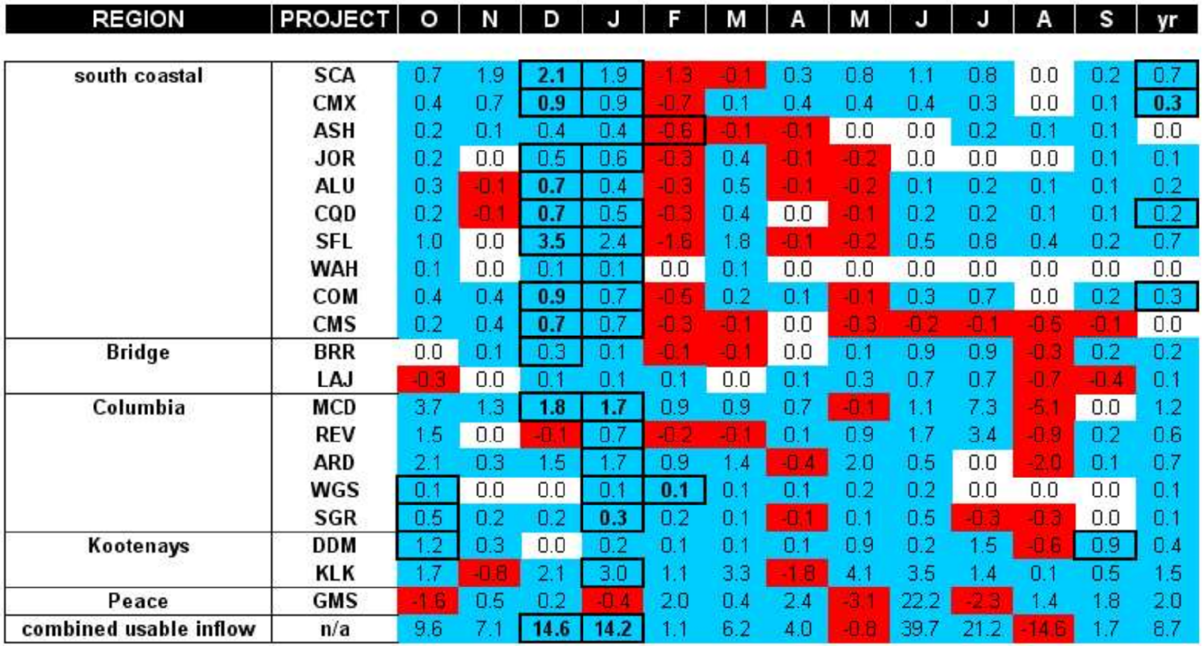
Stahl *et al*. (2008) studied the Bridge system upstream of the BC Hydro reservoir, exploring both a continuation of historical climate (using a stochastic weather generator) and changed climates synthesized by the CGCM3 global climate model under SRES A2 and B1 emissions scenarios (using a statistical downscaling technique). Future glacier coverages were estimated by combining a mass balance model, an empirically calculated but physically based volume-area scaling relationship, and an algorithm for modification of a spatial land use grid. Results were used to drive the semi-distributed HBV-EC watershed model, which has a simple but reasonably effective glacial melt modelling capability. Prediction uncertainty arising from hydrologic model parameter uncertainty was assessed. Overall, it was found that summer streamflow would decrease under all climate scenarios considered, due to continued glacial area loss arising from both future climate changes and ongoing adjustment to past climatic changes. Flow reductions nominally occurred from May forward to October, but were projected to be most acute over July though September. The results imply a concomitant annual mean flow decrease. Spiegelhalter and Stahl (2010) applied the same method to the Canoe River (a tributary to the Columbia River above Mica reservoir) and the Illecillewaet River (a tributary to the Columbia River above Arrow reservoir). An earlier freshet start and lower late summer (July-September) flows were predicted for both. However, Canoe peak freshet flows were projected to shift from July to June and to decrease substantially, whereas those for Illecillewaet did not shift substantially in time and remained unchanged or perhaps increased slightly in magnitude. Annual total runoff was projected to decrease for the Canoe but to remain essentially unchanged for the Illecillewaet. One potential reason for the different predicted responses between these two Columbia headwater basins are local variations in the balance between increased future precipitation *vs*. loss of glacial melt production under a warmed climate (Kerstin Stahl, University of Freiburg, pers. comm., 2010). Schnorbus *et al*. (2010) reported on the Pacific Climate Impacts Consortium's ongoing regional-scale application of the fully distributed VIC hydrologic model across the entire Canadian region of the Columbia and Kootenay basins, using methods discussed above for the Peace and Campbell watersheds. The model will yield flow projections for a multitude of points; preliminary results for Mica were presented by Schnorbus *et al*. (2010). The VIC model does not have an explicit capacity to model glacial melt production, but an approximate technique was devised by Schnorbus *et al*. (2010). The results suggest an earlier and larger melt freshet with flow increases over about March through June, slightly higher winter streamflows, and a significant decrease in late-summer (July and August) flows. An increase in annual mean discharge was projected. Moore *et al*. (2010) and Fleming *et al*. (2010) reported on the most sophisticated work to date on the responses of glacial rivers to climate change in British Columbia. The work, led by R.D. Moore and G.K.C. Clarke of the Western Canadian Cryospheric Network (WC^{2}N), also focuses on inflows to Mica. The study uses both a glacier mass balance model and a fully physically based, fully spatially distributed finite-difference glacier dynamics model. Several combinations of GCMs, emissions scenarios, and runs for a given GCM and emissions scenario (*i.e.*, GCM restarts using perturbed initial conditions) were selected using a number of criteria, ranging from an ability to reproduce weather patterns as determined using Kohonen neural network analysis of daily atmospheric pressure fields, to reproduction of summary climate statistics both globally and regionally. Downscaling was performed separately for the glacier and hydrologic models due to different input data requirements, but both techniques were statistical. The HBV-EC watershed model was employed. Preliminary results are similar to those presented by Schnorbus *et al*. (2010) for Mica in several, though not all, respects. Preliminary results suggest an earlier and larger melt freshet with flow increases over about April through June, slightly higher winter streamflows, and a significant decrease in late-summer (July through September) flows. Annual mean flows were projected to remain essentially unchanged, reflecting a balance between increased precipitation and loss of glacier ice volume potentially available for melt.

How do these modelling results for glaciated catchments compare to the trend analysis results? The matches are variable. The correspondence for the Bridge system is poor (Table 1). The model output is consistent only with the observed August and/or September inflow decreases to Carpenter Lake and La Joie. This discrepancy may arise from the fact that Stahl *et al*. (2008) modelled streamflows at an upstream, alpine part of the Bridge basin, which is far more strongly dominated by glacial runoff – and therefore changes in glacial mass balance and area – than downstream locations near the reservoir, which will experience a comparatively greater impact from such processes as changes in annual precipitation and seasonal snowpack. The general directions of change observed historically for Mica (Table 1) are loosely consistent with those projected by the PCIC and WC^2N studies insofar as both generally suggest increases over winter through early summer and a decrease in late summer, along with no change to a potential small rise in mean annual flow. Mica results (Table 1) are less consistent with model projections for the Canoe River (Spiegelhalter and Stahl, 2010), as the model suggests an annual flow decrease for which there is no historical support, and the model-predicted change in freshet timing is not clearly apparent in the historical analysis either. These discrepancies may reflect the apples-to-oranges nature of the comparison (the Canoe is a headwater tributary to Mica). Note, however, that to the extent there is any empirical evidence for a historical trend in annual total water supply for the Canoe River itself, that trend seems to be an increase (see Section 4.2 and Figure 4), again inconsistent with the Canoe River modelling results. Arrow results (Table 1) appear loosely consistent with the model projections for the Illecillewaet, though this comparison again suffers from the problem of (highly glacierized) headwater *vs*. (relatively less glacierized) downstream locations, as did the comparisons between Bridge headwaters *vs*. Bridge reservoir and Canoe *vs*. Mica. Note that differences in model predictions might also be attributed in part to particular choices of GCM-emissions scenario combinations, although this factor is presumably mitigated by the use of ensembles.

Overall, the findings seem to suggest that modelling studies of glacier-fed rivers speak most clearly to changes in hydroelectric reservoir inflows when they directly consider such inflows, rather than changes in upstream or tributary flows. Given that the statistical results generally tend to find fewer clear changes (particularly if a screening *S:N* level of 0.10 is applied) than what are predicted by the models, the findings may also suggest that modelling studies still underestimate the level of noise in the climate change response (*i.e.*, under-represent total modelling chain uncertainty), at least for Mica. By the same token, however, it seems notable that none of the observed Mica, Arrow, or Bridge trends (except in the December-January period, which seem unlikely to be directly related to glacier change) exhibited *S:N* > 0.10. Thus, as was noted for the Peace, the observational trends against which we are comparing the model projections are again very subtle indeed. It is perhaps then possible that there is a limitation to the use of our empirical baseline, insofar as the historical trend signals might be sufficiently weak that a longer or better-quality observational record would be required to robustly determine their direction and magnitude.

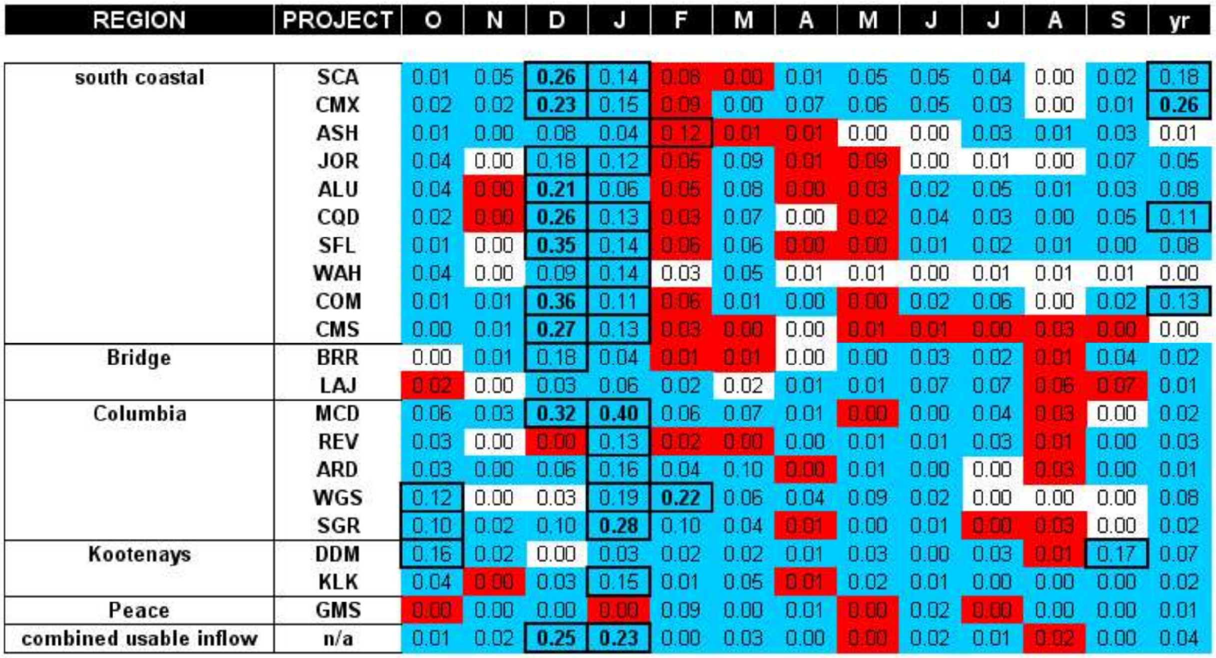
This brings us to three broad conceptual caveats regarding the use of historical data analysis as a method for validating model-generated hydrologic scenarios under future climatic changes. (1) The comparison is in some sense fundamentally apples-to-oranges. The historical record is a single realization of a complex, generally stochastic or chaotic process, of which long-term trend is only one out of many elements. Thus, the signal-to-noise ratio may be quite low (defining as before signal as the long-term trend component, and noise as all other system components or behaviours). In contrast, the signal-to-noise ratio of model output may be inflated, through one or both of two mechanisms. Typically, averages (or some other measure of central tendency) are used for model output visualization and interpretation – for example, mean annual hydrographs in past *vs.* future climates. However, the averaging process increases *S:N* by a factor of $\sqrt{n}$, where *n* is the number of ensemble members (*e.g.*, Telford *et al.*, 1990). Long-term systematic changes may therefore appear in summary model output in an unrealistically clear fashion. The other potential *S:N* inflation mechanism is that a model is by definition a simplification of reality and, therefore, not all variance-producing processes are included. Some potential examples include the complex set of large-scale atmosphere-ocean circulation patterns, which may not be fully reproduced by GCMs; or watershed vegetation changes, such as that associated with the death and regrowth of forests in response to the mountain pine beetle infestation in British Columbia, which are not represented in many or most hydrologic models of climate change impacts. Non-trend variation ("noise") may therefore be underestimated. An important corollary to this first caveat is that while the expectation value (often taken to be the mean) may offer the statistically "best" projection, the single historical realization (and indeed the single future realization that will ultimately be experienced by a watershed) could diverge substantially from the expectation value, without implying that an explanatory model is wrong per se. (2) The second general caveat to our comparisons between historical trend assessments and model-projected changes is that – as discussed in some detail in Section 3.2 – the empirical trend analysis result can depend on a variety of purely methodological factors. Clearly, then, it does not constitute an absolute truth against which model output may be judged. (3) Such comparisons are based on the assumption that future directions and strengths of hydroclimatic change should be broadly reflected in the historical record. This caveat may seem the least limiting of the three: it is well-accepted that GHG-driven global climatic change has been strongly expressed in the late 20th and early 21st century climate against which we are comparing the future projections, and as discussed in Section 4.1.1, the past observed and future projected changes in climate over British Columbia appear to match reasonably well. There could nevertheless be complications, as not all types of environmental change are smooth or even monotonic in character. Tipping-point effects could be one example. Of potential concern in British Columbia is the long-term interplay between climate, glacier, and hydrologic change (*e.g.*, Braun *et al.*, 2000; Moore *et al.*, 2009). At early stages in glacier recession, increased temperatures lead to increased outflows as the melt from large glaciers or icefields is accelerated. Such effects have been observed over the historical record in the large icefields of the St. Elias Mountains in southwest Yukon and northwest British Columbia (Fleming and Clarke, 2003). At later stages in glacier recession, the progressive loss of glacier area and volume available for melting becomes the dominant factor, leading to negative outflow trends. Such effects have been observed over the historical record in southern British Columbia (Stahl and Moore, 2006). In a heavily glacierized watershed, and given sufficient time, it is therefore conceivable that one might observe a trend reversal. This particular issue is unlikely to be of strong concern specifically to BC Hydro reservoir inflow trends, as the significantly glacierized watersheds within BC Hydro's portfolio of hydroelectric resources all lie within southern BC, where glacial streamflows appear to already be in their declining phase. Nevertheless, it is important to explicitly recognize that changes in the direction and rate of hydrologic trend may be possible and, if present, could complicate comparisons between observations of past changes and projections of future changes.

These three important caveats notwithstanding, both scientific rigour and managerial responsibility require that model output be tested before practical decisions are based on it. Data-driven analyses of historical data remain the closest thing available to a ground truth against which to validate model-based projections of the future impacts of climatic change upon reservoir inflows.



REGION	PROJECT	O	N	D	J	F	M	A	M	J	J	A	S	yr
south coastal	**SCA**	0.7	1.9	**2.1**	1.9	-1.3	-0.1	0.3	0.8	1.1	0.8	0.0	0.2	0.7
	CMX	0.4	0.7	**0.9**	0.9	-0.7	0.1	0.4	0.4	0.4	0.3	0.0	0.1	**0.3**
	ASH	0.2	0.1	0.4	0.4	-0.6	-0.1	-0.1	0.0	0.0	0.2	0.1	0.1	0.0
	JOR	0.2	0.0	0.5	0.6	-0.3	0.4	-0.1	-0.2	0.0	0.0	0.0	0.1	0.1
	ALU	0.3	-0.1	**0.7**	0.4	-0.3	0.5	-0.1	-0.2	0.1	0.2	0.1	0.1	0.2
	CQD	0.2	-0.1	**0.7**	0.5	-0.3	0.4	0.0	-0.1	0.2	0.2	0.1	0.1	0.2
	SFL	1.0	0.0	**3.5**	2.4	-1.6	1.8	-0.1	-0.2	0.5	0.8	0.4	0.2	0.7
	WAH	0.1	0.0	0.1	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	COM	0.4	0.4	**0.9**	0.7	-0.5	0.2	0.1	-0.1	0.3	0.7	0.0	0.2	0.3
	CMS	0.2	0.4	**0.7**	0.7	-0.3	-0.1	0.0	-0.3	-0.2	-0.1	-0.5	-0.1	0.0
Bridge	**BRR**	0.0	0.1	0.3	0.1	-0.1	-0.1	0.0	0.1	0.9	0.9	-0.3	0.2	0.2
	LAJ	-0.3	0.0	0.1	0.1	0.1	0.0	0.1	0.3	0.7	0.7	-0.7	-0.4	0.1
Columbia	**MCD**	3.7	1.3	**1.8**	**1.7**	0.9	0.9	0.7	-0.1	1.1	7.3	-5.1	0.0	1.2
	REV	1.5	0.0	-0.1	0.7	-0.2	-0.1	0.1	0.9	1.7	3.4	-0.9	0.2	0.6
	ARD	2.1	0.3	1.5	1.7	0.9	1.4	-0.4	2.0	0.5	0.0	-2.0	0.1	0.7
	WGS	0.1	0.0	0.0	0.1	**0.1**	0.1	0.1	0.2	0.2	0.0	0.0	0.0	0.1
	SGR	0.5	0.2	0.2	**0.3**	0.2	0.1	-0.1	0.1	0.5	-0.3	-0.3	0.0	0.1
Kootenays	**DDM**	1.2	0.3	0.0	0.2	0.1	0.1	0.1	0.9	0.2	1.5	-0.6	0.9	0.4
	KLK	1.7	-0.8	2.1	3.0	1.1	3.3	-1.8	4.1	3.5	1.4	0.1	0.5	1.5
Peace	**GMS**	-1.6	0.5	0.2	-0.4	2.0	0.4	2.4	-3.1	22.2	-2.3	1.4	1.8	2.0
combined usable inflow	**n/a**	9.6	7.1	**14.6**	**14.2**	1.1	6.2	4.0	-0.8	39.7	21.2	-14.6	1.7	8.7

Figure 1 Results from primary trend analysis for all basins and months, as well as annual means and combined usable inflow, presented as linear slopes in m^3/s/yr for mean volumetric flow rates, or GWh/yr for combined usable inflow. Blue and red shading indicate positive and negative water resource trends, respectively; slope values with a magnitude equal to zero within one decimal place are left uncoloured. Trends with signal-to-noise ratios 0.1 or greater are outlined by a box; those with *S:N* ≥ 0.20 are additionally illustrated in bold font. See text for additional details.



REGION	PROJECT	O	N	D	J	F	M	A	M	J	J	A	S	yr
south coastal	**SCA**	0.01	0.05	**0.26**	0.14	0.08	0.00	0.01	0.05	0.05	0.04	0.00	0.02	0.18
	CMX	0.02	0.02	**0.23**	0.15	0.09	0.00	0.07	0.06	0.05	0.03	0.00	0.01	**0.26**
	ASH	0.01	0.00	0.08	0.04	0.12	0.01	0.01	0.00	0.00	0.03	0.01	0.03	0.01
	JOR	0.04	0.00	0.18	0.12	0.05	0.09	0.01	0.09	0.00	0.01	0.00	0.07	0.05
	ALU	0.04	0.00	**0.21**	0.06	0.05	0.08	0.00	0.03	0.02	0.05	0.01	0.03	0.08
	CQD	0.02	0.00	**0.26**	0.13	0.03	0.07	0.00	0.02	0.04	0.03	0.00	0.05	0.11
	SFL	0.01	0.00	**0.35**	0.14	0.06	0.06	0.00	0.00	0.01	0.02	0.01	0.00	0.08
	WAH	0.04	0.00	0.09	0.14	0.03	0.05	0.01	0.01	0.00	0.01	0.01	0.01	0.00
	COM	0.01	0.01	**0.36**	0.11	0.06	0.01	0.00	0.00	0.02	0.06	0.00	0.02	0.13
	CMS	0.00	0.01	**0.27**	0.13	0.03	0.00	0.00	0.01	0.01	0.00	0.03	0.00	0.00
Bridge	**BRR**	0.00	0.01	0.18	0.04	0.01	0.01	0.00	0.00	0.03	0.02	0.01	0.04	0.02
	LAJ	0.02	0.00	0.03	0.06	0.02	0.02	0.01	0.01	0.07	0.07	0.06	0.07	0.01
Columbia	**MCD**	0.06	0.03	**0.32**	**0.40**	0.06	0.07	0.01	0.00	0.00	0.04	0.03	0.00	0.02
	REV	0.03	0.00	0.00	0.13	0.02	0.00	0.00	0.01	0.01	0.03	0.01	0.00	0.03
	ARD	0.03	0.00	0.06	0.16	0.04	0.10	0.00	0.01	0.00	0.00	0.03	0.00	0.01
	WGS	0.12	0.00	0.03	0.19	**0.22**	0.06	0.04	0.09	0.02	0.00	0.00	0.00	0.08
	SGR	0.10	0.02	0.10	**0.28**	0.10	0.04	0.01	0.00	0.01	0.00	0.03	0.00	0.02
Kootenays	**DDM**	0.16	0.02	0.00	0.03	0.02	0.02	0.01	0.03	0.00	0.03	0.01	0.17	0.07
	KLK	0.04	0.00	0.03	0.15	0.01	0.05	0.01	0.02	0.01	0.00	0.00	0.00	0.02
Peace	**GMS**	0.00	0.00	0.00	0.00	0.09	0.00	0.01	0.00	0.02	0.00	0.00	0.00	0.01
combined usable inflow	**n/a**	0.01	0.02	**0.25**	**0.23**	0.00	0.03	0.00	0.00	0.02	0.01	0.02	0.00	0.04

Figure 2 Signal-to-noise ratios from primary trend analyses for all basins and months, as well as annual means and combined usable inflow. Formatting is as described in the caption to Figure 1.

