Evaluation of Long Range Summer Forecasts of Lower Fraser River Discharge and Temperature Conditions

D.A. Patterson and M.J. Hague

Fisheries and Oceans Canada
Science Branch, Pacific Region
Co-operative Resource Management Institute
School of Resource and Environmental Management
Simon Fraser University
Burnaby, BC

2007

Canadian Technical Report of
Fisheries and Aquatic Sciences 2754

Fisheries and Oceans Canada | Pêches et Océans Canada

Canada

Canadian Technical Report of
Fisheries and Aquatic Sciences 2754

2007

EVALUATION OF LONG RANGE SUMMER FORECASTS OF LOWER FRASER RIVER DISCHARGE AND TEMPERATURE CONDITIONS

by

D.A. Patterson and M.J. Hague

Fisheries and Oceans Canada
Science Branch, Pacific Region
CRMI c/o School of Resource and Environmental Management
Simon Fraser University
Burnaby BC
V5A 1S6

Cat. No. Fs 97-6/2754E ISSN 0706-6457

Correct citation for this publication:

Patterson, D.A. and Hague, M.J. 2007. Evaluation of long range summer forecasts of lower Fraser River discharge and temperature conditions. Can. Tech. Rep. Fish. Aquat. Sci. 2754: vii + 34 p.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Patterson, D.A. and Hague, M.J. 2007. Evaluation of long range summer forecasts of lower Fraser River discharge and temperature conditions. Can. Tech. Rep. Fish. Aquat. Sci. 2754: vii + 34 p.

Extreme temperature and discharge conditions in the Fraser River adversely affect adult sockeye salmon migration success. Current fisheries management practices adjust harvest plans based on predicted summer temperature and discharge values. Therefore, the development of long range (~1 to 4 months) forecasts of environmental conditions will aid in pre-season harvest planning. This report evaluated several models used to make long range forecasts of summer conditions. Fraser River sockeye salmon (*Oncorhynchus nerka*) run-timing groups were used as a case study to illustrate model performance. Most models were best-fit using simple, or multiple, linear regressions. We quantified the uncertainty in the temperature and discharge forecasts arising from uncertainty in the model structure, and, where applicable, uncertainty in the predictor variable. Predictor variables include winter precipitation anomalies, spring and summer air temperature anomalies, water volume forecasts, and historic trends in water and air temperatures. Temperature forecast models performed comparably, and consistently predicted summer river temperatures with a coefficient of variation of less than 8% and an approximate standard deviation of 1°C. The precision of discharge forecasts deteriorated throughout the summer, and there was a trade-off between the availability of the forecast method and the precision of the results. In general, increasing the number of days used to calculate the predicted means led to small improvements in model fit, however there was only modest improvement going from 19-day to 31-day means. Recommendations of the most appropriate models were made based on model fit, forecast uncertainty and the timing of data availability.

RESUMÉ

Patterson, D.A. and Hague, M.J. 2007. Evaluation of long range summer forecasts of lower Fraser River discharge and temperature conditions. Can. Tech. Rep. Fish. Aquat. Sci. 2754: vii + 34 p.

Les températures extrêmes et les conditions d’écoulement dans le Fraser sont réputées pour diminuer la réussite de la migration des saumons rouges adultes. Les stratégies actuelles de gestion des pêcheries ajustent les prévisions de pêche aux valeurs prédites concernant les températures estivales et le flux dans le Fraser. Ainsi, le développement de prévisions à long terme (1 à 4 mois) des conditions environnementales va aider à mettre en place des programmes de pêche pour l’avant saison. Ce rapport a évalué plusieurs modèles utilisés pour faire des prévisions à long terme des conditions estivales. L’étude du cas des saumons rouges du Fraser (*Oncorhynchus nerka)*, migrant à différentes époques selon le groupe, a servi à illustrer les performances des modèles. La plupart des modèles étaient le plus en adéquation avec des régressions linéaires simples ou multiples. Nous avons quantifié les incertitudes concernant les prévisions des températures et des écoulements résultant des incertitudes de la structure du modèle, et quand c’était possible, des incertitudes des variables indice. Les variables indice incluent les anomalies des précipitations hivernales, les anomalies de température de l’air du printemps et de l’été, les prédictions des volumes d’eau et les évolutions historiques de la température de l’air et du fleuve. Les modèles de prévision de la température estivale du fleuve ont donné des résultats très proches et tous avec un coefficient de variation de moins de 8% et un écart moyen d’environ 1 °C. La précision des prévisions de l’écoulement a diminué au cours de l’été, et il y eut un compromis entre la commodité de la méthode de prévision et la précision des résultats. En général, l’augmentation du nombre de jours sur lesquels la moyenne est calculée a mené à des améliorations légères de la correspondance du modèle aux résultats, cependant il n’y eut que de modestes améliorations en passant de moyennes de 19 à 31 jours. Les conseils prodigués pour choisir les modèles les plus appropriés se sont basés sur la correspondance du modèle, l’incertitude des prévisions et la période des données disponibles.