4.2 Auxiliary Analyses

Results from the auxiliary analyses are illustrated in Figures 3 to 11. The results are organized on a basin-by-basin basis. To recap, the auxiliary analyses involved more in-depth studies of total annual flow volume for a number of indicator basins, as well as the annual total combined usable inflow. The analysis steps included:

- Linear trend analysis with *S:N* estimation for the period of data common to all datasets included in the primary analysis (denoted "linear analysis over common period" in the figures), which is essentially an identical analysis to that presented in Section 4.1
- Linear trend analysis with *S:N* estimation for the full period of quality-controlled record available for each indicator basin (denoted "linear analysis" in the figures)
- Linear trend analysis with *S:N* estimation for the time series of residuals obtained after linearly removing the PDO signal from the full period of quality-controlled record available for each indicator basin (denoted "linear analysis on PDO residuals" in the figures)
- Linear trend analysis with *S:N* estimation for the full period of record available for each of two WSC streamgauges thought to be relevant to each BC Hydro indicator basin (denoted "linear analysis: *watershed name*" in the figures) (obviously not included for combined usable inflow)
- Spearman rank correlation, preceded by prewhitening as appropriate, for the full period of record available for each indicator basin (denoted "nonparametric test for monotonic trend" in the figures) (note that in no case considered here was the observed lag-1 autocorrelation coefficient ultimately found to be sufficient to warrant deserialization)
- Change point detection using the full period of record available for each indicator basin (denoted "windowed Mann-Whitney analysis" in the figures)
- Visual trend assessment using polynomial fits, of order 1 through 5, to the full period of record available for each indicator basin (denoted "low-pass polynomial filtering" in the figures)
- Three different analyses of changes in variability, performed by splitting the full period of record available for each indicator basin into early and late halves (bottom panel of each figure).

The goal of the auxiliary analyses was to assess, with a focus on annual water supply, the sensitivity of inferred trend direction and strength to different choices of (i) dataset selection and processing, and (ii) trend analysis technique. Additionally, we used this opportunity to consider changes in the year-to-year variability of total water supply. Results were as follows.

While data truncation and processing choices clearly affect the detailed outcomes, the broad result for any given basin seems largely insensitive. Consider the Peace River, for instance (Figure 3). The short record, the full record, and the full record following PDO signal removal all yield nominally positive long-term trends, but all with *S:N* < 0.10. As another example, consider Bridge (Figure 6). Over the truncated common period of record, the trend is nominally positive but with *S:N* << 0.10. Over the full period of record available, the trend becomes nominally negative, but still with *S:N* << 0.10. And over the full period of record available but following PDO signal removal, the trend becomes nominally positive again, but still with *S:N* << 0.10. In other words, irrespective of such data processing choices, no observational trend can be reliably identified for Bridge annual water supply. This lack of sensitivity to dataset length and to PDO signal removal in terms of broad, overall result was observed for all the indicator basins considered in the auxiliary analyses, with one exception. Strathcona exhibited a substantial (*S:N* = 0.16) upward trend over the common period of record (1984-2007), but the signal-to-noise ratio dropped to insignificant levels when the full period of record available for Strathcona was considered (*S:N* = 0.01) and when the full period of record was linearly adjusted for PDO effects (*S:N* = 0.03).

The results obtained using these various derivatives of the available reservoir inflow data also appear generally consistent with those obtained using WSC streamgauge data, although the latter do seem to help elucidate the potential sources of the observed net reservoir inflow behaviours over time. Again considering the Peace River (Figure 3), one of its western tributaries (the Omineca) shows an upward long-term trend with a substantial signal-to-noise ratio of 0.18, whereas a northern tributary (Finlay, making a contribution to total Williston inflows about three times larger than that from the Omineca) also shows a nominal upward long-term trend but with *S:N* ~ 0. In combination, the trends from these tributaries seem fully consistent with a low-amplitude, ambiguous, nominally positive trend in total annual Williston inflow (Figure 3), but suggest that different parts of the basin may be responding differently to climatic drift. Consider another important example: Mica (Figure 4). The Columbia River upstream at Donald and the Canoe River each show possible subtle trends (*S:N* ~ 0.06 to 0.07), but which are opposite in direction. For the sake of argument, let us accept the notion that the streamflow trends at these two upstream locations, though very modest, are genuine. Then the nominally positive but very small (*S:N* ~ 0) net trend observed for Mica inflows, which are seen irrespective of reservoir inflow dataset length or processing choice, may result not from a zero hydroclimatic trend in the basin, but instead from mutual cancellation of different hydroclimatic signals from different parts (Canoe: north, Columbia at Donald: south) of the basin.

Similarly, while the particular method used for trend analysis influences the detailed outcomes, the broad result for any given basin seems largely insensitive here. Consider Jordan, for instance (Figure 10). Linear analysis of the full period of record indicates a nominally positive ($\beta > 0$) but insignificant (*S:N* < 0.10) trend in annual inflow volume.

Analysis of the same dataset using nonparametric statistical hypothesis testing similarly suggests a nominally increasing ($R_S > 0$) but insignificant ($p > 0.05$) trend. By the same token, a change detection algorithm comparing rolling 10-year medians using a rank-based test exhibited mainly positive consecutive differences (med x_{t+} – med $x_{t-} > 0$) consistent with a positive trend, but these were in all cases statistically significant ($p > 0.05$). Polynomial fits of various orders are again visually suggestive of only a mild upward trend, perhaps superimposed upon fluctuations at roughly decadal timescales. Generally comparable results were found for the other indicator basins considered.

An interesting feature of the results, however, is that for some indicator basins (Cheakamus, Strathcona, and Stave), as well as combined usable inflow, the sliding window analyses indicated a statistically significant change, in contrast to the lack of significance implied by the other analytical techniques when applied to the same record. For these three rivers, the significant change point occurred in the early to mid 1990s, and involved a transition to wetter conditions. For combined usable inflow, the change consisted of a transition to wetter conditions in the 1950s. More analysis would be required to rigorously pin down the nature of the changes detected by this method for these quantities, and why statistically significant outcomes were obtained for some indicator basins and not others. At least in loose interpretive terms, however, the partial inconsistency between the methods appears readily resolved: the other methods (and in particular the linear-*S:N* and Spearman rank correlation techniques) search for monotonic trend, whereas the sliding window analysis searches for discrete change points which may or may not be associated with unidirectional change. That is, they aim to detect potentially related but different phenomena, and it thus seems unsurprising that some discrepancies may occur. Note also that all the basins considered in the auxiliary analyses similarly showed a nominal positive step change in the 1990s.

The results of the auxiliary analyses as a whole seem remarkable precisely for their lack of remarkableness. Given the extensive literature devoted to deep questions of trend analysis methodology, it may be unexpected that, overall, the conclusions drawn with respect to annual water supply for a given basin (and for combined usable inflow) are largely to entirely insensitive to substantial modifications of the dataset assessed and the methods employed. This outcome suggests that such technical details may often be a second-order consideration. Perhaps changes in method or data selection and processing might give such modest changes in the analytical result because they materially affect the overall outcome of a trend detection exercise only when the result is marginal, *i.e.*, on the cusp of "significance," however that is defined for a particular study goal. Thus, in cases of strong trend or (as would appear to be generally true for total annual inflow volumes here) little or no trend, minor differences in analytical result between data processing and analysis methodologies may not affect the broad conclusions drawn regarding managerially relevant long-term trends in some quantity of interest. Alternatively, the practical importance of such technical details may in general simply be considerably less than one might assume given the amount of attention devoted to them by the research community. In either case, this result also suggests that straightforward techniques which can be applied quickly, yield compact answers, and do not require extensive case-specific interpretation – such as simple linear regression with signal-to-noise ratio estimation – can be very effectively used (at a minimum) as a screening or reconnaissance method. Perhaps detailed studies might then be reserved for those cases where the result suggests a borderline *S:N* value of perhaps 0.10 or so.

In no case was a statistically significant change in variability detected using any of the three analysis methods. For the two basins (Bridge and Cheakamus) lying in or near the Sea-to-Sky corridor linking interior and coastal parts of the province, the *p*-value for the Ansari-Bradley test fell just outside the 0.05 level usually taken as the cutoff for statistical significance. Even for these two comparatively "promising" watersheds, however, the results from the other two methods (two-sample F-test on the variance, and Monte Carlo bootstrap test on the Shannon entropy) provided no evidence to support rejection of the null hypothesis of no change. A variety of considerations can come into play in determining the inherent predictability of a data series, a subject of substantial interest in the nonlinear dynamics literature, for example. Nevertheless, the outcome from our analysis does imply that annual total water supplies in BC Hydro's operating areas have not systematically grown more volatile or unpredictable over time.

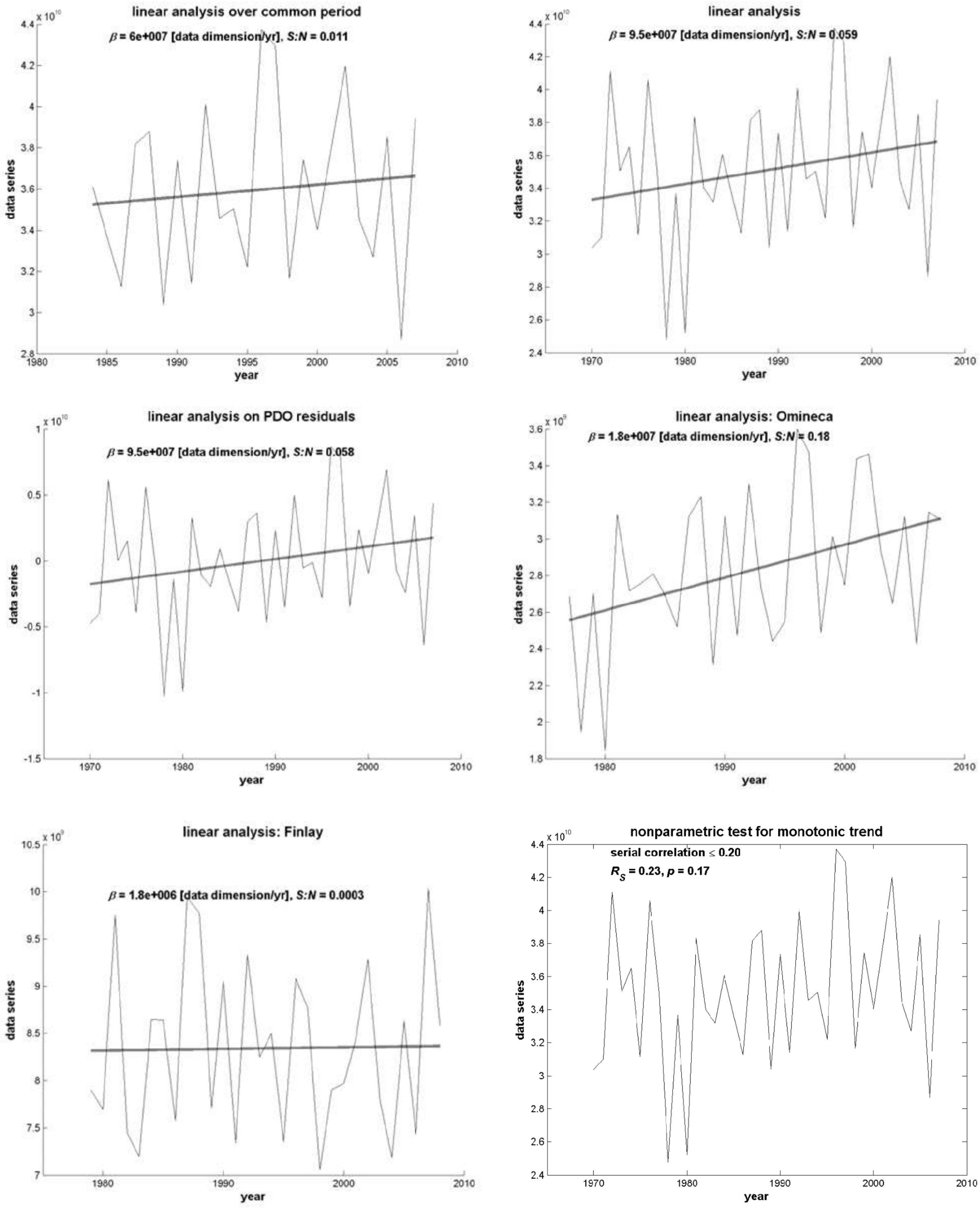


Figure 3 Results of primary and auxiliary analyses for Williston annual inflow volumes in m^3 (continued on next page).

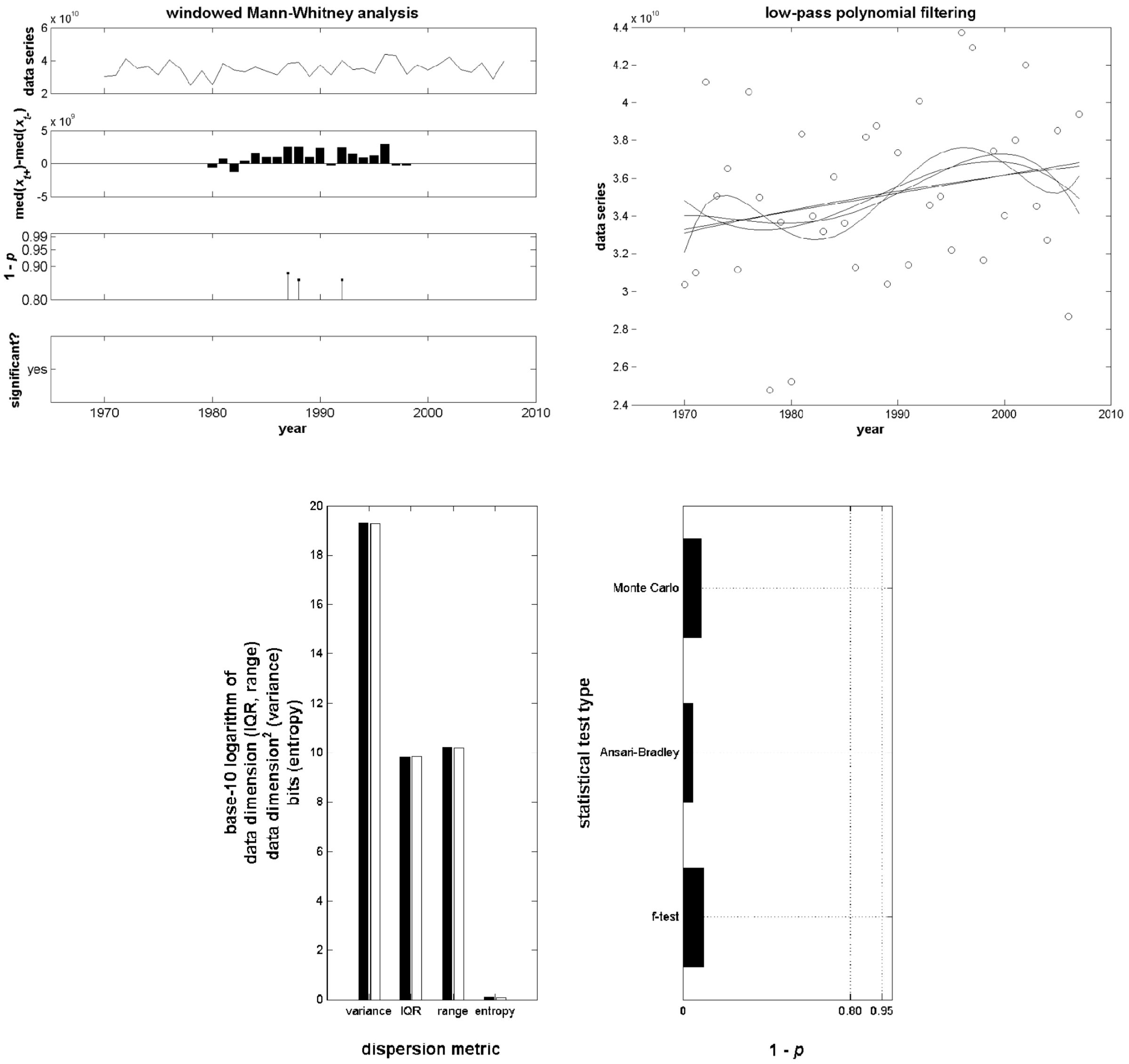
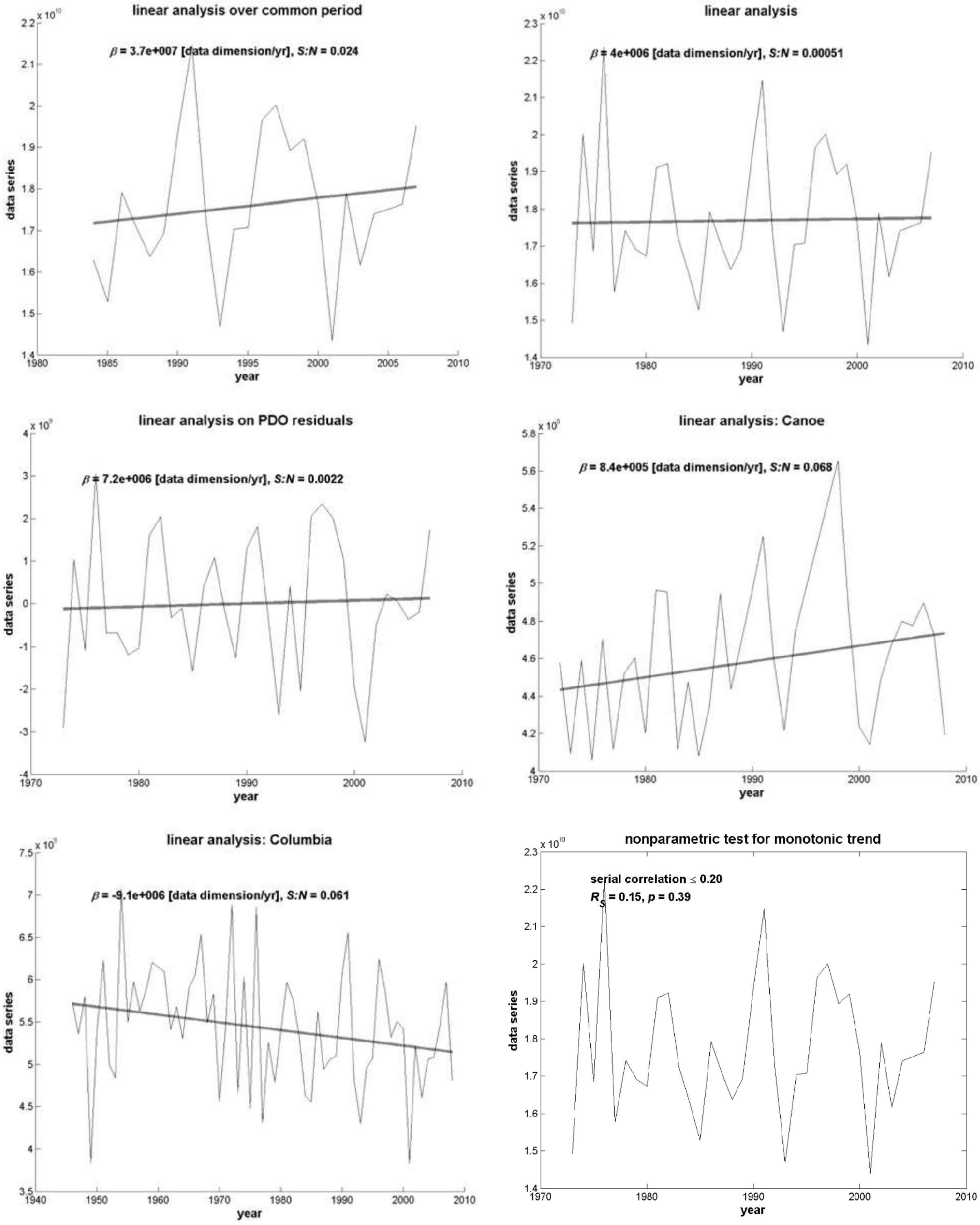


Figure 3 (continued) Results of primary and auxiliary analyses for Williston annual inflow volumes in m^3 (continued from previous page).



linear analysis over common period

linear analysis

linear analysis on PDO residuals

linear analysis: Canoe

linear analysis: Columbia

nonparametric test for monotonic trend

Figure 4 Results of primary and auxiliary analyses for Mica annual inflow volumes in m^3 (continued on next page).

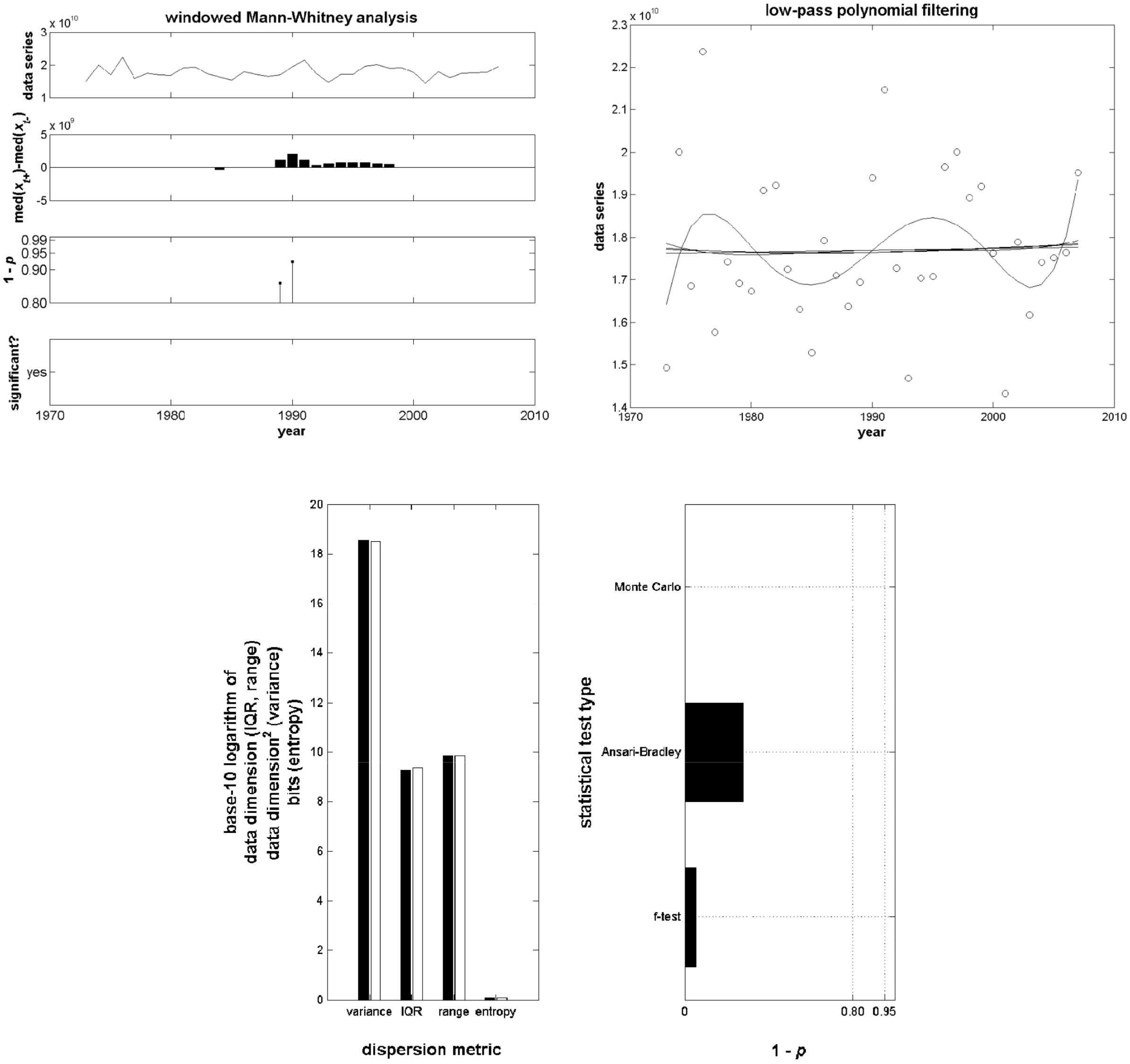


Figure 4 (continued) Results of primary and auxiliary analyses for Mica annual inflow volumes in m^3 (continued from previous page).