INTRODUCTION

Environmental conditions experienced by upstream migrating salmon have a direct influence on their survival (e.g. Quinn et al. 1997; Naughton et al. 2005). In the Fraser River, high en route mortality is correlated to severe environmental conditions experienced by migrating adult sockeye salmon (*Oncorhynchus nerka*) (Macdonald 2000; Macdonald et al. 2000; Patterson et al. 2007a). High temperatures influence sockeye salmon migratory success by impairing swimming ability (Salinger and Anderson 1996; Lee et al. 2003; Naughton et al. 2005) through increasing energy expenditures (Brett 1995), stress (Fagerlund et al. 1995) and susceptibility to disease (Macdonald et al. 2000; Wagner et al. 2005). There is also a negative correlation between high discharge and up-river migration success (Rand and Hinch 1998; Macdonald et al. 2000). High discharge values are associated with high river velocities and therefore slower migration rates (Quinn et al. 1997) and higher energy expenditures (Rand and Hinch 1998). In the Fraser River, velocity barriers are also present near Hells Gate at discharge values in excess of 8000 $m^3 \bullet s^{-1}$ (Macdonald 2000). Given the relationships between salmon survival and environmental conditions, forecasts of Fraser River summer conditions are of particular interest to fisheries managers.

Currently, the Fisheries and Oceans (DFO) Environmental Watch Program provides long range (i.e. 1 – 4 months in advance) forecasts of environmental conditions to fisheries managers for use in their pre-season planning process (D. Patterson, DFO, pers. comm. 2006). Specifically, managers use pre-season (D. Patterson, DFO, pers. comm. 2006), and in-season (Morrison 2005), Fraser River temperature and discharge forecasts to generate estimates of the expected differences between lower river escapements estimated at Mission and spawning ground escapements (after accounting for in-river catch). The current approach uses a Difference Between Estimates (DBE), also known as Management Adjustment (MA), model that fits a non-linear simple or multiple regression relationship between historic temperature and/or discharge data, and differences between potential and actual spawning ground escapement estimates. In general, Early Stuart, Early Summer and Summer run-timing groups use the same MA model framework (I. Guthrie, Pacific Salmon Commission, Vancouver, BC, pers. comm. 2006). Pre-season forecasts of summer river conditions provide an early indication to managers of the expected discrepancy between potential (Mission) and observed spawning abundance; however, the reliability of the DBE or MA estimates are limited by the uncertainty inherent in making environmental predictions derived from long range forecasts (e.g. Moore 2006). Providing fisheries managers with quantified estimates of the reliability of environmental forecasts used in the management adjustment models will facilitate a more informed decision-making process.

The purpose of this technical report is to evaluate the use of different environmental variables to generate long range forecast models for predicting Fraser River summer water temperature and discharge. Diagnostic assessments are used to identify the appropriate model structure for each relationship. Next, uncertainty associated with both the environmental input variable and forecast model structure is quantified. Finally, model sensitivity is explored with respect to seasonality of forecast dates, and time frames used to calculate mean temperature and discharge values. Specifically, this report evaluates eight long range forecasting methods:1) historic river temperature trends, 2) winter precipitation index, 3) snowpack water volume forecasts, 4) an ensemble flow model, 5) temperature – discharge correlations, 6) forecasted summer air temperature anomalies, 7) measured spring air temperature anomalies, and 8) various multiple regression models which combine two of the above methods. The following sections describe model structure, linear regression diagnostics, and methods used to quantify model uncertainty. The report concludes with a retrospective analysis of 2005 forecasts and provides recommendations regarding forecasting methods and future research directions.

METHODS

MODEL STRUCTURE

All models and simulations were performed using the freeware statistical analysis package R (http://cran.r-project.org/). Model flexibility was emphasised, allowing investigators to vary several input sources (e.g. predictor and response variables, median run-timing date, number of days used to calculate mean temperature and discharge), and facilitating sensitivity analyses and data updates.

Fraser River sockeye salmon run-timing groups were used as case specific examples to illustrate the predictive capacity of forecast models over the course of the summer season. The run-timing dates used in the models represent the expected mid-point of temperature/discharge exposure for each salmon run group at a specific station in the lower Fraser River. Specific values for each run-timing group (Early Stuart, Early Summer, and Summer) reflect the expected median passage date for Hells Gate (Table 1). Hells Gate dates were computed by adding five days to the median date calculated from 1977-2005 Mission run-timing information (collected by the Pacific Salmon Commission; Woodey 1987). Temperature and discharge data were calculated as 19-day symmetric means (9-days before and 9-days after historic Hells Gate dates). A 19-day period was selected to represent the average lower-river environmental conditions experienced by the incoming run, and is the length of the period currently used for in-season DBE forecasts.

Table 1. Median run timing dates for Hells Gate.

Run Timing Group	Date
Early Stuart	July 14
Early Summer	August 6
Summer	August 17

Historic temperature and discharge values represent a 1950-2004 time series. Temperature data recorded at Hells Gate was collected over time by the International Pacific Salmon Fisheries Commission, the Pacific Salmon Commission, and most recently by the DFO Environmental Watch Program (Patterson et al. 2007b). Discharge data recorded at Hope was extracted from the online Environment Canada Water Survey of Canada database (http://www.wsc.ec.gc.ca/).

DIAGNOSTIC TESTS

The raw data structure was explored to determine whether the assumptions for linear regressions between environmental forecast predictors and mean temperature and discharge were met. If assumptions were not met then additional models were applied to determine whether the data was better fit using an alternative model structure. A summary of diagnostic tests and corrective procedures are presented in Table 2.