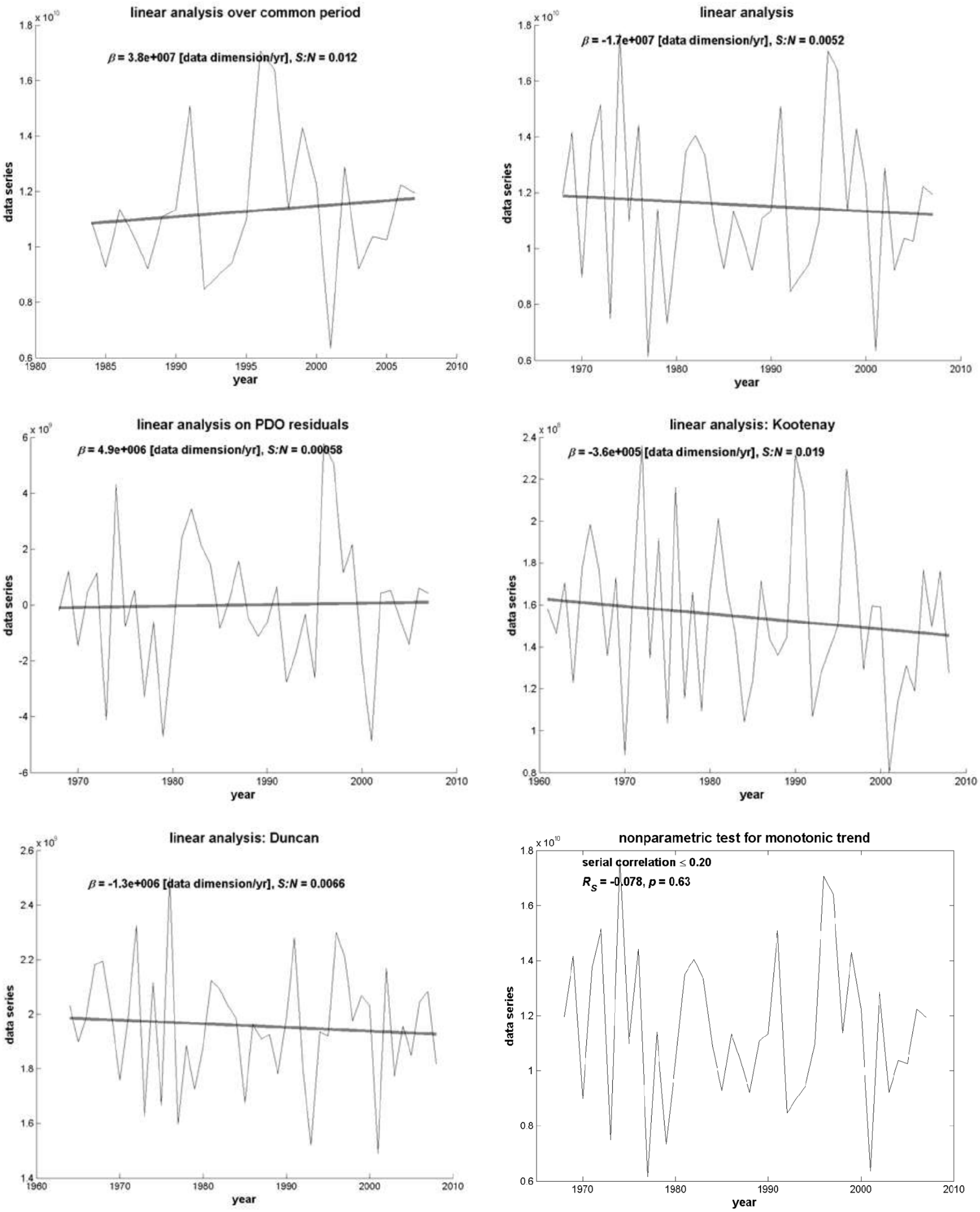


Figure 5 Results of primary and auxiliary analyses for Kootenay Lake annual inflow volumes in m^3 (continued on next page).

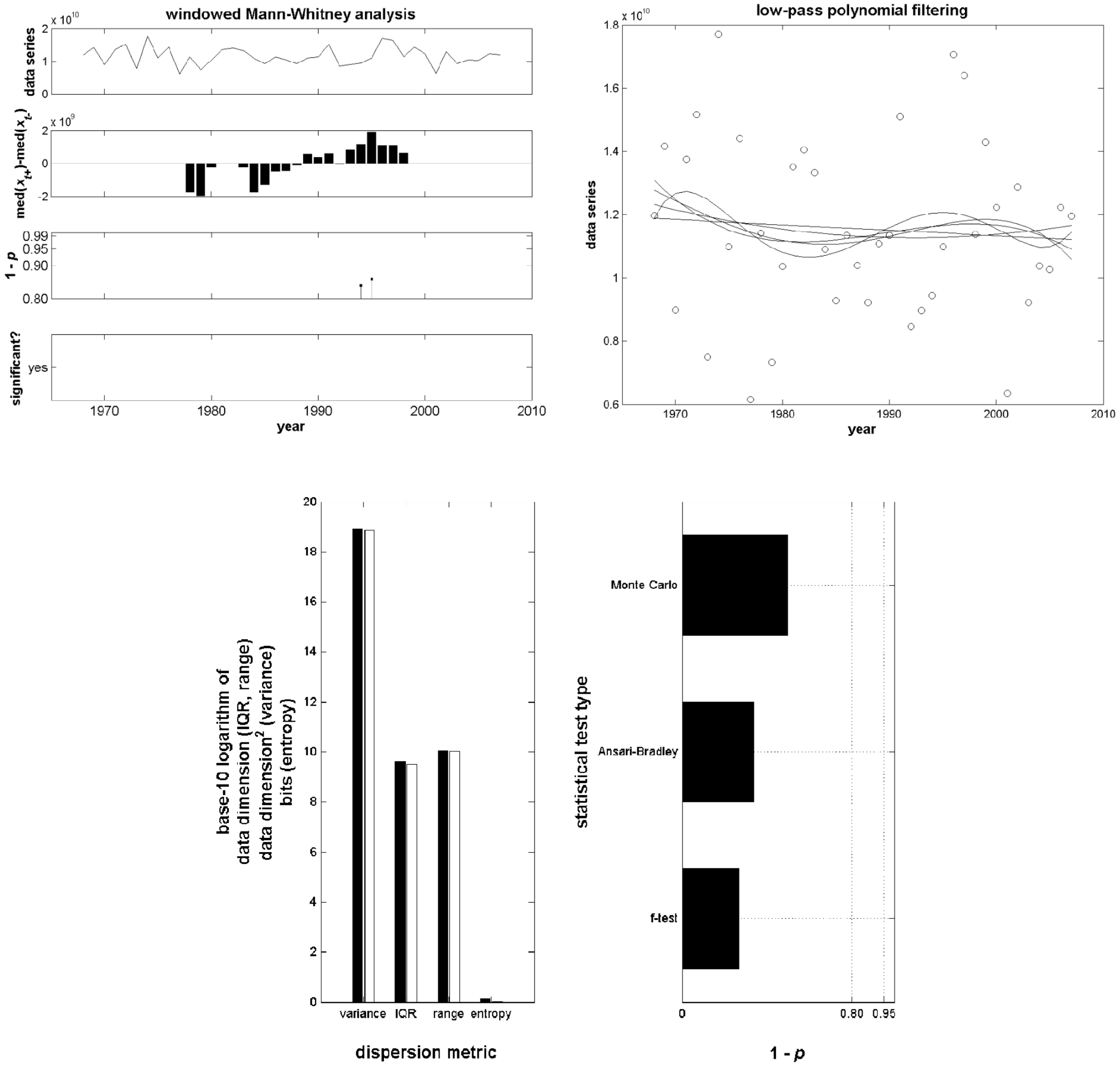


Figure 5 (continued) Results of primary and auxiliary analyses for Kootenay Lake annual inflow volumes in m^3 (continued from previous page).

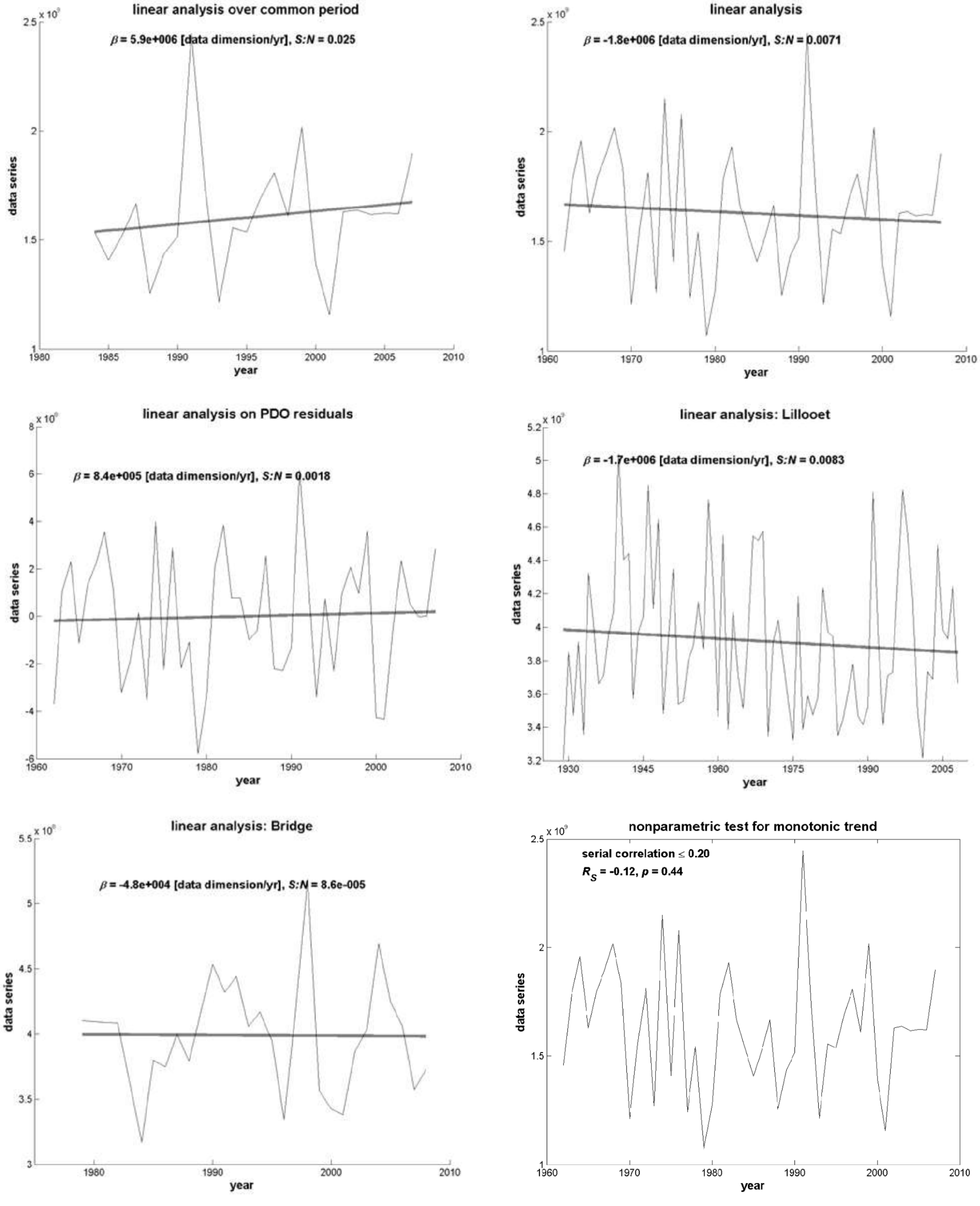


Figure 6 Results of primary and auxiliary analyses for Bridge annual inflow volumes in m^3 (continued on next page).

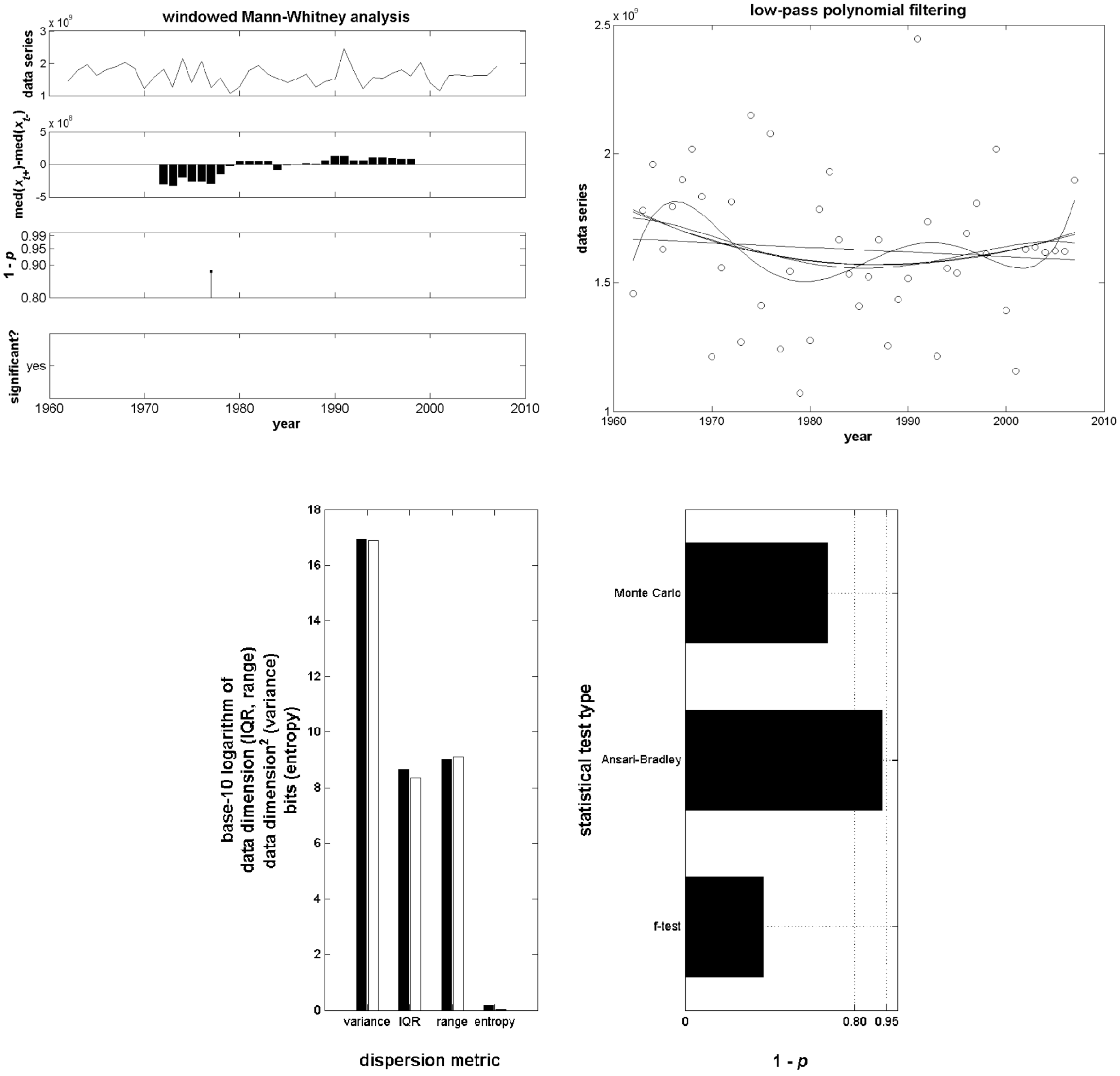


Figure 6 (continued) Results of primary and auxiliary analyses for Bridge annual inflow volumes in m^3 (continued from previous page).

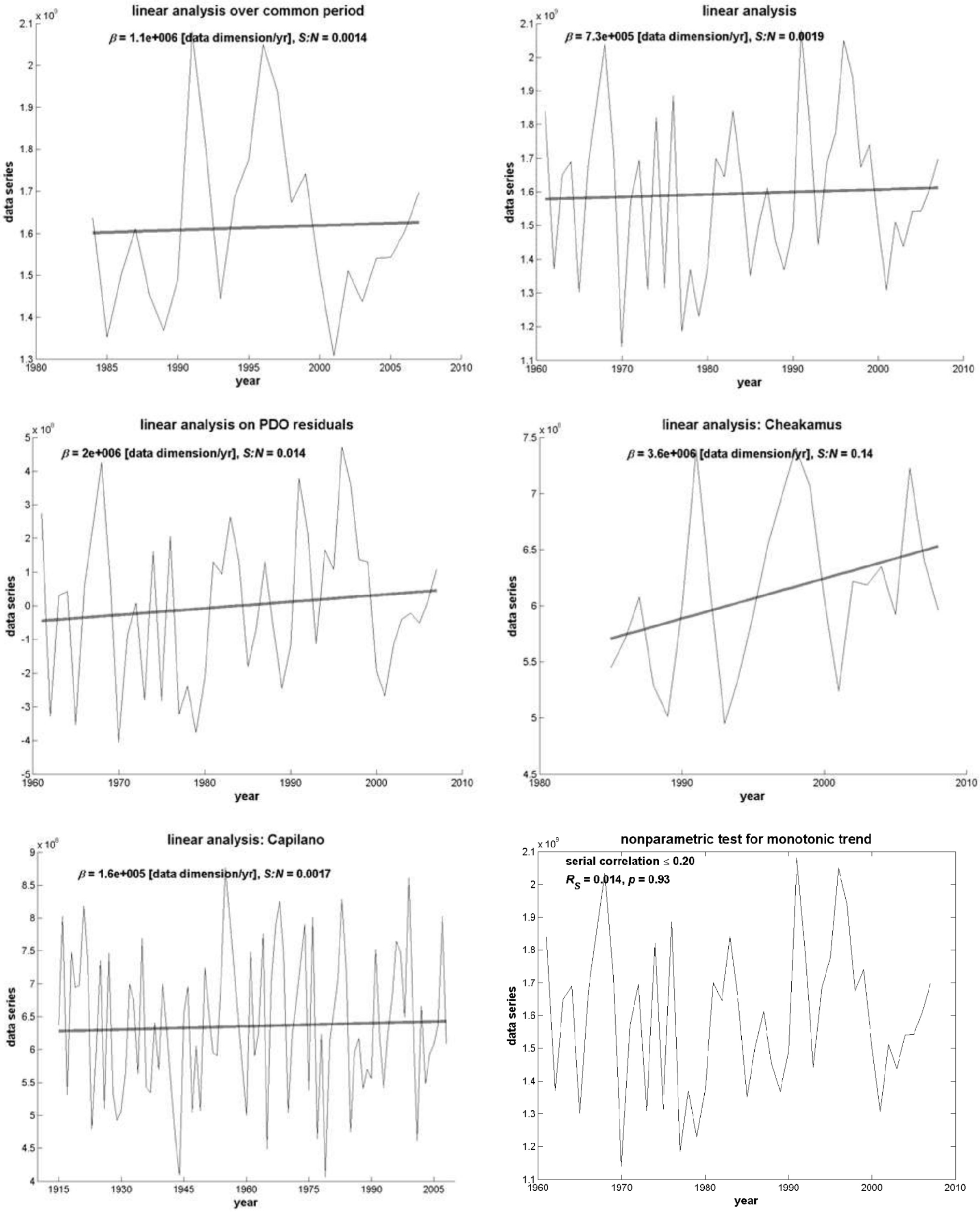


Figure 7 Results of primary and auxiliary analyses for Cheakamus annual inflow volumes in m^3 (continued on next page).

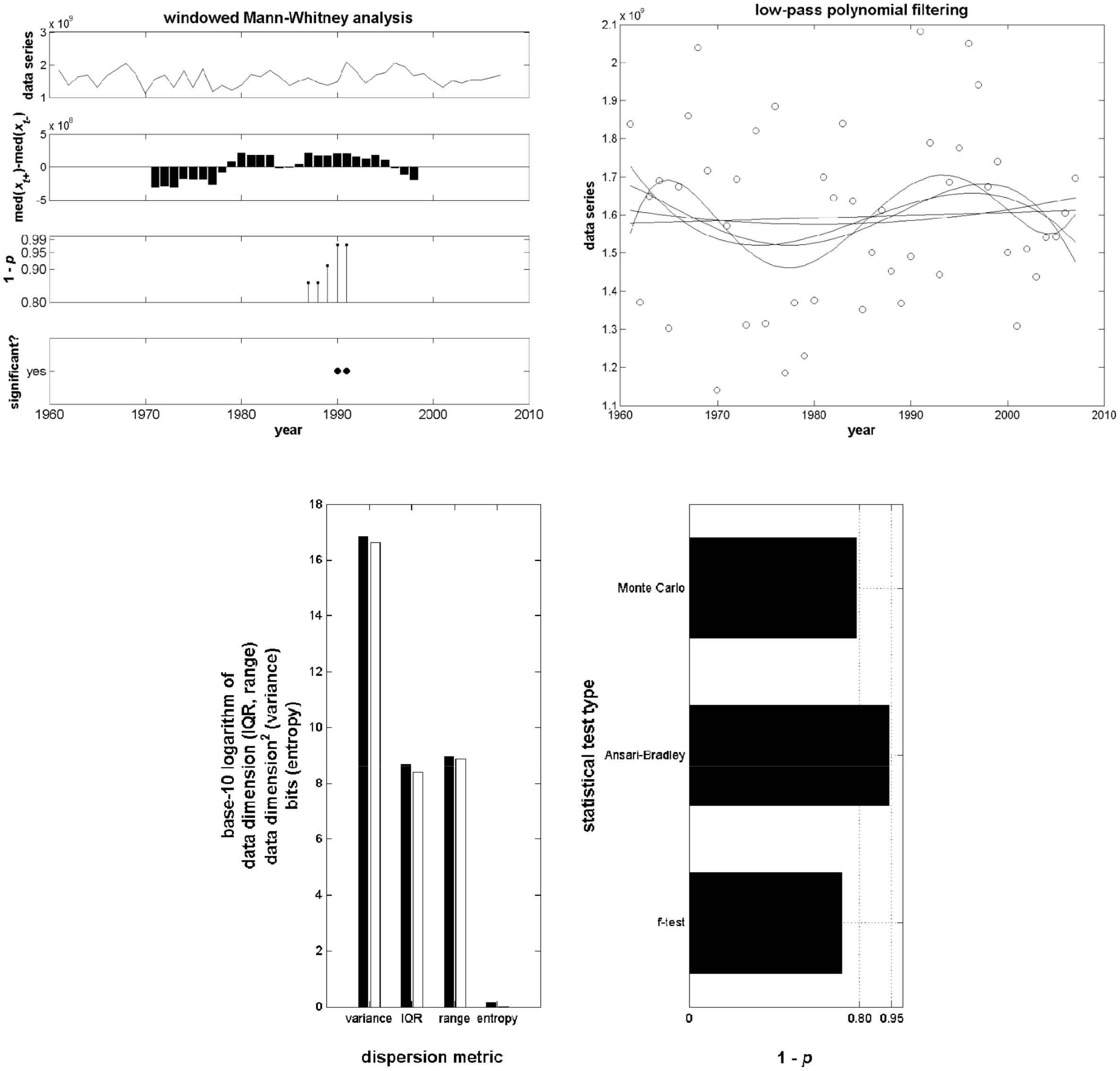


Figure 7 (continued) Results of primary and auxiliary analyses for Cheakamus annual inflow volumes in m^3 (continued from previous page).

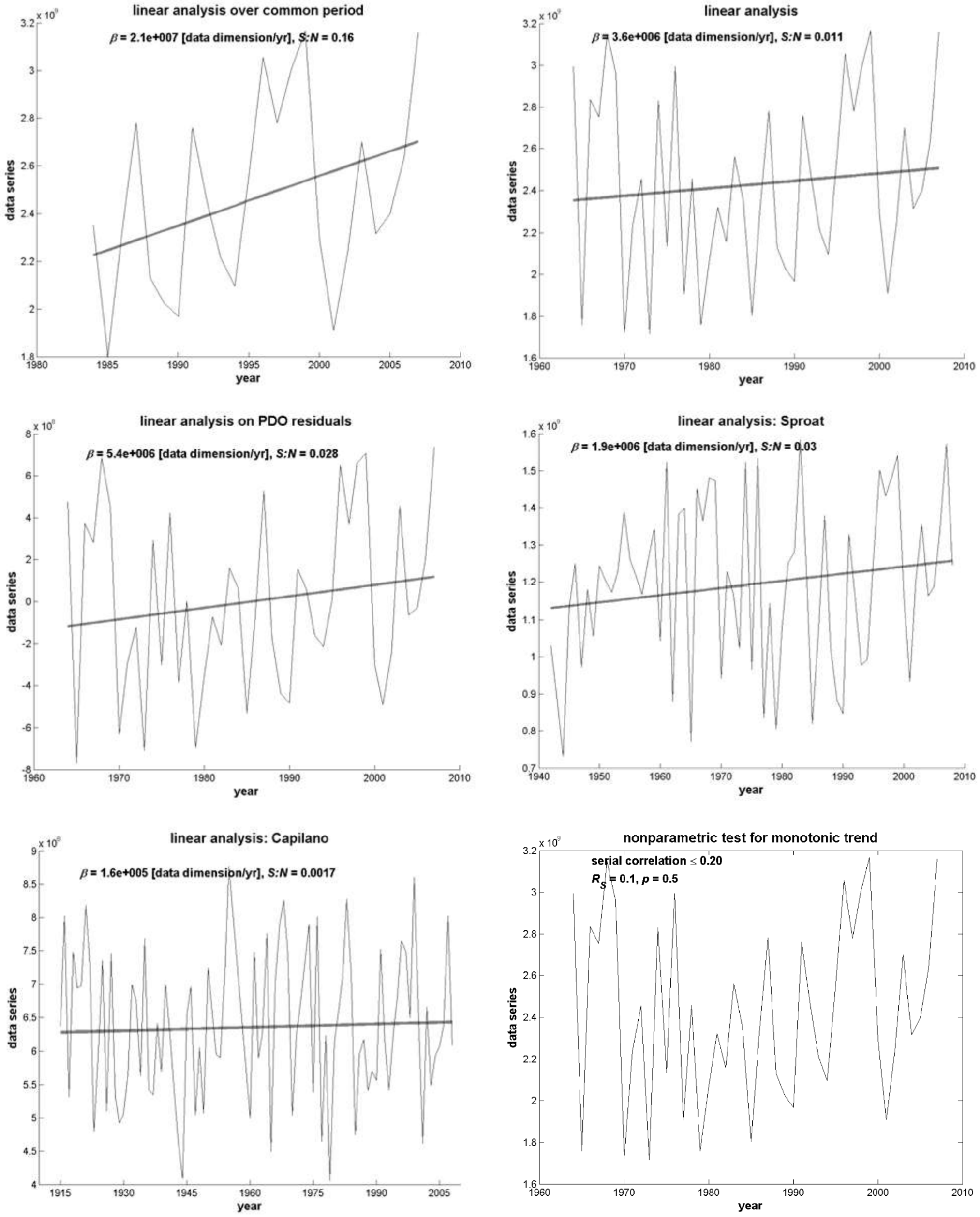


Figure 8 Results of primary and auxiliary analyses for Strathcona annual inflow volumes in m^3 (continued on next page).

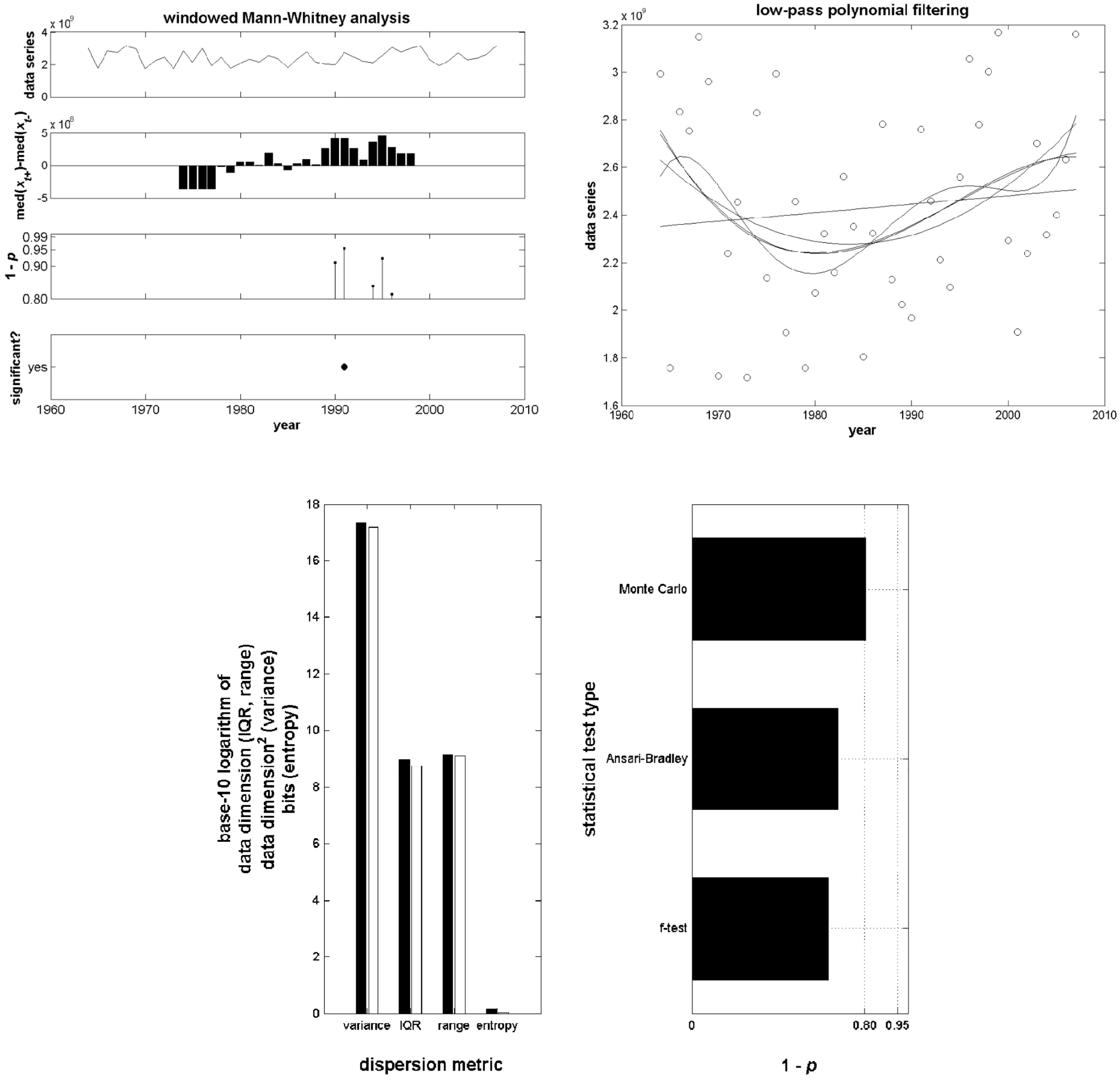


Figure 8 (continued) Results of primary and auxiliary analyses Strathcona annual inflow volumes in m^3 (continued from previous page).

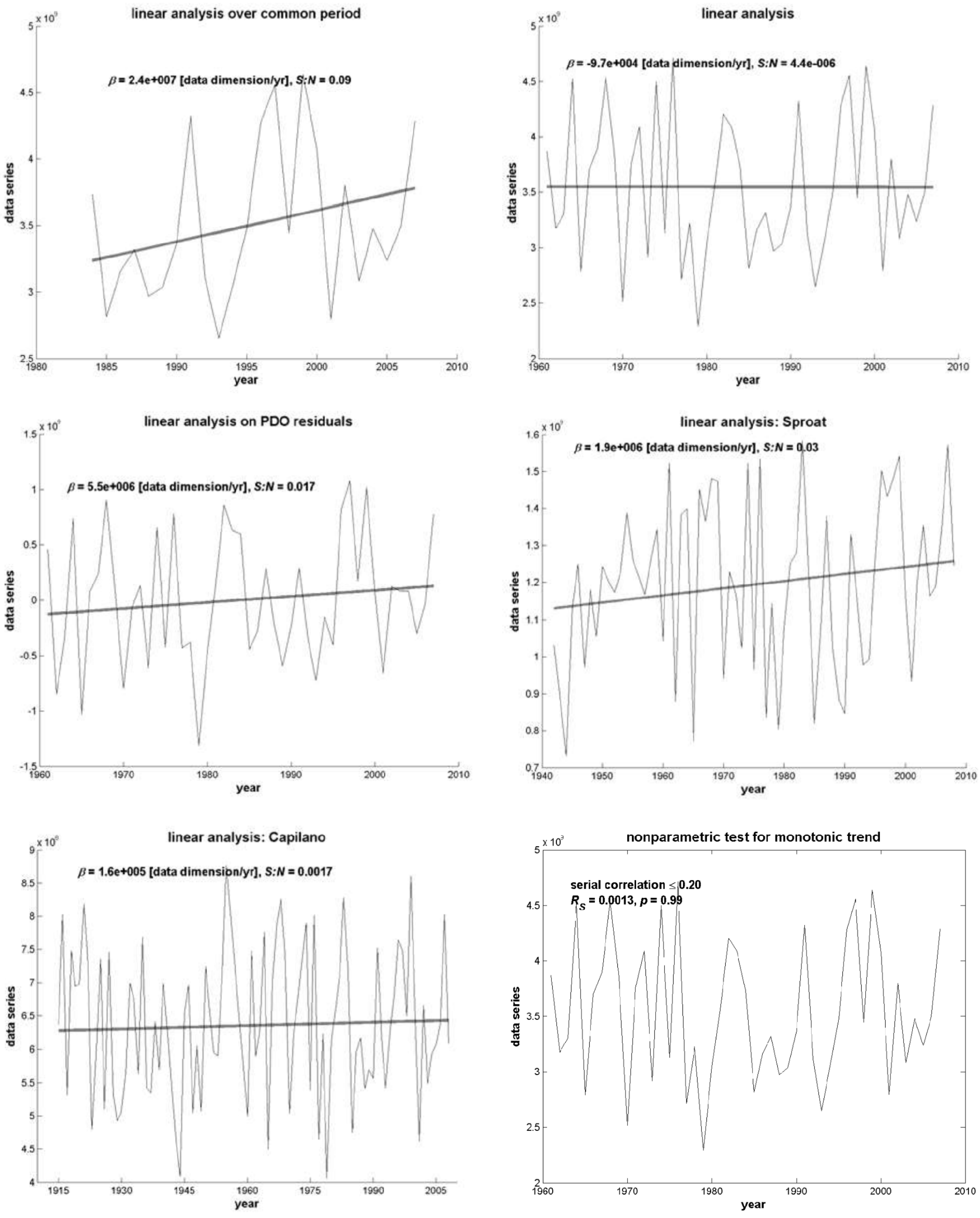


Figure 9 Results of primary and auxiliary analyses for Stave annual inflow volumes in m^3 (continued on next page).

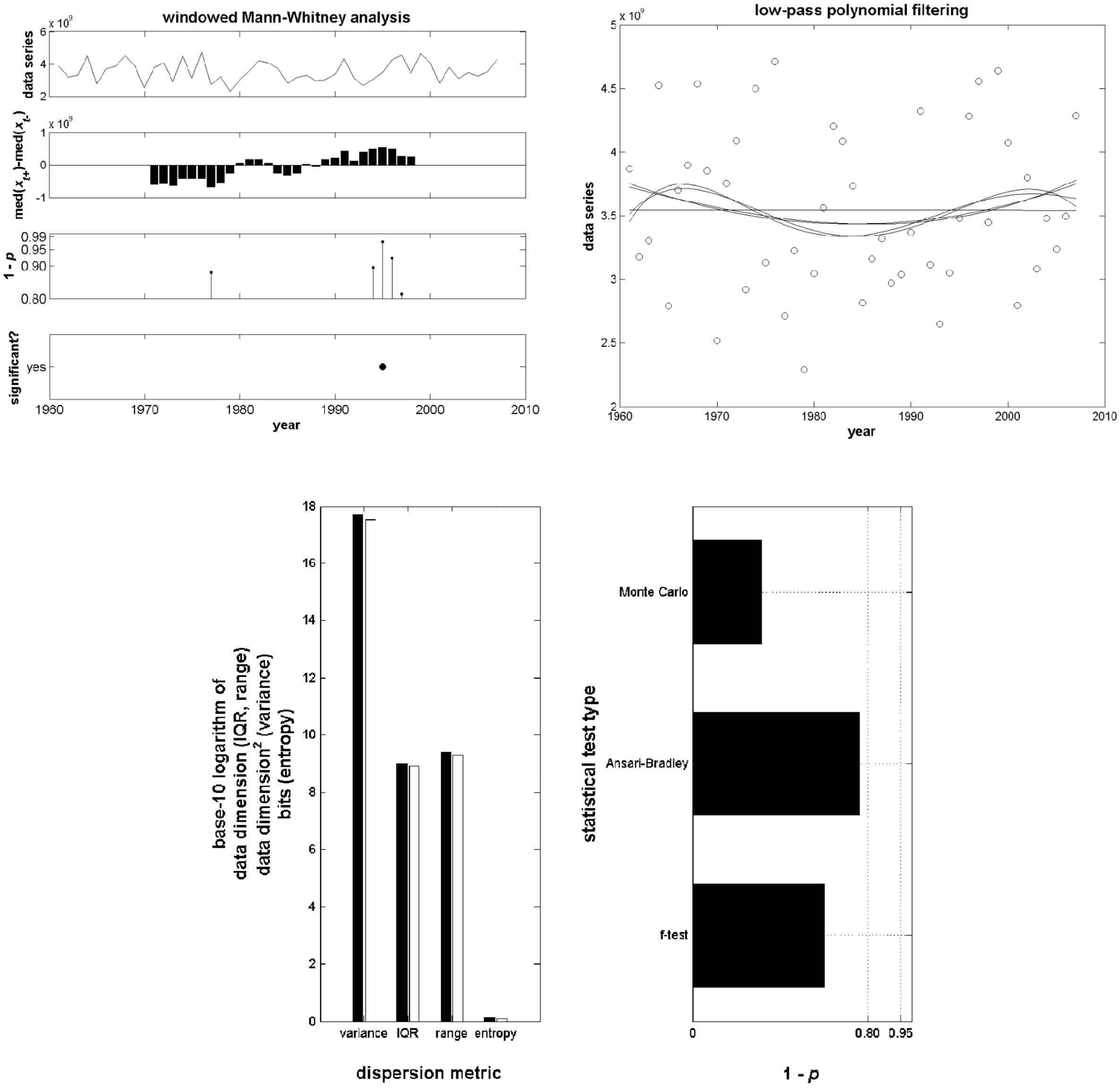


Figure 9 (continued) Results of primary and auxiliary analyses Stave annual inflow volumes in m^3 (continued from previous page).

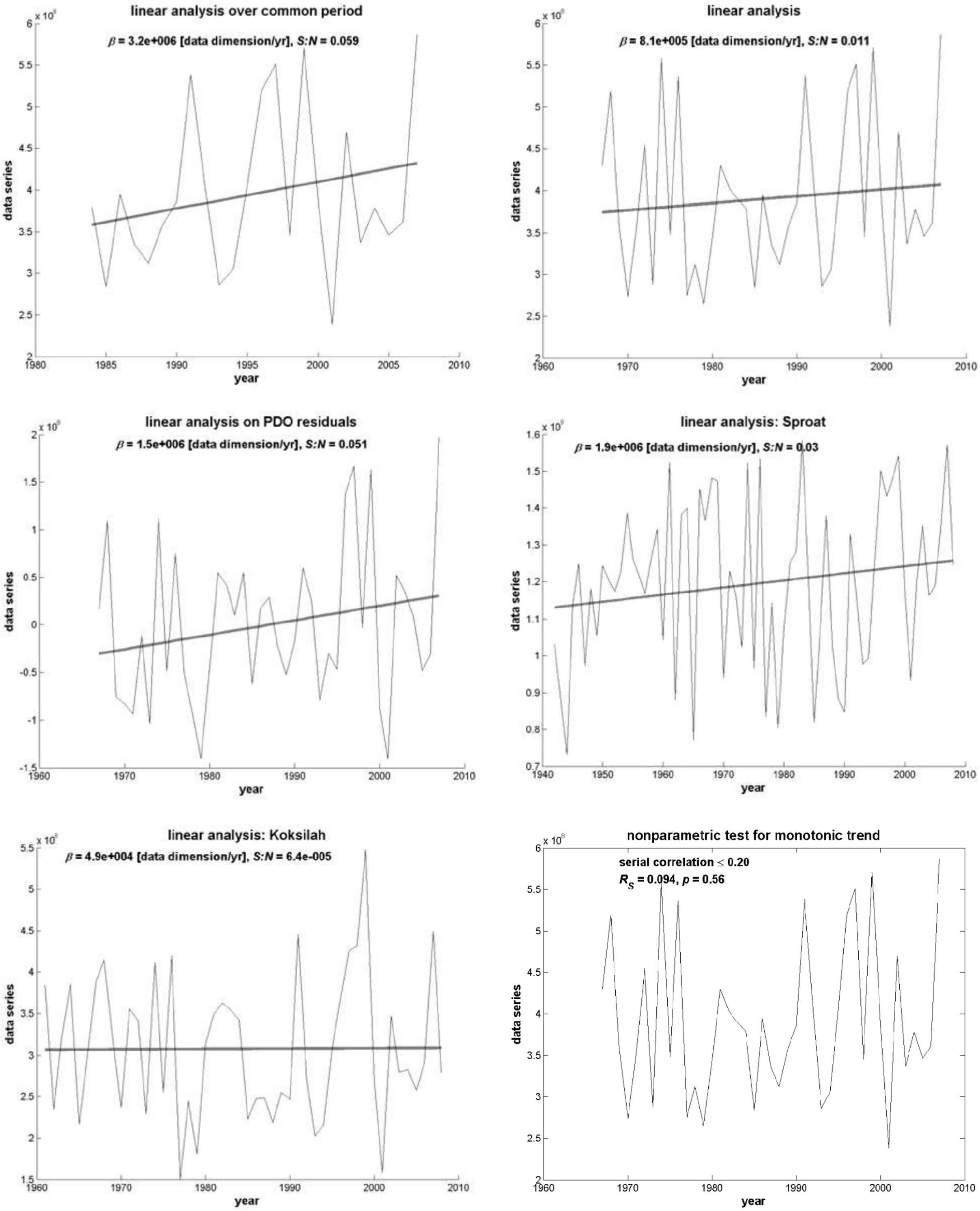


Figure 10 Results of primary and auxiliary analyses for Jordan annual inflow volumes in m^3 (continued on next page).