If the data indicated non-linearities, the fit of the linear regression was compared to the fit of polynomial models, using Akaike's Information Criterion (AIC) (Maindonald and Braun 2003), and log-linear models, through a chi-squared goodness of fit procedure (Zar 1996). If the non-linear models did not provide a significantly better fit to the data, and the linear regression was still significant, then a linear regression was performed. In cases of autocorrelations, the program fit a moving average model to the data and then tested whether the moving average model provided a statistically better fit than the linear regression (Venables and Ripley 2002; Maindonald and Braun 2003). If the fits were not significantly different, then the program defaulted to a simple linear regression. If the data showed evidence of heteroscedasticity a log transformation or a weighted least squares analysis (Zar 1996) was performed.

Table 2. Diagnostic tools for linear regression assumptions and potential correction techniques.

Model Assumption	Evidence	Correction
Linearity – the data is best fit assuming a linear relationship	Plot residual vs. predicted values. A bowed pattern indicates non-linearity.	A linear model may not be appropriate for this data, and a non-linear model (e.g. polynomial regression) may be more appropriate. Alternatively, try transforming the data (e.g. log transform).
Independence of Errors – values of 'y' are independent of each other	Plot an autocorrelation function of the residuals. If autocorrelations fall outside the 95% confidence limits, there maybe autocorrelation (often occurs in time series data). Test using a Ljung-Box test.	Apply a moving average regression model.
Homoscedasticity – equal variances around each 'y' value	Plot residuals vs. predicted values. If residuals show increasing spread over predicted values there is evidence of heteroscedasticity.	Apply a weighted least squares regression or a log transformation.
Normality – errors come from a normal distribution	Create a normal probability plot of residuals; distribution should be the same as for a random normal distribution.	Apply a general linear model; try transforming for non-linearity; remove unexplainable outliers.
Errors in Variables - observations of x are obtained without error.	Fails for several provided datasets where predictions are based on forecasted values.	Apply a mean functional regression or bootstrap technique.

FORECASTING METHODS

The following section provides a description of each forecast method, the nature of the data used for each method and the procedures used to quantify the uncertainty in the temperature or discharge prediction. The statistical techniques utilised to estimate forecast uncertainty for each method are described. Forecasted or measured environmental variables for 2005 were used as examples in all cases.

Temperature forecast models

Historic temperature trends

Water temperatures in the lower Fraser have been increasing since the first continuous records were established in 1953 (Foreman et al. 2001; Patterson et al. 2007b). Therefore, a linear regression fit between years and historic 19-day mean river temperature was evaluated to determine if the historic temporal trend could be used to make predictions of water temperature in the following year for each run timing group. Uncertainty in the temperature forecast was quantified using simple linear regression prediction intervals and residual standard errors (standard deviation of the error distribution).

Temperature-discharge correlation

Past observations have often noted negative correlations between river temperature and discharge (Quinn et al. 1997; Macdonald et al. 2000; Naughton 2005; Moore 2006). Predicting temperature, and its associated error, directly from forecasted discharge may eliminate the need to forecast temperature from other data sources, such as the historic trend or spring and summer air anomalies. As the current discharge forecast will be uncertain, this estimation procedure violates the linear regression assumption of negligible error in the predictor variable. Therefore, the estimation of temperature uncertainty required quantification of both model and predictor variable uncertainty.

Uncertainty in forecasted discharge was modelled as a normal distribution with error equivalent to the standard deviation derived from the applied discharge forecasting method (see Discharge forecast models). Temperature was then predicted from each of 500 bootstrapped discharges using a simple linear regression model fit between 19-day mean discharge and 19-day mean temperature. Model uncertainty was incorporated into the final prediction by performing a non-parametric bootstrap of model residuals (Chernick 1999). A bootstrapped residual error term was added to each predicted temperature value. Eighty-percent confidence intervals were calculated using the 10th and 90th percentiles of the bootstrapped data. The generation of bootstrapped prediction intervals involves a more complex statistical analysis (Stine 1985), which was not completed here. Error structure was also presented as the standard deviation of the bootstrapped predictions.

Summer air anomaly

On June 1, Environment Canada (EC) posts predictions of summer air temperature, and reports this data in the form of an air temperature anomaly, varying from the mean, or “normal”, temperature averaged over 1951-1980 (http://weather.ec.gc.ca/saisons/index_e.html).

Environment Canada generates twelve anomaly values using two different forecasting models. Each forecast is specified as “near normal”, “below normal”, or “above normal”, and the 12 model results are divided into each category. The category thresholds are spaced equidistant apart, at intervals equal to 0.43 times the inter-annual seasonal temperature standard deviation. However, these forecasts are highly uncertain, and the probability that the true forecast falls within a given range is un-calibrated (not compared to true historic trends) due to a lack of available data (http://weather.ec.gc.ca/saisons/image_e.html?img=pc_dyn_jja1_temp).

Previous attempts to utilise the summer air temperature anomaly forecast include selecting a reasonable percentile from the historic anomalies given the probabilistic forecasts, or creating a categorical river temperature prediction matrix given the relationship between river temperature and predicted flow and air temperature ("low", "normal", "high") (J. Morrison, VYNX design, Sidney, unpub. data). This report evaluates two new approaches for quantifying air temperature forecast uncertainty using bootstrap methods. The first method attempts to mimic the actual forecasting process as outlined on the EC website. The second method predicts forecast error using the historic summer air anomaly trend.