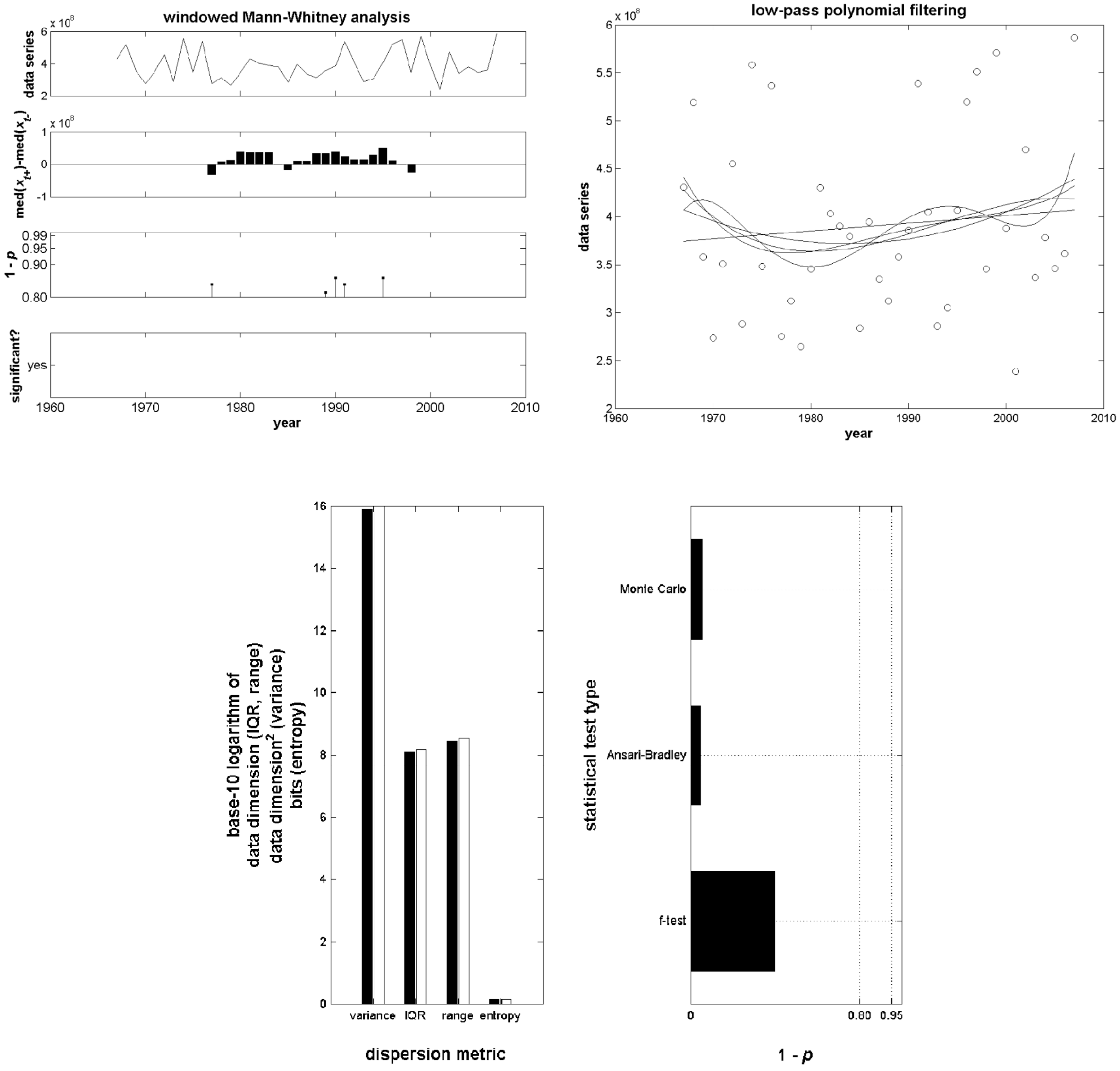


Figure 10 (continued) Results of primary and auxiliary analyses for Jordan annual inflow volumes in m^3 (continued from previous page).

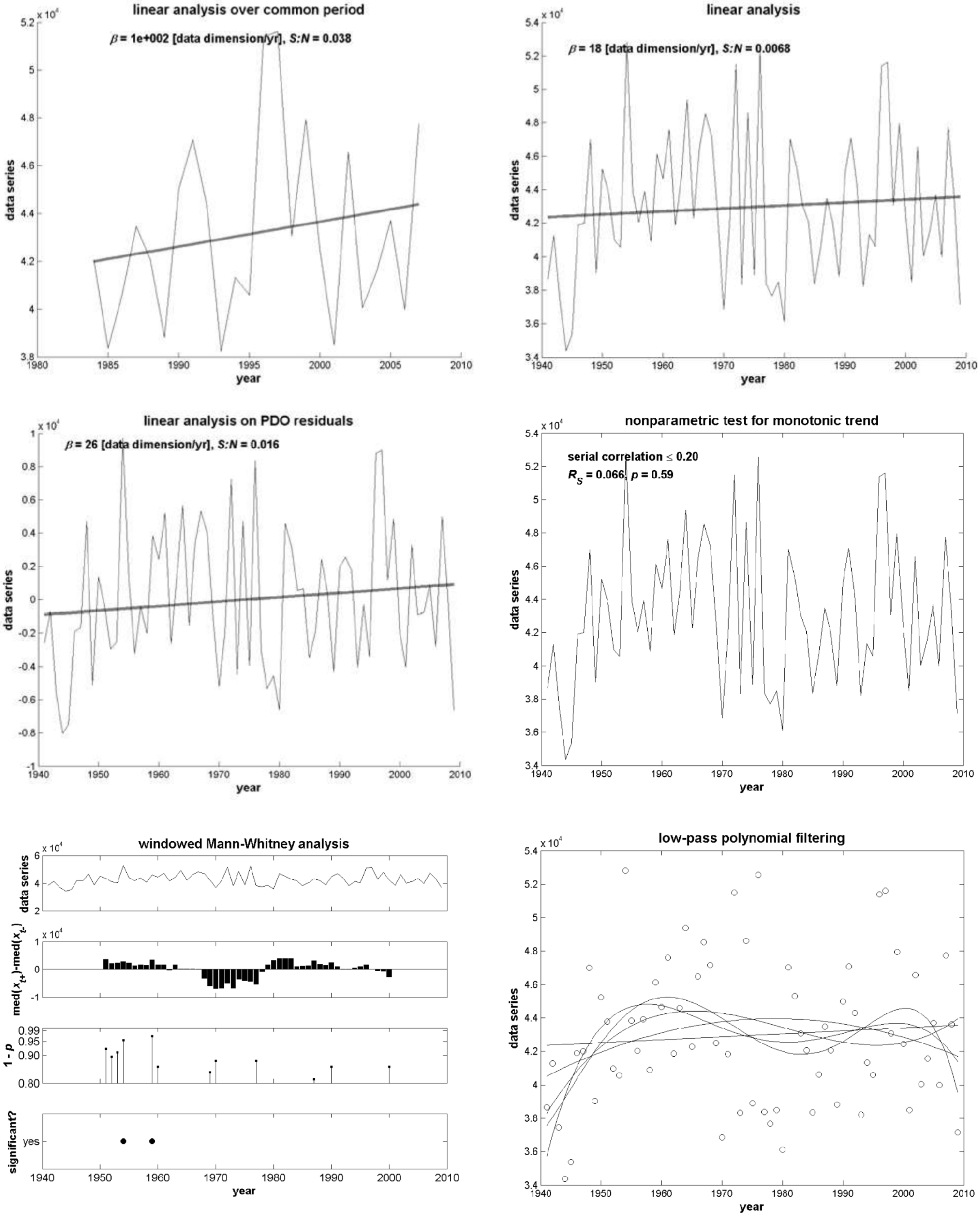


Figure 11 Results of primary and auxiliary analyses for combined usable inflow in approximate equivalent GWh (continued on next page).

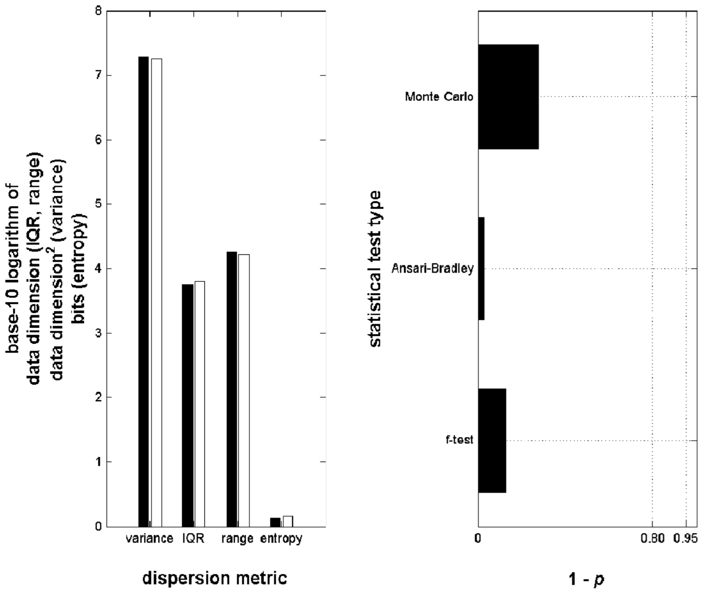


Figure 11 (continued) Results of primary and auxiliary analyses for combined usable inflow in approximate equivalent GWh (continued from previous page).

5 Conclusions and Recommendations

This report presents what appears to be the first thorough study undertaken of long-term historical trends in BC Hydro reservoir inflow volumes. It was motivated largely by concerns over the potential impacts of climatic change. The study was conducted in two phases. The first phase involved application of a relatively straightforward linear analysis method, coupled with direct estimation of signal-to-noise ratio, in a broad reconnaissance analysis of seasonal and annual changes across most of the BC Hydro system. Monthly and annual mean values of local inflows to 20 reservoirs, and combined usable inflow, were considered. Reservoir selection was determined by the availability of carefully quality-controlled and corrected reservoir inflow data from a previous recent project in Hydrology & Technical Services. A common period of record across all the basins (1984-2007) was employed. The second phase of the study involved a detailed assessment of sensitivity of the outcomes to methodological choices around data selection, processing, and analysis. In the interest of managing project scope, the second phase focussed on annual total inflow volume (i.e., total yearly water supply) over the entire period of record available for each of a subset of eight reservoirs representative of different geographic regions and hydroclimatic regimes across the BC Hydro system, in addition to total annual combined usable inflow.

The main findings of the study may be summarized as follows:

- A strong upward trend in autumn-winter flow, especially in December and January, was apparent for combined usable inflow and for 16 out of 20 projects. This trend likely reflects the impact of historically observed and climate model-predicted warmer temperatures, leading to higher rain-to-snow ratios; and perhaps heavier precipitation.
- No clear evidence was found for declining annual water supplies over the historical records, and some evidence was found for increases in some basins. This result is also consistent with atmospheric behaviour observed historically and predicted by climate models for BC, in which precipitation likely remains constant or increases slightly.
- Some other patterns also appeared but are more ambiguous. For instance, a nominal decline in late-summer baseflow was found for most basins which are driven primarily by melt of glaciers and seasonal snowpack, in particular in the Bridge and Columbia regions, but in no case did the trend exhibit a signal-to-noise ratio of 0.10 or higher.
- The foregoing results are consistent with outcomes from GCM-driven studies of the impacts of long-term climate changes upon BC Hydro (or comparable) watersheds. Because the empirical analysis results provide the closest thing available to a ground truth against which modelling studies can be validated, this finding provides some reassurance that the modelling studies are on track.
- Conversely, a number of enigmatic features and inconsistencies were also observed. Although signal-to-noise ratios were low, a pattern of late winter-early spring decline in

inflows to coastal reservoirs may be apparent. This pattern seems difficult to explain physically in terms of current understanding of the atmospheric impacts of climatic changes in BC, and does not appear consistent with GCM-driven hydrologic modelling studies. The earlier melt freshet that one might expect due to warming temperatures, and which is predicted in modelling studies, is not clearly seen in the historical inflow record, at least at the monthly temporal resolution considered here. Additionally, coastal basins are expected on physical grounds and from modelling studies to show a late-summer flow decrease which was not observationally detectable in the FLOCAL data over the historical period.

- Additionally, most of the trends found in this study are quite subtle relative to other time series components, as expressed (for example) by generally low signal-to-noise ratios. Potential explanations include inflow measurement errors; large year-to-year fluctuations, such as those arising from dynamical modes like ENSO or from random variations, which mask long-term shifts; and relatively short inflow records, which compromise the ability to detect the signature of long-term monotonic change. Another pertinent issue in this regard is that modelling studies may artificially inflate the signal-to-noise ratio by suppressing noise – for instance, if all non-trend components of interannual variation are not fully accounted for by the GCM or watershed model, or if ensemble averaging is used for the display and interpretation of the modelling results. Such noise suppression might in some sense be desirable, but it may lead to unrealistic expectations for signal clarity which may be unrealizable in the single trace of historical or future climate, i.e., some signals predicted by model-based studies may be true but so small in amplitude relative to other processes that may be practically undetectable in reality.
- Basic conclusions with respect to annual water supply were found, for the basin subset considered in the auxiliary analysis, to be largely insensitive to dataset selection, processing, and analysis methods. This finding is comforting insofar as it adds further confidence to the results obtained from the primary analysis.
- No evidence was found for a change in the dispersion generally, the variance specifically, or the five-state Shannon entropy of annual total water supply for the basin subset considered in the auxiliary analysis. This result suggests that there has been no significant change in the year-to-year volatility and predictability of water supplies.

Although it is believed that the present study provides a solid view of overall historical trends in BC Hydro reservoir inflows, as always, in the course of performing the study it has become apparent that there is room for improvement and additional work. A number of recommendations for further analysis are as follows:

- The study should be recompleted once every 5-10 yr, for two reasons. First, the historical record will be 5-10 yr longer, which is important because even holding all else constant, the signal-to-noise ratio of monotonic trend increases with record length (Fleming, 2010). Thus, to the extent that long-term trends are an underlying process, a longer dataset will reveal that signal more clearly. Second, an important motivation for conducting historical trend analysis is to serve as a sort of benchmark against which to compare climate change modelling results, and such models evolve considerably over 5-10 yr.

- The vast majority of work performed on climatic and hydroclimatic trend analysis focuses on identifying first-order nonstationarity. However, changes in year-to-year variability are also key, with implications to the reliability and predictability of water supplies and hydroelectric power. We have made a preliminary assessment of such changes in this study, including the introduction of a novel analysis technique using information theory, but much more work should be done. Note that this is a broadly important question of fundamental science which requires attention from the research community at large.
- Better ensuring (or at least understanding) the comparability of data-driven historical hydroclimate studies on the one hand, to model-driven future hydroclimate studies on the other, is very important. Model-based studies are key for interpreting historical hydroclimatic changes, and data-driven studies are key for validating model-predicted hydroclimatic changes. Some important issues to be addressed include the practical question of performing modelling studies for reservoir inflows, rather than headwater locations for instance; and the more fundamental questions of assessing potential *S:N*-inflation issues in modelling studies, and finding more meaningful and representative ways for comparing the single realization of a historical hydroclimatic record to the ensemble output of a modelling study.
- An interesting and potentially valuable extension of the work would be to perform trend analysis upon other relevant hydrologic metrics derived from the reservoir inflow data, such as annual peak flows, freshet timing, ecological low flows, and so forth. However, ability to perform these analyses in a meaningful way may be limited by data quality issues (see also immediately below).
- One of the most important limiting factors on the reliability of empirical trend analysis results is the accuracy of the historical reservoir inflow data used. Sophisticated methods for estimating reservoir inflows, and for quality control and correction of these datasets, are in place. Nevertheless, these data likely remain (in general) of poorer quality relative to those obtained from hydrometric stations, particularly at the very high and very low flows and at the daily timescales required for analysis of changes in extreme hydrologic events. Further improvements to the procedures used to back-calculate and quality-control reservoir inflows would no doubt be useful, but perhaps the single most promising and productive avenue is to improve the reliability of reservoir level measurements, for instance by installing more water level sensors to additively cancel noise and to correct for hydraulic effects such as seiches. Of course, improvements to these data would also have strongly positive implications for other BC Hydro tasks, such as daily and seasonal reservoir inflow forecasting or water license compliance monitoring.

- Maintaining the continuity and consistency of the monitoring network for other hydroclimatic variables is also key. For many of the same reasons listed above, trend analysis of measures like precipitation, temperature, and snowpack (*e.g.*, Rodenhuis *et al.*, 2007) should similarly be recompleted every 5-10 years as new data become available. However, such analyses are only possible if the climate monitoring network maintains its integrity.

Acknowledgements

A BC Hydro project steering committee consisting of F. Weber, S. Smith, and M. Rucker provided valuable feedback. The author additionally thanks A. Gobena for data support.

References

Adelman DE. 2004. Harmonizing methods of scientific inference with the precautionary principle: opportunities and constraints. *Environmental Law Reporter*, 34, 10131-10141.

Amorocho J, Espildora B. 1972. Entropy in the assessment of uncertainty in hydrologic systems and models. *Water Resources Research*, 9, 1511-1522.

Bayazit M, Önöz B. 2007. To prewhiten or not to prewhiten in trend analysis? *Hydrological Sciences Journal*, 52, 611-624.

Braun LN, Weber M, Schulz M. 2000. Consequences of climate change for runoff from Alpine regions. *Annals of Glaciology*, 31, 19-25.

Buishand TA. 1984. Tests for detecting a shift in the mean of hydrological time series. *Journal of Hydrology*, 73, 51-69.

Casassa G, López P, Pouyaud B, Escobar F. 2009. Detection of changes in glacial run-off in alpine basins: examples from North America, the Alps, central Asia, and the Andes. *Hydrological Processes*, 23, 31-41.

Chatfield C. 1996. *The Analysis of Time Series, 5th Ed.* Chapman and Hall, London.

Chen J, Kumar P. 2002. Role of terrestrial hydrologic memory in modulating ENSO impacts in North America. *Journal of Climate*, 15, 3569-3585.

Chen Z, Grasby SE. 2009. Impact of decadal and century-scale oscillations on hydroclimate trend analyses. *Journal of Hydrology*, 365, 122-133.

Cunderlik JM, Burn DH. 2002. Local and regional trends in monthly maximum flows in southern British Columbia. *Canadian Water Resources Journal*, 27, 191-212.

Cunderlik JM, McBean E, Day G, Thiemann M, Kouwen N, Jenkinson W, Quick M, Lence B, Li Y. 2010. *Intercomparison Study of Process-Oriented Watershed Models*. Report prepared for British Columbia Hydro and Power Authority. Conestoga-Rovers and Associates, Richmond, BC.

De Jongh ILM, Verhoest NEC, De Troch FP. 2006. Analysis of a 105-year time series of precipitation observed at Uccle, Belgium. *International Journal of Climatology*, 26, 2023-2039.

Dodds WK. 2002. *Freshwater Ecology*. Academic Press, San Diego, CA

Ducré-Robitaille J-F, Vincent LA, Boulet G. 2003. Comparison of techniques for detection of discontinuities in temperature series. *International Journal of Climatology*, 23, 1087-1101.

Easterling DR, Peterson TC. 1995. A new method for detecting undocumented discontinuities in climatological time series. *International Journal of Climatology*, 15, 369-377.

Eskridge RE, Ku JY, Rao ST, Porter PS, Zurbenko IG. 1997. Separating different scales of motion in time series of meteorological variables. *Bulletin of the American Meteorological Society*, 1473-1483.

Fleming SW. 2007a. Climatic influences on Markovian transition matrices for Vancouver daily rainfall occurrence. *Atmosphere-Ocean*, 45, 163-171.

Fleming SW. 2007b. Artificial neural network forecasting of nonlinear Markov processes. *Canadian Journal of Physics*, 85, 279-294.

Fleming SW. 2007c. An information theoretic perspective on mesoscale seasonal variations in ground-level ozone. *Atmospheric Environment*, 41, 5746-5755.

Fleming SW. 2008. Approximate record length constraints for experimental identification of dynamical fractals. *Annalen der Physik*, 17, 955-969.

Fleming SW. 2009. Exploring the nature of Pacific climate variability using a “toy” nonlinear stochastic model. *Canadian Journal of Physics*, 87, 1127-1131.

Fleming SW. 2010. Signal-to-noise ratios of geophysical and environmental time series. *Environmental and Engineering Geoscience*, 26, 389-399.

Fleming SW, Clarke GKC. 2002. Autoregressive noise, deserialization, and trend detection and quantification in annual river discharge time series. *Canadian Water Resources Journal*, 27, 335–354.

Fleming SW, Clarke GKC. 2003. Glacial control of water resource and related environmental responses to climatic warming: empirical analysis using historical streamflow data from northwestern Canada. *Canadian Water Resources Journal*, 28, 69-86.

Fleming SW, Clarke GKC. 2005. Attenuation of high-frequency interannual streamflow variability by watershed glacial cover. *ASCE Journal of Hydraulic Engineering*, 131, 615-618.

Fleming SW, Moore RD. 2008. *Geophysical trend detection as a signal-to-noise ratio problem, with application to ecological low flows in Cowichan-region watersheds*. North American Lake Management Society Symposium on Lake Management in a Changing Environment, Lake Louise, AB.

Fleming SW, Moore RD, Clarke GKC. 2006. Glacier-mediated streamflow teleconnections to the Arctic Oscillation. *International Journal of Climatology*, 26, 619-636.

Fleming SW, Moore RD, Clarke GKC, Werner AT, Weber FA. 2010. *Projections of Columbia River inflows to Mica dam under potential future trajectories of climate and glacier change*. Canadian Water Resources Association National Conference, 15-18 June 2010, Vancouver, BC.

Fleming SW, Weber FA, Weston S. 2010. *Multiobjective, manifoldly constrained Monte Carlo optimization and uncertainty estimation for an operational hydrologic forecast model*. American Meteorological Society Annual Meeting, Atlanta, GA. Extended abstract online at ams.confex.com/ams/90annual/techprogram/paper_160170.htm.

Fleming SW, Whitfield PH, Moore RD, Quilty EJ. 2007. Regime-dependent streamflow sensitivities to Pacific climate modes across the Georgia-Puget transboundary ecoregion. *Hydrological Processes*, 21, 3264-3287.

Franzke C. 2009. Multi-scale analysis of teleconnection indices: climate noise and nonlinear trend analysis. *Nonlinear Processes in Geophysics*, 16, 65-76.

Gobena AK. 2010. *Teleconnections between Large-Scale Climate Modes and the Hydroclimate of BC Hydro Watersheds*. Technical Report, Hydrology and Technical Services, BC Hydro, Burnaby, BC.

Gobena A, Weber FA, Fleming SW. 2010. *Teleconnections between large-scale climate modes and the hydroclimatic data of BC Hydro watersheds*. Canadian Water Resources Association Annual Conference, Vancouver, BC.

Harvey KD, Pilon PJ, Yuzyk TR. 1999. *Canada's Reference Hydrometric Basin Network (RHBN)*. Canadian Water Resources Association Annual Conference, Halifax, NS.

Hodgkins GA. 2009. Streamflow changes in Alaska between the cool phase (1947-1976) and the warm phase (1977-2006) of the Pacific Decadal Oscillation: the influence of glaciers. *Water Resources Research*, 45, doi: 10.1029/2008WR007575.

Jeong Y, Sanders BF, Grant SB. 2006. The information content of high-frequency environmental monitoring signals pollution events in the coastal ocean. *Environmental Science and Technology*, 40, 6215-6220.

JISAO, 2010. *PDO Index*. Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Seattle. http://jisao.washington.edu/pdo/PDO.latest.

Kallache M, Rust HW, Kropp J. 2005. Trend assessment: applications for hydrology and climate research. *Nonlinear Process in Geophysics*, 12, 201–210.

Khaliq MN, Ouarda TBMJ, Gachon P, Sushama L, St-Hilaire A. 2009. Identification of hydrological trends in the presence of serial and cross correlations: a review of selected methods and their application to annual flow regimes of Canadian rivers. *Journal of Hydrology*, 368, 117-130.

Krasovskaia I. 1995. Quantification of the stability of river flow regimes. *Hydrological Sciences Journal*, 40, 587-598.

Kundzewicz ZW, Robson AJ. 2004. Change detection in hydrological records – a review of the methodology. *Hydrological Sciences Journal*, 49, 7-19.

Lafrenière M, Sharp M. 2003. Wavelet analysis of inter-annual variability in the runoff regimes of glacial and nival stream catchments, Bow Lake, Alberta. *Hydrological Processes*, 17, 1093-1118.

Lanzante JR. 1996. Resistant, robust and non-parametric techniques for the analysis of climate data: theory and examples, including applications to historical radiosonde station data. *International Journal of Climatology*, 16, 1197-1226.

Lee AFS, Heghinian SM. 1977. A shift of the mean level in a sequence of independent normal random variables – a Bayesian approach. *Technometrics*, 19, 503-506.

Livezey RE, Chen WY, 1983. Statistical field significance and its determination by Monte Carlo techniques. *Monthly Weather Review*, 111, 46-59.

Mahajan S. 2010. *Street-Fighting Mathematics: the Art of Educated Guessing and Opportunistic Problem Solving*. MIT Press, Cambridge, MA.

Mantua NJ, Hare SR, Zhang Y, Wallace JM, Francis RC. 1997. A Pacific interdecadal climate oscillation with impacts on salmon production. *Bulletin of the American Meteorological Society*, 78, 1069-1079.

Mauget SA. 2003. Multidecadal regime shifts in U.S. streamflow, precipitation, and temperature at the end of the twentieth century. *Journal of Climate*, 16, 3905-3916.

McCloskey D. 1995. The insignificance of statistical significance. *Scientific American*, 272, 32-33.

Moore RD, Jost G, Radic V, Anslow F, Jarosch A, Clarke GKC, Menounos B, Wheate R, Murdock T, Werner A. 2010. *Past and future contributions of glacier melt to Columbia River streamflow*. Canadian Geophysical Union/Canadian Meteorological and Oceanographic Society Joint Congress, 31 May – June 4, 2010, Ottawa, ON.

Moore RD, Fleming SW, Menounos B, Wheate R, Fountain A, Stahl K, Holm K, Jakob M. 2009. Glacier change in western North America: influences on hydrology, geomorphic hazards and water quality. *Hydrologic Processes*, 23, 42-61.

Neal EG, Walter MT, Coffeen C. 2002. Linking the Pacific Decadal Oscillation to seasonal discharge patterns in southeast Alaska. *Journal of Hydrology*, 263, 188-197.

Nicholls N. 2000. The insignificance of significance testing. *Bulletin of the American Meteorological Society*, 81, 981-986.

Pierce JR. 1980. *An Introduction to Information Theory: Symbols, Signals, and Noise, 2nd Ed*. Dover, New York.

Radionov SN. 2004. A sequential algorithm for testing climate regime shifts. *Geophysical Research Letters*, 31, L09204, doi:10.1029/2004GL019448.

Rao ST, Zurbenko IG, Neagu R, Porter PS, Ku JY, Henry RF. 1997. Space and time scales in ambient ozone data. *Bulletin of the American Meteorological Society*, 78, 2153-2166.

Rodenhuis DR, Bennett KE, Werner AT, Murdock TQ, Bronaugh D. 2007. *Hydro-climatology and future climate impacts in British Columbia.* Pacific Climate Impacts Consortium, University of Victoria, Victoria, BC.

Schnorbus M, Bennett KE, Werner A, Berland A. 2010. *Hydrologic impacts in BC watersheds: Peace, Campbell, and Upper Columbia.* Presentation at joint PCIC-BC Hydro workshop, Assessing Hydrologic Impacts on Water Resources in BC: Current Accomplishments and Future Vision, 20 April 2010, Burnaby, BC. Available online at http://www.pacificclimate.org/resources/presentations/.

Shannon CE. 1948. A mathematical theory of communication. *The Bell System Technical Journal*, 27, 379-423, 623-656.

Shannon CE, Weaver W. 1949. *The Mathematical Theory of Communication*. University of Illinois Press, Urbana. Reprinted in 1963 and 1998.

Spiegelhalter K, Stahl K. 2010. Modelling the coupled influence of climate and glacier change on discharge. *European Geosciences Union General Assembly*, 2-7 May 2010, Vienna, Austria.

Stahl K, Moore RD. 2006. Influence of watershed glacial coverage on summer streamflow in British Columbia, Canada. *Water Resources Research*, 42, doi: 10.1029/2006WR005022.

Stahl K, Moore RD, Shea JM, Hutchinson D, Cannon AJ. 2008. Coupled modelling of glacier and streamflow response to future climate scenarios. *Water Resources Research*, 44, doi: 10.1029.2007WR005956.

Telford WM, Geldart LP, Sheriff RE. 1990. *Applied Geophysics, 2nd Ed.* Cambridge University Press, Cambridge, UK.

Vingarzan R, Taylor B. 2003. Trend analysis of ground level ozone in the greater Vancouver/Fraser Valley area of British Columbia. *Atmospheric Environment*, 37, 2159-2171.

von Storch H. 1995. Misuses of statistical analysis in climate research. In: *Analysis of Climate Variability: Applications of Statistical Techniques*. H. von Storch and A. Navarra (Eds). Springer-Verlag, Berlin, pp. 11–26.

Whitfield PH. 2001. Linked hydrologic and climate variations in British Columbia and Yukon. *Environmental Monitoring and Assessment*, 67, 217-238.

Whitfield PH, Cannon AJ. 2000. Recent variations in climate and hydrology in Canada. *Canadian Water Resources Journal*, 25, 19-65.

Whitfield PH, Moore RD, Fleming SW, Zawadzki A. 2010. Pacific Decadal Oscillation and the hydroclimatology of western Canada: review and prospects. *Canadian Water Resources Journal*, 35, 1-27.

Whitfield PH, Taylor E. 1998. *Apparent recent changes in hydrology and climate of coastal British Columbia.* Canadian Water Resources Association Annual Conference, Victoria, BC.

WSC, 2010a. *Archived Hydrometric Data*. Water Survey of Canada, Environment Canada, Ottawa. http://www.wsc.ec.gc.ca/hydat/H2O/index_e.cfm?cname=main_e.cfm.

WSC, 2010b. *Station Information*. Water Survey of Canada, Environment Canada, Ottawa. http://www.wsc.ec.gc.ca/hydex/main_e.cfm?cname=StationList_e.cfm

Yue S, Pilon P, Cavadias G. 2002b. Power of the Mann-Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. *Journal of Hydrology*, 259, 254-271.

Yue S, Pilon P, Phinney B, Cavadias G. 2002a. The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrological Processes*, 16, 1807–1829.

Yue S, Wang C. 2004. Reply to comment by Xuebin Zhang and Francis W. Zwiers on "Applicability of prewhitening to eliminate the influence of serial correlation on the Mann-Kendall test." *Water Resources Research*, 40, W03806, doi: 10.1029/2003WR002547.

Zhang X, Harvey KD, Hogg WD, Yuzyk TR. 2001. Trends in Canadian streamflow. *Water Resources Research*, 37, 987-998.

Zhang X, Vincent LA, Hogg WD, Niitsoo A. 2000. Temperature and precipitation trends in Canada during the 20th century. *Atmosphere-Ocean*, 38, 395-429.

Zhang X, Zwiers, FW. 2004. Comment on "Applicability of prewhitening to eliminate the influence of serial correlation on the Mann-Kendall test" by Sheng Yue and Chun Yuan Wang. *Water Resources Research*, 40, W03805, doi:10.1029/2003WR002073.

Zurbenko I, Porter PS, Rao ST, Ku JY, Gui R, Eskridge RE. 1996. Detecting discontinuities in time series of upper-air data: development and demonstration of an adaptive filter technique. *Journal of Climate*, 9, 3548-3560.

Appendix

As noted, there is a lack of consensus in the literature regarding how to best deal with the impacts of serial correlation upon null-hypothesis significance tests for monotonic trend. Broadly speaking, to assess the various methods proposed to date, to decide which is best used (or not used) in some particular practical application, and ultimately to define directions for future work, it is clearly desirable to have some appreciation for the basic issues at play. However, despite the often-substantial statistical rigour and sophistication of the literature on the subject, in general it seems surprisingly vague about what, fundamentally, is behind either the departure of real from nominal significance levels when serial correlation is present, or the problems with common methods for correcting that behaviour. Here, we offer relatively intuitive and direct interpretations of those two issues. These interpretations may not be unique, and they certainly do not constitute the final words on these matters. Nevertheless, we believe they are valid and useful perspectives which have not been very clearly expressed previously if at all.

We begin with the question of why autocorrelation can create problems for statistical tests for monotonic trend. The basic notion behind this approach to trend detection is as follows. Any observational time series will have a trend which (if measured to sufficiently high precision) is non-zero. To decide whether this trend is significant or not, one considers how likely it is that a trend of the observed magnitude would have happened by accident in a white-noise dataset. More specifically, it is assumed that there are two possibilities. The null hypothesis, H_0, holds that no trend is present, or more precisely, that the time series consists of uncorrelated random numbers – that is, white noise. The alternative hypothesis, H_1, holds that a structural monotonic trend is present above and beyond any white noise. Through some mechanism, which depends on the specific methodology of a given test technique, a p-value is estimated from the dataset. This value is the probability that one would be wrong in rejecting H_0 in favour of H_1. A critical p-value, called the significance level of the test, is defined. By convention 0.05 (sometimes denoted 5%, or incorrectly as a "confidence level" of 95%) is normally used, although the choice is somewhat arbitrary and other possibilities are sometimes employed. If p estimated for the dataset is less than this critical value, the trend is declared significant. A potential problem here is when the background noise is red (see immediately below) rather than white. It is widely acknowledged that in this case, the p-value estimated by a procedure assuming white noise is lower than the real value for an autocorrelated noise series, leading to a higher-than-reported risk of incorrectly concluding trend to be present.

Why does this happen? A useful interpretation is that the likelihood of a chance trend is higher for a red-noise process than for a white-noise process. A red-noise process (so-called because its Fourier spectrum contains more power at low frequencies than at high frequencies) is a stochastic process in which successive data values are related. The classic example is the first-order autoregressive process, denoted AR(1), which arises directly from the governing physics in some applications, and which is also frequently used as an approximate parameterization of memory (see for example Fleming, 2008 and references therein):

$$x_{i+1} = \alpha\, x_i + \varepsilon_{i+1} \tag{A.1}$$

In Equation A.1, $x_i \; \forall \; i = 1,N$ are the consecutive values of the time series, $0 < \alpha < 1$ is the lag-1 serial correlation coefficient, and $\varepsilon_i \; \forall \; i = 1,N$ is white (serially uncorrelated) noise. Because each successive value of x shows a direct linear relationship to the preceding one, the current value “remembers” prior values (as such, it is a type of Markov process). One physical implication of the fact that each value is related to the previous one is that, while the total dataset variance is not decreased, the severity of fluctuations from one time step to the next is: subsequent values tend to track together, at least for a while before a particularly sizable random shock from the ε term kicks the process to a different part of the state space. This in turn tends to give rise to long runs of values above or below the mean, or long runs of increasing or decreasing values – that is, trends. One the one hand, such trends are chance occurrences, rather than a reflection of some underlying trending structure, such as a long-term, unidirectional climate change signal. On the other hand, this dynamic tends to produce stronger and more numerous chance trends than those which occur in a white noise process. The consequence is that, if the noise is in fact red instead of white, a significance test based on the assumption of a white-noise null hypothesis underestimates the likelihood of a chance trend of the same magnitude as the observed trend, giving an elevated rate of false positive trend detections.

Figures A.1, A.2, and A.3 illustrate some of these considerations at work. Figure A.1 shows a synthetic AR(1) process generated using $\alpha = 0.8$ and $\varepsilon \sim \mathcal{N}(0,1)$, with x rescaled after-the-fact to zero mean and unit variance for convenience. We can see a clear downward trend in the dataset. The trend is real, but it is does not reflect any specific underlying structural process that could be expected to result in its indefinite continuation into the future: this is readily seen in Figure A.2, which is precisely the same time series, but run forward for several hundred more data points. The red noise-induced temporary downward trend is replaced by a temporary upward trend, in turn replaced at about $i \sim 250$ by broad oscillations. Also compare this behaviour to a white-noise series, $\varepsilon \sim \mathcal{N}(0,1)$, an example of which is illustrated in Figure A.3. Chance trends can also occur in a white-noise process – this is, in fact, assumed by the standard statistical testing procedure – but we can clearly see that such chance trends are much smaller, more short-lived, and less clear than those apparent in a red-noise series. If we were to use a time series of the type illustrated in Figure A.3 to estimate the probability of a chance trend, but in reality the noise was of the type illustrated in Figures A.1 and A.2, we would underestimate that likelihood. This in turn would corrupt our attempts to estimate the *p*-value in the usual null-hypothesis tests for the significance of monotonic trend.

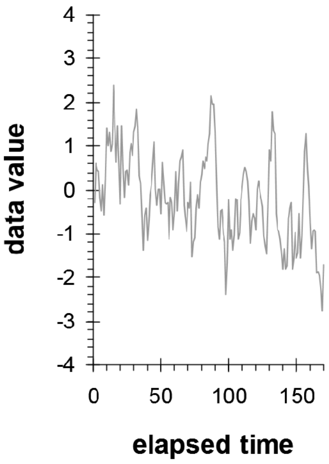


Figure A.1 Illustration of an AR(1) process

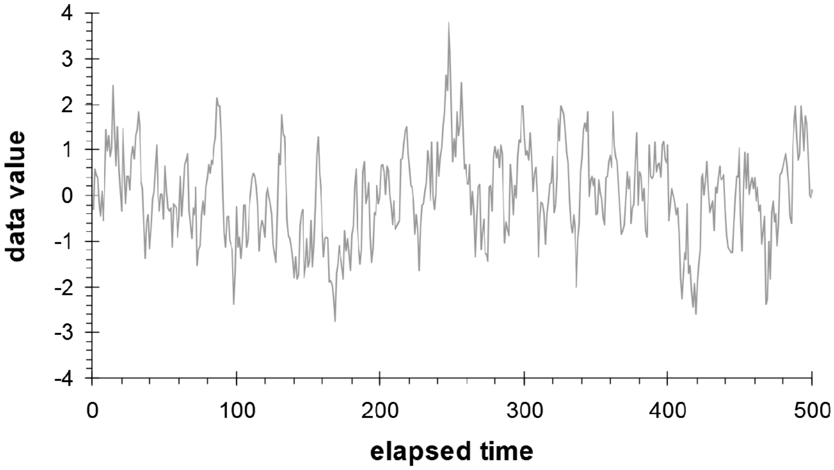


Figure A.2 Illustration of the same AR(1) process, carried further forward in time

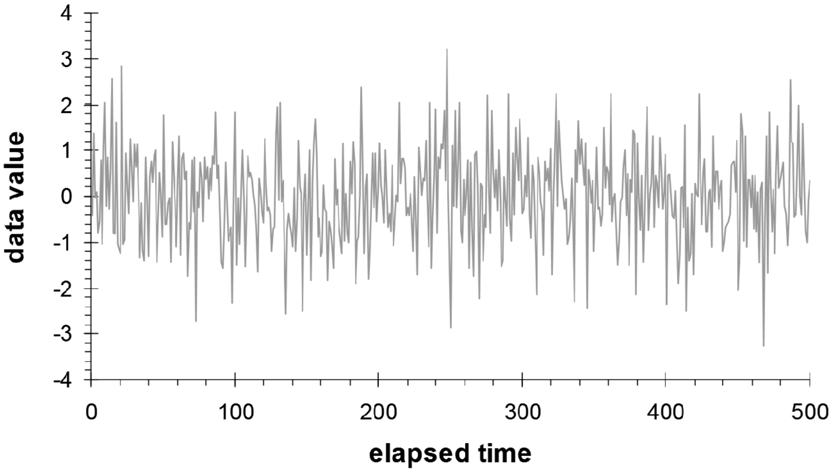


Figure A.3 White-noise process with same mean and variance as AR(1) process shown above

Several methods have been proposed in the statistical climatology and hydrology literature for circumventing this limitation. The most broadly employed of these, which also serves in one way or another as a core component of most other methods, is to remove the red noise from the data prior to analysis in a procedure called prewhitening (von Storch, 1995):

$$^{PW}x_{i+1} = x_{i+1} - \alpha\, x_i \tag{A.2}$$

Here, x is the observed data series, ^{PW}x is the prewhitened data upon which the significance test for trend will then be performed, and α is again the serial correlation coefficient, in this case estimated directly from the observational dataset as the Pearson linear correlation coefficient between the time series and a lagged copy of itself (*e.g.*, Chatfield, 1996):

$$\alpha = \frac{\sum_{i=1}^{N-1}(x_i - \bar{x})(x_{i+1} - \bar{x})}{\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{A.3}$$

On the conditions that the dataset, x, consists of an AR(1) process and (in particular) that the observed value of α for this dataset reflects the AR(1) process parameter as in Equation A.1, then it is clear that the prewhitening operation of Equation A.2 simply “undoes” the serial dependence captured in (A.1). That is, Equation A.2 renders the problem compliant again with the framework assumed by standard statistical tests for trend, and in particular, the null hypothesis of uncorrelated random samples.

But what is the impact, upon statistical tests for trend significance, of using prewhitening when a trend actually is present? Several workers have previously demonstrated through numerical experiments that it can produce errors of opposite direction to those encountered when AR(1) noise is present without correction: a potentially large decrease in the statistical power of the test is seen, that is, an increase in the frequency of false negative trend detections. The problem with prewhitening lies in the fact that the aforementioned conditions upon which it is premised may very commonly not be met. In particular, a structural trend induces serial correlation in a dataset, and this observed serial correlation has nothing to do with an AR(1) process. Thus, if H_1 is true, then the prewhitening procedure of (A.2) removes serial dependence even if there is no red noise present – an effect very different from what was intended.

It seems fundamentally important to demonstrate and understand how autocorrelation is produced by a trending process. We propose that this can be accomplished as follows. Conceptually, it is not difficult to envision that (for example) an upward trend might induce some kind of serial dependence-like behaviour: because there is a progressive upward drift in the data, two points close together in time are more likely to have values similar to each other than two points located far apart from each in time. We can more formally derive this result, however. Our concern here lies with the sample statistics, and one can always calculate and remove the average value for a finite-length observational dataset, even if the driving process is first-order nonstationary and the population mean therefore constantly changes (as is the case for a monotonic trend). We can therefore assume for convenience that the sample mean is zero. Then from Equation A.3 we have:

$$\alpha = \frac{\sum_{i=1}^{N-1}(x_i)(x_{i+1})}{\sum_{i=1}^{N}(x_i)^2} \tag{A.4}$$

Next assume a dataset which consists solely of linear temporal trend, $x(t) = \beta_0 + \beta_1 t$, and which is evenly sampled in time. This relationship may be recast as an iterated map, $x_{i+1} = x_i + \beta_1 \Delta t$. Taking for convenience the sampling rate, Δt, to be unity (say, $[t]$ = yr for an annual flow record), and substituting the resulting expression for x_{i+1} into (A.4), gives after some algebra:

$$\alpha = \frac{\sum_{i=1}^{N-1}(x_i)^2 + \beta_1 \sum_{i=1}^{N-1} x_i}{\sum_{i=1}^{N}(x_i)^2} \tag{A.5}$$

If the summations of x_i^2 over N and N-1 are about the same, which they should be for large sample sizes, then (A.5) reduces to:

$$\alpha = 1 + \frac{\beta \sum_{i=1}^{N-1} x_i}{\sum_{i=1}^{N} (x_i)^2} \tag{A.6}$$

However, if $\bar{x}$ =0 as assumed above, then the summation of x_i over large N goes to zero, giving our final result:

$$\alpha \sim 1 \tag{A.7}$$

Thus, the lag-1 serial correlation coefficient of a noise-free linear trend is a non-zero positive value, and in particular, unity. Numerical experiments, not shown here for conciseness, confirm this analytical result, and also show that the observational autocorrelation coefficient remains non-zero (but declines in magnitude) as white noise is added to the trend.

The implication of this finding for the validity of the prewhitening procedure should be clear. When a trend is in fact present in the data, a non-zero observational serial correlation coefficient will occur, with a magnitude that increases with the signal-to-noise ratio of the trend. In such a case, the prewhitening procedure misattributes serial correlation to an AR(1) process, and in fact proceeds to remove a non-existent AR(1) process, in what amounts to a deliberate corruption of the data. The net effect, as has been previously demonstrated by others (*e.g.*, Fleming and Clarke, 2002) is to reduce the ability of the test to detect real trends and to incorrectly reduce the estimated magnitudes of those trends.