To mimic the forecasting process described by Environment Canada, 500 counts of "near normal", "below normal", and "above normal" forecasts were generated using a multinomial distribution and 2005 forecast probabilities. The probability that the true anomaly will fall correctly into each category was determined from un-calibrated EC forecast probability maps, and then roughly adjusted using forecast reliability plots provided on the EC website. For example, if the probability that the true anomaly is "above normal" is 20-30% but the reliability data indicates that the *observed* frequencies of "above normal" anomalies are underestimated when the *forecasted* probability is <30%, then the new forecast was roughly calibrated as the upper bound of the prediction interval (30%). As the reliability maps use data averaged over all of Canada, the calibration for a specific region is approximate.

The count data generated from the multinomial simulations was converted into temperature anomalies by non-parametrically bootstrapping the historic anomaly values occurring within the range specified by each category. Predicted water temperatures were produced for each bootstrapped air anomaly by fitting a simple linear regression to historic summer anomaly and river temperature data. Model uncertainty was incorporated into each prediction by adding a bootstrapped model residual term to the predicted water temperature value and calculating the standard deviation of the bootstrapped values, as previously described.

The second method used a linear relationship fit between years and historic summer air anomalies (http://www.msc-smc.ec.gc.ca/ccrm/bulletin/archive_e.cfm). The mean and residual standard error for the current year's anomaly was predicted from the historic trend, as opposed to using the Environment Canada forecast. A parametric bootstrap assuming a normal distribution (~N (mean anomaly, residual standard error)) was used to generate 500 anomaly values. As above, predicted water temperatures were then predicted for each bootstrapped air anomaly.

Spring air anomaly

Spring air anomalies are available in early June from Environment Canada (http://www.msc-smc.ec.gc.ca/ccrm/bulletin/national_e.cfm). Because the spring anomalies are directly measured, the analysis is much simpler than for the summer air forecasts. Prediction intervals were directly computed from the linear regression between spring air anomalies and water temperature.

Multiple regression

Because river temperature is correlated to several external factors, such as discharge, air temperature and climate trends, we hypothesised that a multiple regression model may explain a greater percentage of the observed variance in historic river temperature than one of the previously described single variable regressions. Several combinations of predictor variables were explored for the temperature multiple regression analysis, including the historic temperature trend and snowpack water volume, historic trend and ensemble flow data, and summer air anomaly and snowpack water volume. Performance measures were altered slightly for the multiple regression models, and adjusted r^2 values were used to evaluate model fit while adjusting for additional predictor variables (Zar 1996). A step-wise AIC procedure was used to evaluate the multiple regression models (Maindonald and Braun 2003).

Discharge forecasts

Winter precipitation index

The winter precipitation index represents the combined precipitation fallen over December, January and February (http://www.msc-smc.ec.gc.ca/ccrm/bulletin/national_e.cfm). The index is presented as a percent anomaly from the mean "normal" precipitation averaged over 1951-1980 (Meteorological Service of Canada, MSC) for the BC South Mountain region. The percent anomaly was regressed against the 19-day historic discharges during the summer period. Because observed decreases in historic winter precipitation were unparalleled by similar decreases in mean summer discharge, observed autocorrelations were calculated for the de-trended time series (Venables and Ripley 2002; Maindonald and Braun 2003). Forecast uncertainty was quantified using simple linear regression prediction intervals and residual standard errors.

Snowpack water volume forecast

The River Forecast Centre (RFC) (BC provincial government) uses hydrological models and statistical regression to provide annual basin-specific forecasts of water volume due to snowfall (http://www.env.gov.bc.ca/rfc/river_forecast/interpret.htm). Forecasts represent the expected volume of snowmelt water to pass by the Fraser River at Hope from the date of the forecast to September 30th, assuming average weather conditions during this time. Forecasts are typically available for April 1st and May 1st. The historic total volume of water to pass Hope was calculated using daily discharge values for these two specified time periods. The summer volume forecast was then regressed against the mean 19-day discharge or temperature for each of the run timing groups.

The use of snowpack water volume forecasts of river discharge violates two linear model assumptions. The first issue arises because the calculation of the historic water volume is partially dependent on values of water discharge. Therefore, a cyclical relationship exists between the two variables, violating the linear model assumption of independence. As such, the model likely underestimates the true uncertainty in predicted discharge. The second violation arises because the current year's predictor variable (volume) has non-negligible error. Volume forecast uncertainty was modelled using a parametric bootstrap, assuming a normal error distribution. The standard deviation of the normal distribution was calculated as the standard error for the 2005 forecast, estimated from the 80% snowpack water volume confidence intervals (http://www.env.gov.bc.ca/rfc/river_forecast/forecast_apr06.htm). Discharge or temperature forecasts and associated error structures were then calculated using the bootstrapping methods previously described (e.g. see Temperature-discharge correlations)

As described above, the RFC typically provides the latest snowpack water volume forecast on May 1st. However, June 1st forecasts can be generated by subtracting the total May Hope water volume (volume = 0.0864*discharge) from the May 1st forecast. The May 80% confidence intervals are then adjusted proportionally. This method assumes that precipitation that falls in May, both snowfall accumulation at high elevation and rainfall, is normal. In other words, the majority of May discharge results directly from snowmelt runoff and not from large deviations in May precipitation. Once the updated forecast was calculated, discharge and temperature values were predicted as described above.

Ensemble flow technique

The RFC also produces simulations of future daily river discharge forecasts using a combination of current snowpack conditions, the previous 50 years of historic weather data and forecasted meteorological conditions for the current year (A. Chapman, RFC, Victoria, BC, pers. comm. 2005). In previous years, the first forecast was produced in April, and then the ensemble procedure was repeated with updated information roughly every two weeks until June. At this time, the RFC is undergoing revisions to their current ensemble model; therefore, future ensemble predictions may produce different results to the models analysed in this report. Uncertainty was incorporated into daily estimates by calculating the mean and standard deviation of the discharge for a given day over all 50 years of simulated data. The 19-day mean discharge for the current year was calculated by averaging mean daily discharge values over all years. The standard deviation of the current discharge was estimated from the standard deviation over all annual 19-day means.