Glacier and Streamflow Response to Future Climate Scenarios, Mica Basin, British Columbia

Final Report Prepared for BC Hydro

21 April 2011

R.D. (Dan) Moore, Georg Jost
Department of Geography
The University of British Columbia

Garry K.C. Clarke, Faron Anslow, Valentina Radic
Department of Earth and Ocean Sciences
The University of British Columbia

Brian Menounos, Roger Wheate
Geography Program
University of Northern British Columbia

Arelia Werner, Trevor Murdock
Pacific Climate Impacts Consortium, Victoria

Alex Cannon
Environment Canada, Pacific and Yukon Region

Executive Summary

This project was motivated by long-term planning needs of the Generation Resource Management group at BC Hydro, particularly in relation to the upcoming review of the current Columbia River Treaty and the associated need to understand how reservoir inflows may change in response to a range of scenarios for future climatic conditions. The Climate Impacts Group at the University of Washington is applying the Variable Infiltration Capacity (VIC) model to make these required projections. However, the VIC model does not contain a glacier routine, so an alternative model must be used to assess the contributions of glacier runoff to Columbia River streamflow and how it may change under future climate scenarios.

Glaciers in the Columbia basin are currently experiencing negative mass balance and, even if the current climate continued with no further warming, glaciers would likely continue to retreat for one or more decades. The overall objective of this study was, therefore, to develop and apply computational tools to generate projections of (a) future glacier coverage and (b) streamflow, including contributions from glacier melt, given a range of future climate scenarios. The study focused specifically on generating scenarios for inflows to the Mica Reservoir. The work was conducted collaboratively by five groups: (1) UBC Geography (R.D. Moore, Georg Jost); (2) UBC Earth and Ocean Sciences (Garry Clarke, Faron Anslow, Valentina Radiç); (3) UNBC Geography (Brian Menounos, Roger Wheate); (4) Pacific Climate Impacts Consortium (PCIC) (Trevor Murdock, Arelia Werner); and (5) Environment Canada, Pacific and Yukon Region (Alex Cannon).

This project employed the UBC Regional Glaciation Model (UBC-RGM), which has been recently developed by Garry Clarke and co-workers in the Department of Earth and Ocean Sciences at UBC, along with the HBV-EC hydrologic model, which was previously developed as a collaboration between scientists at Environment Canada Pacific and Yukon Region and the University of British Columbia. This study employed the version of HBV-EC that was incorporated into the Green Kenue modelling platform by the Canadian Hydraulics Centre. Models were calibrated and tested using historic weather and climate data as forcing. Calibration and testing of the glaciation model was based on available historic mass balance records from glaciers in British Columbia and the Alberta Rockies. The hydrologic model was calibrated using historic streamflow data and changes in glacier volume. Effects of uncertainty in the calibrated parameters and in recent glacier volume change were assessed using an approach based on the Generalized Likelihood Uncertainty Estimation procedure.

The potential effects of future climate scenarios were simulated using output from Global Climate Models (GCMs) using a variety of greenhouse gas emission scenarios from the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report. Selection of candidate GCMs was based on their ability to reproduce climate for both western North America and globally, based on air temperature, precipitation and synoptic-scale sea-level pressure patterns. Six GCMs met all selection criteria: Canadian Global Coupled Model Version 3, with two resolutions – CGCM3.1 (T47) and CGCM3.1 (T63); the Model for Integrated Research on Climate – MIROC3.2 (hires), abbreviated here as MIRO; and models developed at the Max Planck Institutes for Meteorology (ECHAM5/MPI-OM, abbreviated here as ECHAM), the Commonwealth Scientific and Industrial Research Organization (CSIRO-Mk3.0), and the Geophysical Fluid Dynamics Laboratory (GFDL-CM2.0). Output was obtained for the following greenhouse gas emission scenarios: A1B, B1 and A2. Selection of scenarios was constrained by the availability of daily output, as this time resolution is required to drive the hydrologic model. Daily output through the twenty-first century was available for CGCM3.1 (T47), CGCM3.1 (T63) and ECHAM. For all the other GCMs, daily output was only available for two time slices, 2050-2065 and 2085-2100.

UBC-RGM requires spatially distributed forcing at a typical resolution of 200 m, which was the resolution used in this study. Considering the large modelling domain involved in this project, it was not computationally feasible to generate spatially and temporally distributed transient weather forcings from future climate scenarios. As an alternative, a baseline set of distributed mass balance forcings was computed using historic information; this baseline mass balance field was then adjusted on the basis of projected air temperature and precipitation from the GCM output using a modified "delta" approach. In this approach, the "deltas" were the differences between the GCM's mean monthly climatology over the chosen historic baseline period and the transient monthly GCM run for 2001-2100. The time series of "deltas" were then added to monthly climatologies from a climate reanalysis dataset. In this way, the glacier mass balance model was not directly driven by GCM output but by transient delta-adjusted climate reanalysis projected through the 21st century. For each future climate scenario, digital maps of glacier extents and elevations were output every ten years.

HBV-EC uses data from one or more weather stations, which are then interpolated/extrapolated over the catchment using routines built into the model, based on the definition of climate zones and vertical gradients for temperature and precipitation within each zone. In this application, five stations within and near the Mica basin were used. Therefore, the downscaling approach had to generate daily weather data at the locations of the five weather stations used to calibrate and test HBV-EC. Downscaling from GCM output to the weather station locations was conducted using the TreeGen routine developed by Alex

Cannon of Environment Canada, Pacific and Yukon Region, which was used successfully in a similar application in the Bridge River catchment. Because TreeGen incorporates information on synoptic-scale conditions at a daily resolution, it has a key advantage over "delta" approaches for making hydrologic projections: it can explicitly accommodate changes in the sequencing of weather systems, which can be critical controls on streamflow generation. The application of TreeGen in this study included the use of an option that ensures that the trends in downscaled weather data are consistent with the trends in the GCM output, so that these trends are consistent with those used in the climate forcing for the glaciation model.

Like most hydrologic model codes, HBV-EC treats land cover as static. To accommodate changes in the glacier extents and elevations through time, HBV-EC was run using scripts that would stop the simulation every ten years, read in the new glacier extents and elevations, update the definitions of Grouped Response Units and state variables, then continue the simulation, including a spin-up period.

All scenarios indicate continued warming to the year 2100. Depending on the choice of GCM and emission scenario the range of temperature increase by the 2050s is from 1.1 °C to 3.7 °C, with a mean increase of 2.1 °C. Projected temperature increases for the 2001-2100 period range from 1.4 °C to 5.6 °C, with a mean increase of 3.4 °C. Projected precipitation changes for 2001-2100 range from a 4% increase to a 25% increase, with a mean increase of 14%.

Glacier cover was between 5 and 6% of the catchment area in 2000. Glaciers in the Mica basin are projected to continue retreating under all of the future climate scenarios. Depending on the GCM and emission scenario, glacier cover in 2100 is projected to decrease by 44% to 100% of the 2000 coverage, with a mean decrease of 93%.

The future warming is projected to generate an earlier onset of spring melt and lower flows in late summer and early autumn, consistent with other studies focused on climate change impacts on streamflow in snow-dominated catchments. Glacier ice melt currently contributes between 3 and 9% of annual runoff, over and above contributions from snowmelt and rainfall runoff. Ice melt contributions are more important in August and September, when they range up to 25% and 35% of monthly runoff, respectively. Ice melt contributions to annual runoff decline in all future scenarios. Ice melt contributions to August streamflow also decline under most future climate scenarios, although some increases are projected for the 2050-2065 time slice under the B2 emission scenario, for which glacier retreat is not as extreme as for other scenarios. The decrease in ice melt contributions to August streamflow exacerbates the effect of an earlier snowmelt in producing low flows in late summer.

There is substantial uncertainty in the future projections, arising from (1) variations among GCMs, (2) variations among emission scenarios, and (3) uncertainty inherent with the use of any geophysical model, captured here as uncertainty in the calibrated parameters in the hydrologic model. The uncertainty in recent glacier volume change is also captured by the modelling approach used here. However, the projections consistently suggest that streamflow will increase in March and April and decrease in August and September regardless of the GCM and emission scenario selected.

This study has made a number of pioneering contributions to the assessment of the hydrologic consequences of climate change in glacier-fed catchments. We would like to highlight the following:

1. It is the first study, to our knowledge, to use a physically based glacier dynamics model to make projections of the transient response of glaciers to future climate scenarios to provide boundary conditions for hydrologic modelling. Previous studies have used volume-area scaling (Stahl *et al.* 2008), adjustment of glacier thickness using a parameterized model (Huss *et al.* 2008), depletion of static glaciers with assumed thickness distributions (Rees and Collins 2006), adjustment of future glacier area based on a constant accumulation area ratio (Horton *et al.* 2006), or arbitrary assumptions about future decreases in glacier area for future time slices (e.g., Loukas *et al.* 2002; Hagg *et* al. 2006; Akhtar *et al.* 2008).
2. It is the first study to use changes in glacier volume and area to assist in hydrologic model calibration using a GLUE-type approach for a large catchment. This point is important given the lack of direct measurements of mass balance over most of the world's mountain regions, and especially in western Canada.
3. It is the first study in a glacier-fed catchment that has included a comprehensive assessment of uncertainties associated with uncertainties in hydrologic parameters, variation among emission scenarios and variation among GCMs.

Table of Contents

Acknowledgments

The work presented in this report has been supported by a grant from the Canadian Foundation for Climate and Atmospheric Science to the Western Canadian Cryospheric Network and by a contract from BC Hydro Generation Resource Management. The research builds upon previous research by the principal investigators that was funded by Discovery Grants from the Natural Sciences and Engineering Research Council. Tobias Bolch and Erik Schiefer contributed to the development of historic glacier masks and digital elevation models. Joel Trubilowicz performed GIS analysis to process the glaciological model output for use in the hydrologic model. Alex Jarosch contributed substantially to the development of the glaciological models. Sean Fleming, Frank Weber and Scott Weston of BC Hydro Generation Resource Management assisted with access to data and project administration. Sean Fleming and Frank Weber provided detailed reviews of earlier drafts of this report.

1 Introduction

This project was motivated by long-term planning needs of the Generation Resource Management group at BC Hydro, particularly in relation to the upcoming review of the current Columbia River Treaty and the associated need to understand how reservoir inflows may change in response to a range of scenarios for future climatic conditions. The Climate Impacts Group at the University of Washington is applying the Variable Infiltration Capacity (VIC) model to make these required projections. However, the VIC model does not contain a glacier routine, and there is, therefore, a need to apply an alternative hydrologic model. Glaciers in the Columbia basin are currently experiencing negative mass balance and, even if the current climate continued with no further warming, glaciers would likely continue to retreat for one or more decades. The overall objective of this study is, therefore, to develop and apply computational tools to generate projections of (a) future glacier coverage and (b) streamflow, including contributions from glacier melt, given a range of future climate scenarios. The study focused specifically on generating scenarios for inflows to the Mica Reservoir. Tasks involved in meeting the overall objectives are listed in Table 1.1.

Sean Fleming and Frank Weber supervised the contract on behalf of BC Hydro. R.D. Moore (UBC Geography) served as project manager. The work was conducted collaboratively by five groups:

1. UBC Geography (R.D. Moore, Georg Jost)
2. UBC Earth and Ocean Sciences (Garry Clarke, Faron Anslow, Valentina Radiç)
3. UNBC Geography (Brian Menounos, Roger Wheate)
4. Pacific Climate Impacts Consortium (PCIC) (Trevor Murdock, Arelia Werner)
5. Environment Canada, Pacific and Yukon Region (Alex Cannon)

The remainder of this report is organized into eight sections (2 to 9). Section 2 provides an overview of the modelling approaches. Section 3 provides details on the hydrologic model, including its calibration and testing, along with an interpretation of historic contributions of glacier runoff to inflow. Section 4 presents a description of the glacier dynamics model, including the approach to generating climatic forcing and the calibration and testing of the model. The generation of future climate scenarios is presented in section 5. Section 6 presents the simulated responses of glaciers to future climate scenarios. Section 7 presents the projections of streamflow under the various future climate scenarios, accounting for glacier response as simulated by the UBC-RGM. Section 8 provides discussion of some methodological issues addressed in this work and provides suggestions for future research. Section 9 summarizes the key findings.

Table 1.1. Summary of tasks and responsibilities associated with the project.

Task number	Description	Team members responsible for task
1	generation of future climate scenarios, involving selection of combinations of General Circulation Models and emission scenarios	Radiç, Werner
2	production of digital maps and digital elevation models (DEMs) for glacier coverage within the Mica basin	Menounos, Wheate
3	downscaling from GCM output to provide weather forcing for the hydrological model	Cannon, Murdock
4	downscaling from GCM output to provide forcing for the glacier dynamics model	Anslow, Radiç
5	running the glacier dynamics model to generate future glacier response	Anslow, Radiç, Clarke
6	calibrating the hydrological model using historic streamflow data and estimates of changes in glacier volume and extent	Jost, Moore
7	running the hydrological model using downscaled weather forcing and simulated changes in glacier coverage to generate future streamflow scenarios	Jost, Moore

2 Overview of modelling approach

This project employed the UBC Regional Glaciation Model (UBC-RGM) to make projections of future glacier cover. UBC-RGM has been recently developed by Garry Clarke and collaborators based in the Earth and Ocean Sciences Department at the University of British Columbia. The model is physically based and represents an increase in sophistication relative to approaches used in previous studies of glacier and streamflow response to future climate scenarios, which have included volume-area scaling (Stahl *et al.* 2008), adjustment of glacier thickness using a parameterized model (Huss *et al.* 2008), depletion of static glaciers with assumed thickness distributions (Rees and Collins 2006), adjustment of future glacier area based on a constant accumulation area ratio (Horton *et al.* 2006), or arbitrary assumptions about future decreases in glacier area for future time slices (e.g., Loukas *et al.* 2002; Hagg *et* al. 2006; Akhtar *et al.* 2008).

The HBV-EC model was used to translate the climatic scenarios and glacier responses into streamflow scenarios. Given the time frame in which the project had to be completed, it was not feasible to develop a new model and, thus, an existing model had to be used. HBV-EC was previously developed as a collaboration between scientists at Environment Canada Pacific and Yukon Region and the University of British Columbia. It was chosen for this study because it has a proven track record for modelling streamflow in British Columbia, particularly in glacier-fed catchments (Moore, 1993; Stahl et al., 2008; Cunderlik et al., 2010; Fleming et al., 2010). Both the glaciological and the hydrological models were calibrated and tested using historic weather and climate data as forcing.

The potential effects of future climate scenarios were simulated using output from Global Climate Models (GCMs) using a variety of emission scenarios from the IPCC Fourth Assessment Report (IPCC-AR4). Selection of candidate GCMs was based on their ability to reproduce climate, for both the western North America region and globally, based on air temperature, precipitation and synoptic-scale sea-level pressure patterns. Selection of emission scenarios was based on the availability of daily output, as this resolution is required to drive the hydrologic model. Daily output through the twenty-first century was available for CGCM, CGCMn and ECHAM. For all the other GCMs, daily output was only available for two time slices, 2050-2065 and 2085-2100.

Figure 2.1 provides an overview of the flow of information among the models, emphasizing the need to use different downscaling approaches for UBC-RGM and HBV-EC. UBC-RGM requires spatially distributed forcing at a typical resolution of 200 m, which was the resolution used in this study. Considering the large modelling domain involved in this project, it was not computationally feasible to generate new distributed (in space and time) weather forcings to compute spatially distributed mass balances for all of the future climate scenarios. As an alternative, a baseline set of distributed mass balance forcings was computed using historic information; this baseline mass balance field was then adjusted on the basis of projected air temperature and precipitation from the GCM output using a modified "delta" approach. Details are provided in section 4. For each future climate scenario, digital maps of glacier extents and elevations were output every ten years.

HBV-EC uses data from one or more weather stations, which are then interpolated/extrapolated over the catchment using routines built into the model, based on the definition of climate zones and vertical gradients for temperature and precipitation. Therefore, the downscaling approach had to generate daily weather data at the locations of the five weather stations used to calibrate and test HBV-EC in the Mica basin. Downscaling from GCM output to the weather station locations was conducted using the TreeGen routine developed by Alex Cannon (Environment Canada, Pacific and Yukon Region), which was used successfully in a similar application in the Bridge River catchment (Stahl *et al.* 2008). TreeGen combines elements of three downscaling approaches: synoptic weather typing, multiple regression and analogue re-sampling. Because TreeGen incorporates information on synoptic-scale conditions at a daily resolution, it has a key advantage over "delta" approaches for making hydrologic projections: it explicitly accommodates changes in the sequencing of weather systems, which can be critical controls on streamflow generation. The version of TreeGen used in this study ensures that trends in downscaled weather match trends in temperature and precipitation modelled by the GCMs, and thus ensures consistency with the climate forcing for the glaciological models.

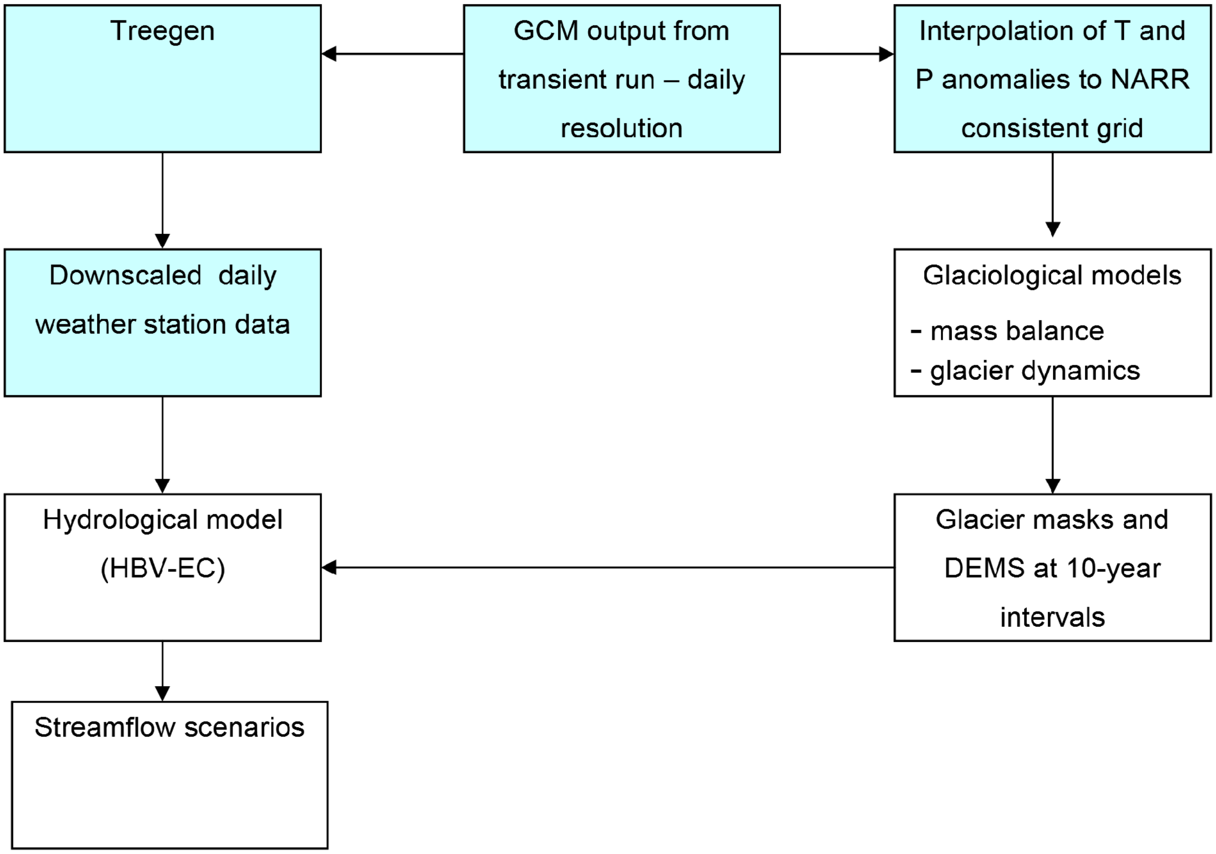


Figure 2.1. Overview of the flow of information amongst the models. Boxes shaded in blue refer to climate and weather data.

Like most hydrologic model codes, HBV-EC treats land cover as static. To accommodate changes in the glacier extents and elevations through time, HBV-EC was run through scripts that would stop the simulation every ten years, read in the new glacier extents and elevations, update the definitions of Grouped Response Units and state variables within the model, then continue the simulation, including a spin-up period.

3 Hydrologic modelling

3.1 The HBV-EC hydrologic model

The HBV model is a popular conceptual-parametric hydrologic model that strives to minimize computer run times and data requirements. The original HBV model was developed by the Swedish Meteorological and Hydrological Institute (Lindstrom et al. 1997). To minimize computational effort, HBV uses the concept of grouped response units (GRUs) to discretize a catchment into groups of grid cells with similar elevation, aspect, slope, and land cover. Variations of the HBV model have been developed by several groups around the world, and it has been demonstrated to provide relatively accurate simulations of streamflow in mountainous catchments, especially those with glacier cover (e.g., Braun and Aellen, 1990; Moore, 1993; Konz and Seibert 2010). The applicability of HBV-EC to modelling streamflow in British Columbia was demonstrated in an intercomparison study of process-oriented watershed models for operational river forecasting (Cunderlik et al., 2010; Fleming et al., 2010).

The HBV-EC model was originally developed in a partnership between Environment Canada and UBC (Hamilton *et al.* 2000). The model has since been integrated by the Canadian Hydraulics Centre into the EnSim Hydrologic modelling environment (now known as Green Kenue), which is a modelling environment and simple GIS system. Green Kenue facilitates the creation of model input files and visualization of inputs and outputs. The Green Kenue version of HBV-EC was used in this study.

HBV-EC differs in several ways from the traditional HBV model as described in Lindstrom *et al.* (1997). To better represent lateral climate gradients, HBV-EC allows for subdividing a basin into different climate zones, whereby each climate zone is associated with a single climate station and a unique parameter set. This feature can also be used to approximate the spatial variation of parameters. For example, HBV-EC can only model four land cover types (open, forest, glacier, lake) within one climate zone. Within a climate zone, the model does not account for variation caused by different forest types. However, by discretizing a basin into more than one climate zone, HBV-EC can be set up so that different forest types can be modelled. Unlike the traditional HBV model, HBV-EC can model snowmelt as a function of slope and aspect and thus can account for topographically forced spatial variations in melt rates. To predict the discharge for a given time step, HBV-EC does not use a weighting function to delay reservoir releases as the original HBV does (the *MAXBAS* parameter in HBV), but only sums output from each reservoir.

HBV-EC models glaciers as a separate land cover, using a different routing approach compared to non-glacier land cover. While HBV-EC does not compute glacier mass balance explicitly as a separate state variable, mass balance can be computed by post-processing model output. Glaciers in HBV-EC are treated as static in time; i.e., the glacier covered area cannot change during a model run. This constraint imposes a challenge for long term model simulations with changing glaciers. To overcome this constraint, Stahl *et al.* (2008) developed a version of HBV-EC coded in the IDL programming language (HBV-IDL) that includes a routine to use volume-area scaling to allow for glaciers to advance or retreat in response to positive or negative mass balance, respectively. In the current application, glacier response was simulated externally to the hydrological model using more sophisticated methods. Spatial masks representing glacier cover and surface elevation were input to the Green Kenue interface to update the GRUs at intervals throughout each future scenario.

Another difference between HBV-EC and the original HBV model is that the temperature-index-based snow melt algorithm was adapted by Hamilton *et al.* (2000) to account for the effects of slope, s, aspect, a, and forest cover. In this modified version, snowmelt (M) is calculated by:

$$M(t) = C_0(t) \times MRF \times (1 - AM \times \sin(s) \times \cos(a)) \times T_{air} \qquad (3.1)$$

where C_0 is a base melt factor that varies sinusoidally between a minimum value ($CMIN$) at the winter solstice to a maximum value at the summer solstice ($CMIN + DC$) to account for seasonal variations in solar radiation, DC is the increase in melt factor between winter and summer solstices, s is the slope of a GRU, a is the aspect of a GRU, and AM is a factor controlling the influence of aspect. T_{air} is the air temperature above a threshold temperature for snowmelt. The melt reduction factor MRF, which ranges between 0 and 1, reduces melt rates under forests compared to melt at open sites.The melt factor for glacier GRUs is multiplied by the coefficient MRG, which typically ranges between 1 and 2, once seasonal snowpack has ablated to reflect a reduction in surface albedo associated with a bare ice cover.

3.2 Spatial discretization of the Mica basin

The Green Kenue platform was used to discretize Mica basin and create model input files. All calculations were based on a DEM with 200 m pixel size, which is the same resolution as used in the glacier modelling. The basin boundaries were delineated by the A^t algorithm in Green Kenue. This algorithm gave a closer match to the basin delineation of BC Hydro than the alternative depressionless flow algorithm. The total drainage area of the Mica basin is 20,742 km^2.

Data from four climate stations within Mica basin and one additional climate station just outside the catchment boundary were available for modelling (Figure 3.1). Mica dam climate station (MCA) has the longest climate record, dating back to 1965. Rogers Pass climate station (RGR) data start in 1967, Radium climate station (RAD) in 1969, Molson Creek climate station (MOL) in 1986, and Floe Lake climate station in 1993. Records for MCA and RGR were backfilled to 1960 by BC Hydro. To extend the records for MOL, FLK and RAD back to 1960, we computed factors for each three-month quarter to rescale precipitation data from MCA. Air temperature was rescaled based on linear regressions for each quarter. Backfilled climate data were used to calculate historical changes in streamflow. For model calibration and testing all climate data were measured, except for 8 years of backfilled data from FLK (1985-1993).

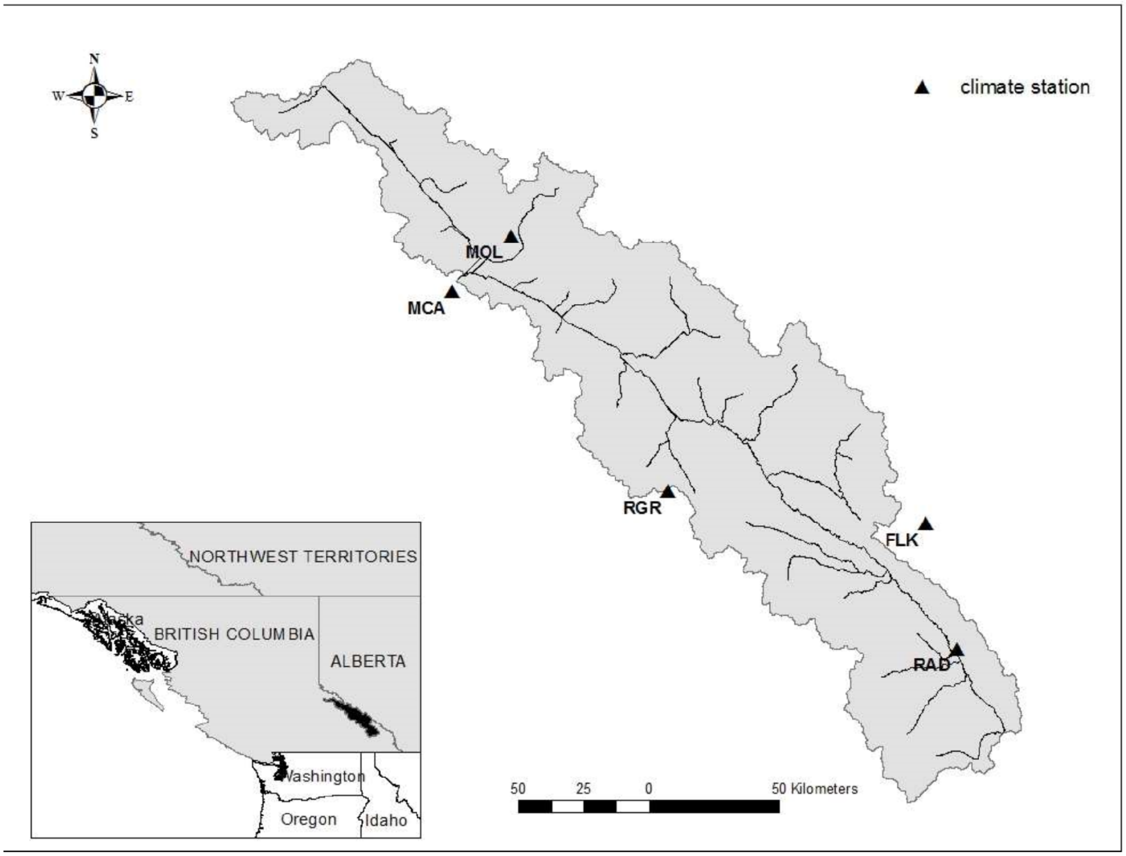


Figure 3.1. The Mica basin showing locations of climate stations used to force the hydrological model.

Mica basin was subdivided into six climate zones, two at low elevation and four at high elevation. This discretization accounted for lateral climate gradients, and also minimized the potential for errors in extrapolating air temperature from valley bottom stations to higher sites under inversion conditions. A smoothed 1200 m contour line was used to separate low from high elevation climate zones. Isolated polygons were erased with a modal filter. Each climate zone was forced by a separate climate station, except for zones 5 and 6, which were both forced by RGR (Figure 3.2).

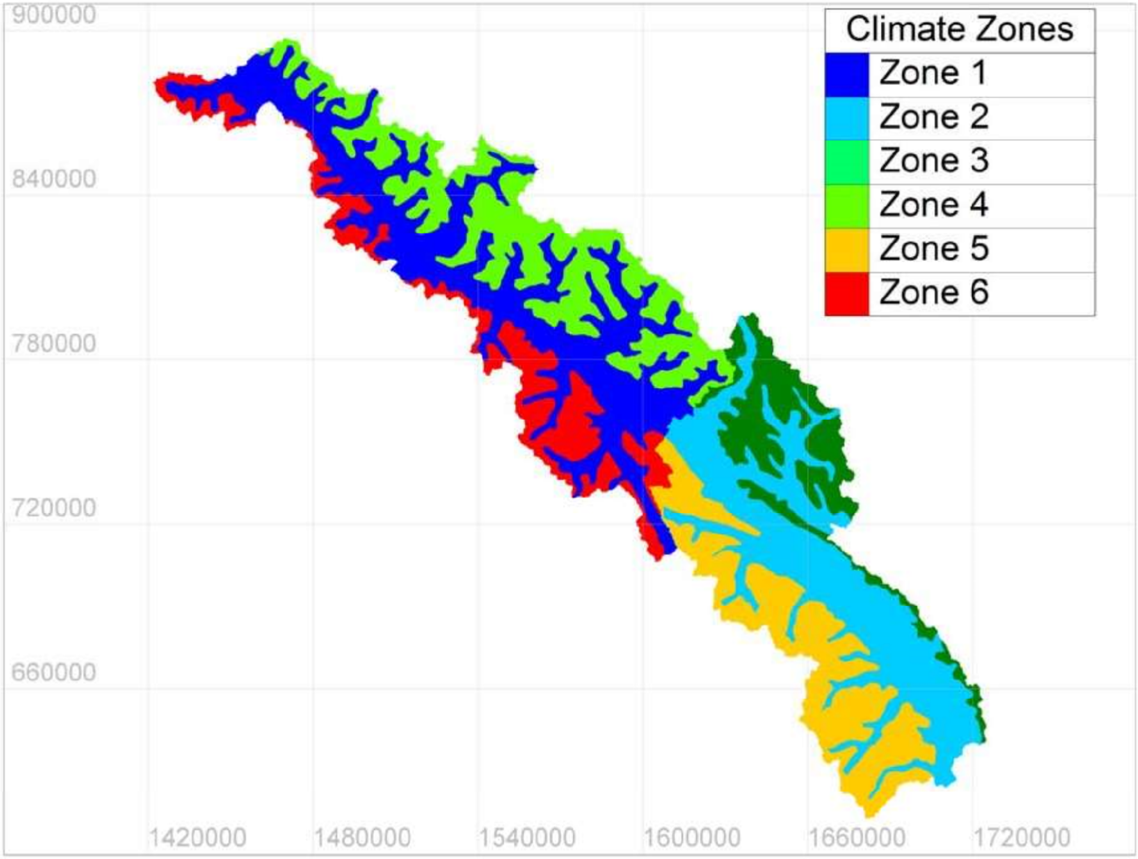


Figure 3.2. Delineation of climate zones used for hydrological modelling.

There is a trade-off between spatial resolution and model run-time. We tested the model performance using various combinations of elevation bands, aspect classes and slope classes. The highest spatial resolution (ten elevation bands, three slope classes and four aspect classes) and the lowest spatial resolution (four elevation bands, two slope and two aspect classes) generated Nash-Sutcliffe efficiencies within 0.01 of each other. Since the highest spatial resolution required run-times that were too long to perform the high number of model runs that are required for uncertainty analysis in a reasonable time, we decided to keep a high resolution for elevation (ten bands) (Figure 3.3), but reduced the slope and aspect resolution to two classes each. A high resolution for elevation is necessary for accurate modelling of the potentially shrinking accumulation zones of glaciers. For example, using a low number of elevation bands could result in a complete loss of accumulation zones in the model when, in reality, accumulation zones would still exist.

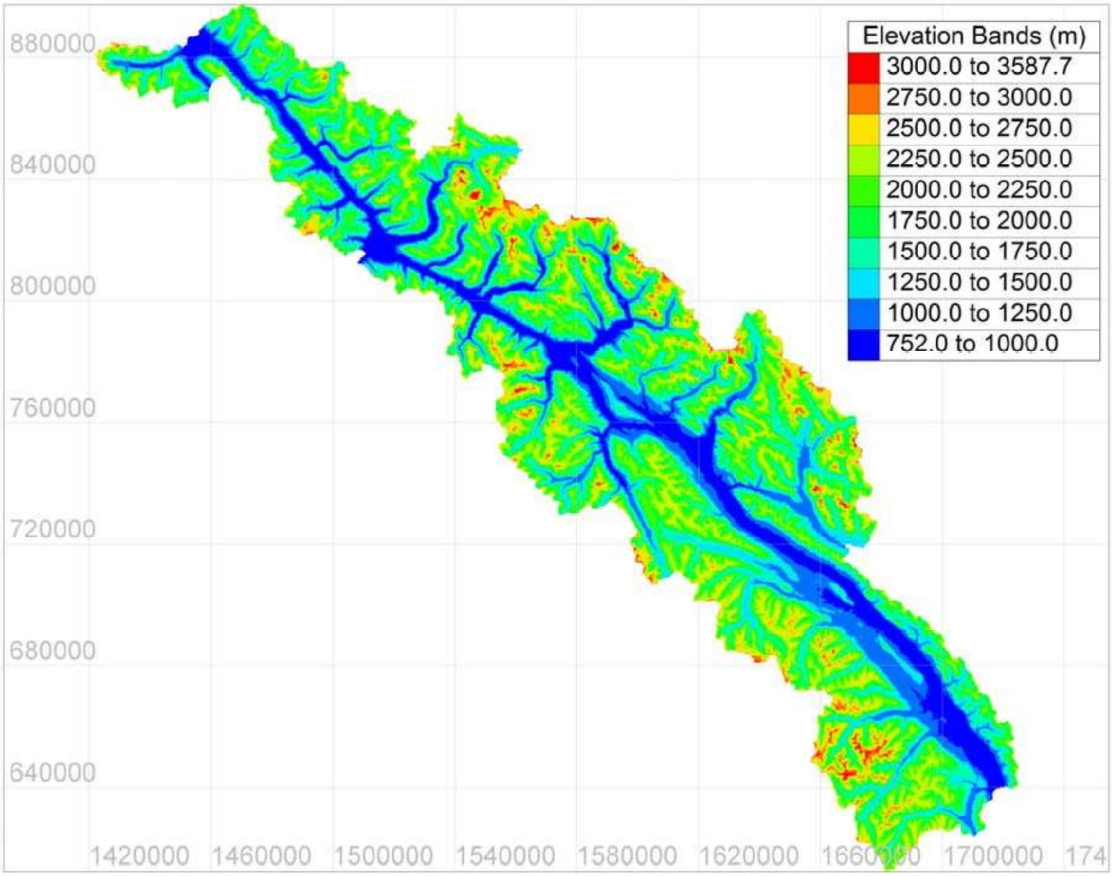


Figure 3.3. Elevation bands used to define grouped response units for hydrological modelling.

In 1985, glaciers covered 1268 km^2 in the Mica basin, representing 6.1% of the total basin area (Table 3.1, Figure 3.4). Between 1985 and 2000 the glacier area decreased by 101 km^2. Between 2000 and 2005, an additional 80 km^2 of glacier area was lost, reducing glacier cover to 5.2% of the basin area. Compared to other studies on the influence of glacier changes on hydrology, the glacierized area at Mica basin is relatively small. For example, Bridge Glacier currently covers about 60% of the catchment area of the Bridge River at the Water Survey of Canada gauge (Stahl et al., 2008). However, the results of this study are likely to be broadly representative of the practical downstream water resource impacts of glacier recession in large basins.

Table 3.1. Historic variations in glacier area in Mica basin.

Year	Glacier area (km^2)	% change relative to 1985	% of Mica basin
1985	1268.8		6.1
2000	1168.0	-7.9	5.6
2005	1088.6	-14.2	5.2

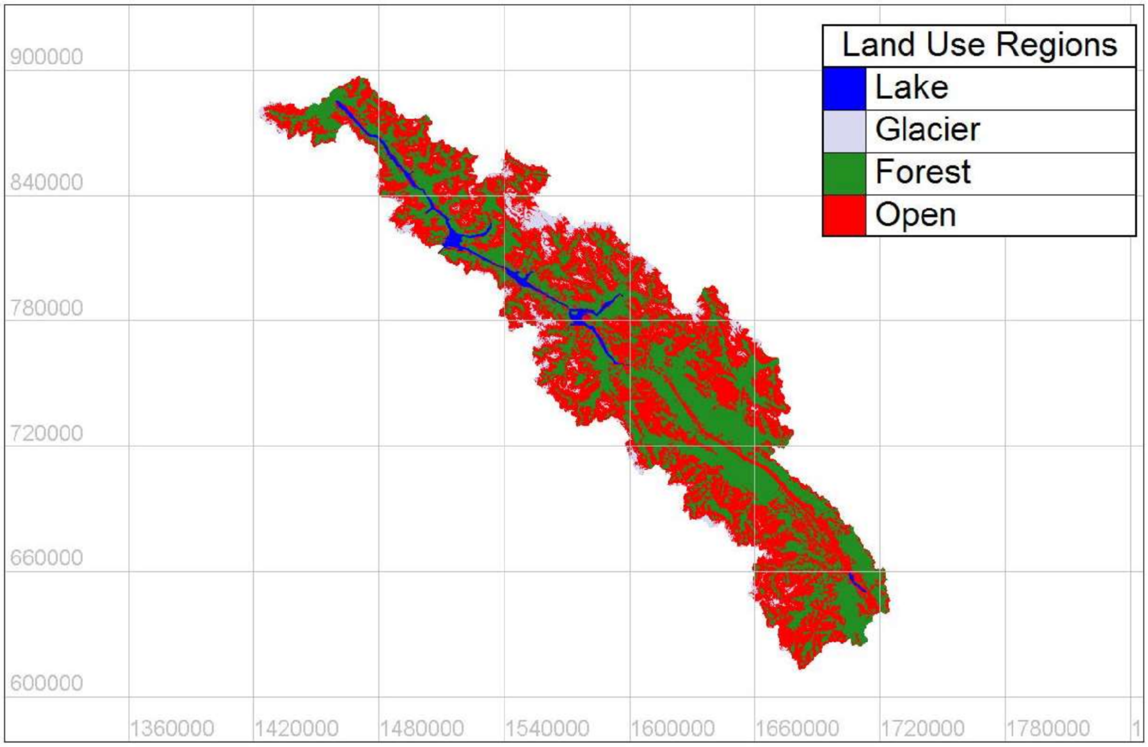


Figure 3.4. Land cover in Mica basin with glacier extents from 1985.

3.3 Model spin-up

Because the initial values of model state variables such as soil moisture are not known, it is common to include a spin-up period in a simulation run to minimize the effect of incorrect initial conditions on simulated streamflow. Model output from the spin-up period is not used in the calculation of model performance statistics used for calibration and testing.

We investigated how long spin-up times need to be in order for the slow reservoir storage in HBV-EC, which we believe is the most "sluggish" state variable, to equilibrate with the forcing data. With an initial slow reservoir discharge of 100 m^3/s, it takes three winter seasons for the slow reservoir storage to equilibrate. It seems intuitive that longer spin-up times should produce more robust results for a given time slice because state variables are more likely not to be influenced by erroneous initial conditions. However, in glacierized catchments, longer spin-up times can be problematic for the following reason: because HBV-EC does not convert snow into ice over time, long-term simulations result in the accumulation of deep snowpack above the ELA rather than the accumulation of firn and ice. As a result, glacier areas, which in reality have ice exposed during the ablation season, will instead be snow covered in the model. Melt from those areas will be under-simulated, as snowmelt rates are typically lower than icemelt rates. This effect increases as the value of the parameter *MRG* increases. This effect is reinforced by the fact that HBV-EC does not transport snow downslope where it would melt at a higher rate. This problem is common to all hydrological model codes that do not explicitly include glacier dynamics.

It is also important to adjust the parameter that sets the maximum amount of liquid water that can be retained by a snowpack, *LWR*. In the Green Kenue version of HBV-EC, the default value of *LWR* is 2500 mm, which is unrealistically high and can strongly influence the dependence of model results on spin-up time because the snow that accumulates above the ELA will retain most of its melt water, thus reducing simulated runoff from those elevation bands. In this work, *LWR* was set to 200 mm, which is equivalent to the water retention capacity of a snowpack with a water equivalence of 4,000 to 10,000 mm (based on a snowpack being able to hold 2 to 5% liquid water by mass).

In the context of the points made above, we set the spin-up period for all model runs (i.e., for each time slice when performing transient model runs, see hereafter) to five years. This spin-up time is sufficient to ensure that the slow reservoir storage equilibrates with the forcing data but also minimizes the problems associated with excess snow accumulation and water retention capacity.

3.4 Model calibration and uncertainty analysis

3.4.1 A guided GLUE approach

Model calibration and uncertainty analysis were based on a multi-criteria and multi-step procedure that can best be described as a "guided GLUE" approach. The approach combined a Generalized Likelihood Uncertainty Estimation (GLUE) type procedure (Beven and Freer 2001; Beven and Binley 1992) with an evolutionary optimization algorithm. GLUE type approaches are based on the philosophy that there is no single "optimal" parameter set for a given model and data set. Rather, these approaches identify multiple parameter sets that provide acceptable simulation accuracy for the calibration data. These parameter sets are termed "behavioural." When applying the model to make predictions outside the calibration data set, all parameter sets are used, resulting in an ensemble of predictions that represents the effect of uncertainty in the parameters.

Previous studies have demonstrated the benefits of using glacier mass balance data to constrain hydrological model calibration in glacier-fed catchments (e.g., Konz and Seibert 2010; Stahl *et al.* 2008; Schaefli *et al.* 2005). However, most of those studies relied on direct mass balance measurements, or reconstructions based on direct measurements. Direct measurements of glacier mass balance are not available for Mica basin. However, even if mass balance data were available, it is unclear how to extrapolate from a single glacier, or even a small number of glaciers, to the entire Mica basin, given the diversity of glacier morphologies (including elevation ranges, slope and aspect) and climatic conditions. An alternative approach is to use information on glacier volume change throughout the basin. For Mica basin, Schiefer *et al.* (2007) estimated the glacier volume loss from 1985-1999 to be 7.75 km^3. Volume loss was calculated from the 1999 Shuttle Radar Topography Mission (SRTM) data and from digital terrain models obtained from aerial photographs. The photographs were taken between 1982 and 1988 with a median weighted date for Mica basin of 1985. Observed glacier areas for Mica basin were available for 1985, 2000, and 2005.

In order to constrain model parameters with this estimate of glacier volume loss, the model was calibrated for the period 1985-1999. The period 2000-2007 was used as an independent test period, using glacier cover based on data from 2005. Model calibration runs were split into two steps, each with a five-year