SENSITIVITY ANALYSES

Long range environmental forecast model performance depends on a number of factors in addition to those incorporated into the current uncertainty analysis. Therefore, the historic trend – snowpack water volume multiple regression and the ensemble flow models were used as case studies to evaluate the robustness of temperature and discharge predictions, respectively, to changes in the number of days used to calculate the environmental averages and the assumed Hells Gate 50% date for each run.

In the uncertainty analyses, we assumed 19 days adequately captures the period of environmental exposure at Hells Gate for the incoming run. However, there is substantial among-group and inter-annual variability in the number of consecutive days a run enters the river (M. Hague, DFO, unpub. data). The magnitude and precision of temperature and discharge forecasts were compared using 3, 11, 19, and 31-day symmetric means to determine whether changes in the assumed exposure length causes significant changes in environmental forecast accuracy.

Pre-season run-timing estimates for Fraser sockeye salmon are derived from models evaluating Alaskan sea surface temperature and ocean currents, and the historic relationship between the timing of different run groups (Blackbourn 1987; Thomson et al. 1994; M. Folkes, DFO, Nanaimo, BC, pers. comm. 2006; I. Guthrie, PSC, Vancouver, BC, pers. comm. 2006). However, timing can be difficult to predict, even in-season, until after the peak of the run physically passes the Mission hydroacoustic facility (I. Guthrie, PSC, pers. comm. 2006). The large uncertainty in long range 50% date forecasts illustrates the importance of evaluating the sensitivity of environmental forecasts to changes in timing. We examined the sensitivity of regression statistics and prediction values to changing 50% date, and assessed the range of dates over which long range environmental forecast methods provide reliable predictions. The years – snowpack multiple regression and ensemble flow models were repeated over 50% dates changing in daily increments from July 10th – September 6th and we recorded model statistics (e.g. r^2 or adjusted r^2; and p-values) and prediction information (e.g. coefficient of variation, standard deviation, mean prediction, % bias compared to 2005 observed data).

MODEL COMPARISON

Ultimately, researchers will be interested in knowing which long range environmental forecast models are best suited to their needs. We summarised the performance of the various temperature and discharge models based on three performance measures: 1) model fit (using adjusted r^2), 2) the precision of the forecasted value (using coefficient of variation; CV), and 3) availability (when does the information required to use the forecast method become available?). Performance measures for each model were ranked and averaged over all run-timing groups, and recommendations for the best temperature and discharge models were provided.

RESULTS AND DISCUSSION

DIAGNOSTIC RESULTS

Preliminary examination of each forecasting method, with the exception of the ensemble flow technique, assumed simple linear regression models, based on a visual examination of data trends. Therefore, the first stage in the current assessment of pre-season environmental forecasts was to determine if data relationships indeed met the assumptions of simple linear regressions (see Table 2).

A summary of diagnostic results for each run-timing group, and each forecasting method, is provided in Table 3. With the exception of the relationship between historic temperature and discharge data, which consistently showed an improved fit using logged variables, the remaining long range forecasts were modelled using simple linear regressions. Although, on occasion, goodness-of-fit tests (either chi-square, or AIC) recommended alternative model fits, visual examination of linear, polynomial, and log-linear trends showed negligible differences, and a linear regression was still applied.

Table 3. Summary of linear regression diagnostic results for all forecasting methods and run-timing groups.

Forecast Method	**Historic temperature trend**			**Winter precipitation index**			**Snowpack water volume - discharge**			**Snowpack water volume - temperature**		
Group	*ESt*	*ES*	*S*	*ESt*	*ES*	*S*	*ESt*	*ES*	*S*	*ESt*	*ES*	*S*
Test												
Linearity	Pass (poly)	Pass	Pass (poly)	Pass	Pass	Pass	Pass	Pass	Pass	Pass (log)	Pass	Pass
Error Independence	Pass	Auto (insig)	Auto (insig)	Pass	Pass	Auto (insig)	Pass	Pass	Pass	Auto (insig)	Auto (sig)	Auto (sig)
Equal Variance	Pass	Pass	Pass	Pass	Pass	WLS	Pass	Pass	WLS	Pass	Pass	Pass
Normality	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass
Errors in Variables	Pass	Pass	Pass	Pass	Pass	Pass	Boot	Boot	Boot	Boot	Boot	Boot
Forecast Method	**Summer air anomaly**			**Spring air anomaly**			**Temperature - discharge**					
Group	*ESt*	*ES*	*S*	*ESt*	*ES*	*S*	*ESt*	*ES*	*S*			
Test												
Linearity	Pass (poly)	Pass	Pass	Pass (poly)	Pass (poly)	Pass (poly)	Log	Log	Log			
Error Independence	Auto (insig)	Pass	Pass	Pass	Pass	Auto (insig)	Auto (insig)	Pass	Auto (insig)			
Equal Variance	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass			
Normality	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass			
Errors in Variables	Boot	Boot	Boot	Pass	Pass	Pass	Boot	Boot	Boot			

ESt = Early Stuart; **ES** = Early Summer; **S** = Summer
Pass = met assumption for simple linear regression
(poly) = a polynomial model was selected through a step-wise AIC but was not employed
Log = a log-linear model provided a better fit than a linear model (chi-squared goodness of fit)
Auto (sig/insig) = a significant autocorrelation was observed; a follow-up Ljung-Box test showed whether a moving average model was a more significant fit to the data than a simple linear regression
WLS = variances showed significant heteroscedasticity; may want to consider a weighted least squares (or log-linear) modelling approach
Boot = there was significant error in the predictor variable, and a bootstrap approach was applied to incorporate this additional uncertainty

TEMPERATURE FORECAST MODELS

Historic temperature trends

There was a significant increase in water temperature over time for all three run-timing groups (r^2 = 9-17%; Figure 1). For the range of dates tested, the historic trend explained more of the variability in historic 19-day mean temperature later in the summer season.

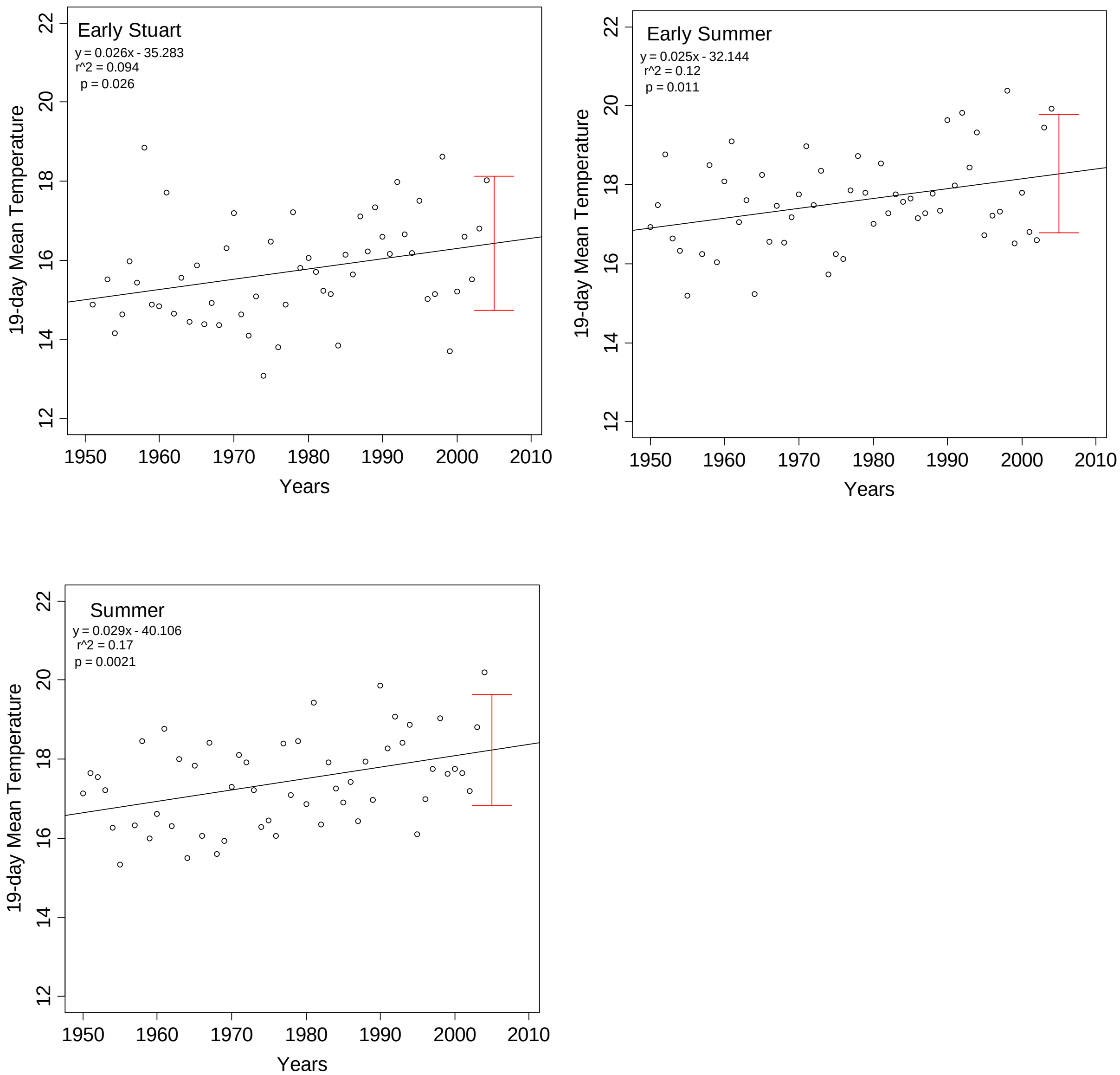

Figure 1. Historic 19-day mean temperature trends with fitted regression line and 80% prediction intervals for the forecasted year (2005). Prediction intervals extend, on average, by ±1.5°C.

Temperature-discharge correlations

There was a significant negative relationship between temperature and discharge for all run-timing groups (r^2 = 19-59%; Figure 2). However, results of a chi-squared goodness of fit test indicated that a log-linear model might provide a slightly improved fit. A few outliers may have influenced the model selection; therefore, we recommend further research into the true nature of the relationship between river temperature and discharge.

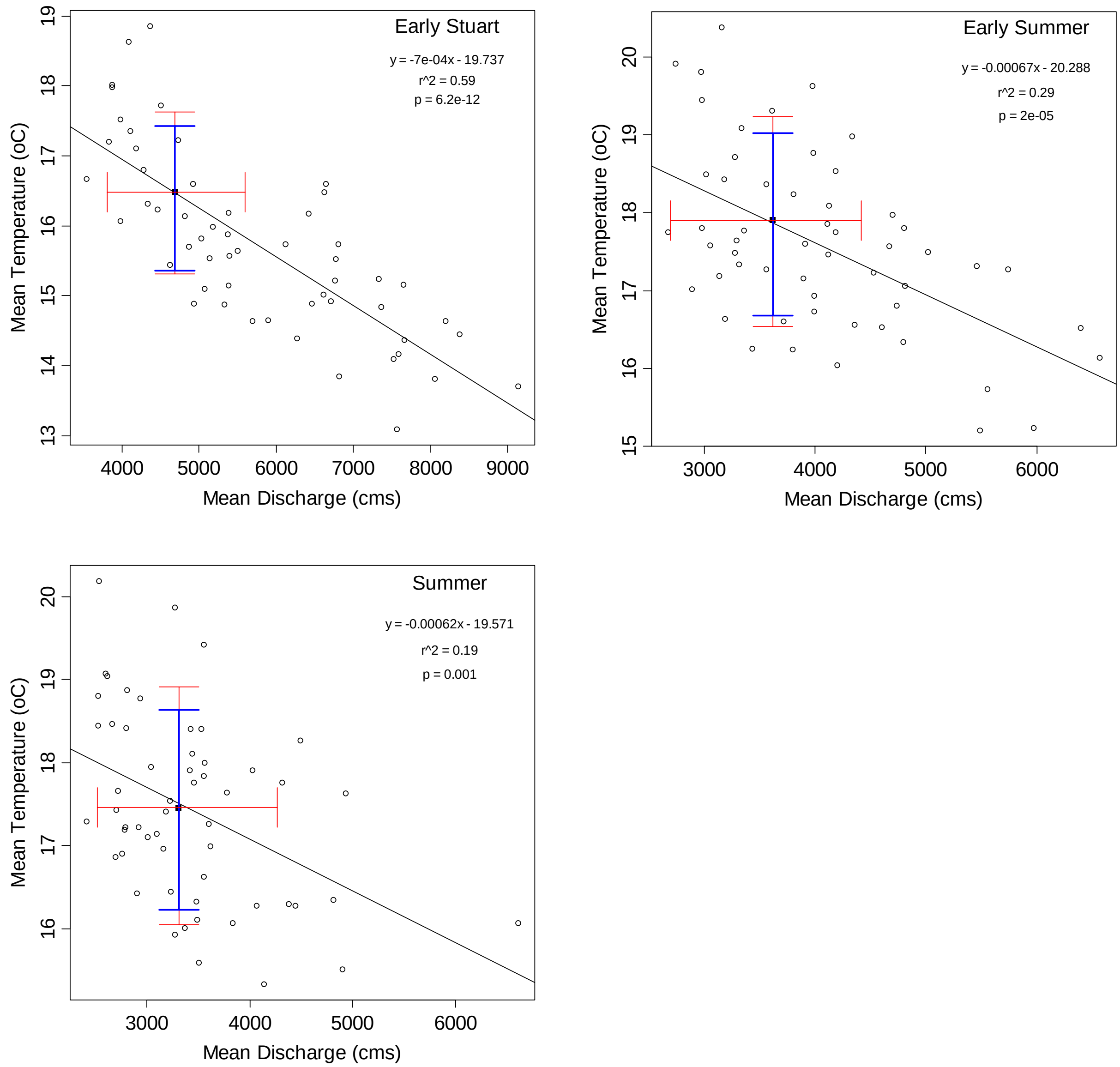

Figure 2. Historic 19-day mean discharge vs. 19-day mean temperature (open points) with best fit regression line. Bootstrapped 80% confidence limits for the mean predicted discharge and temperature are shown in red. The 80% prediction intervals generated by the linear regression are overlaid in blue. Forecasts correspond to discharges predicted from the June 2005 ensemble flow model.

Summer air anomaly

Bootstrap method I: Environment Canada approach

Summer air temperatures were a significant predictor of water temperature throughout the season (r^2 = 45-51%; Figure 3).

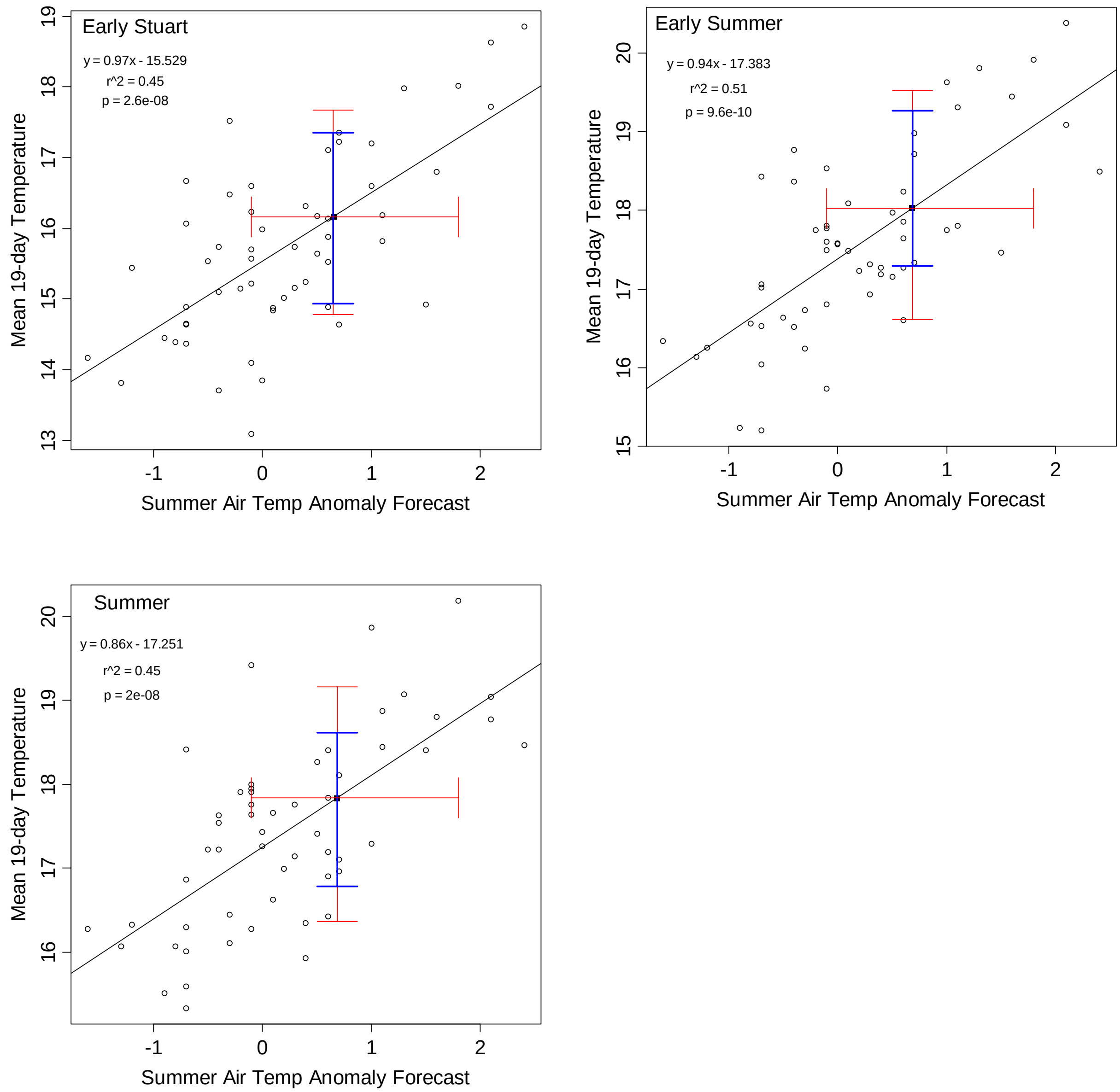

Figure 3. Summer air anomaly vs. 19-day mean temperature trends with fitted regression line. Bootstrapped 80% confidence limits for the mean predicted air anomaly and temperature are shown in red. The 80% prediction intervals generated by the linear regression are overlaid in blue. Forecasts correspond to the 60% “above normal”, 30% “near normal” and 10% “below normal” air temperature anomalies predicted by Environment Canada in 2005.

Bootstrap method II: historic trend approach

There was a significant positive linear relationship between years and historic summer air anomaly (r^2 = 10%; Figure 4). Therefore, the historic regression was used to predict uncertainty in the summer anomaly forecast. Uncertainty in river temperature predicted from the historically forecasted summer air anomaly was similar to the uncertainty predicted using the Environment Canada method.

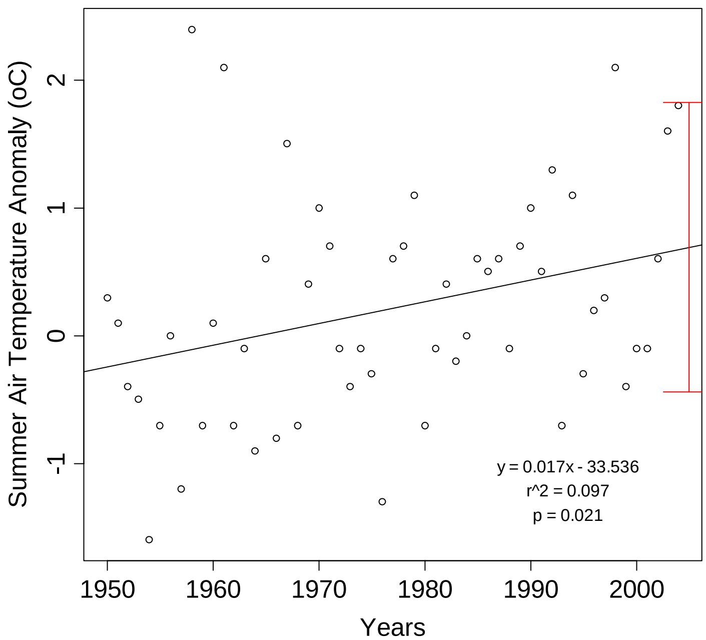

Figure 4. Historic trend in summer air temperature anomaly shown with fitted regression line and 80% prediction intervals for 2005 corresponding to a mean anomaly of 0.69°C.

Spring air anomaly

The positive relationship between summer river temperatures and spring air anomalies was significant, but weaker (r^2 = 25-34%) than the relationship using summer anomalies (Figure 5). Because the spring anomaly is a measured value, prediction intervals were derived directly from the linear regression.

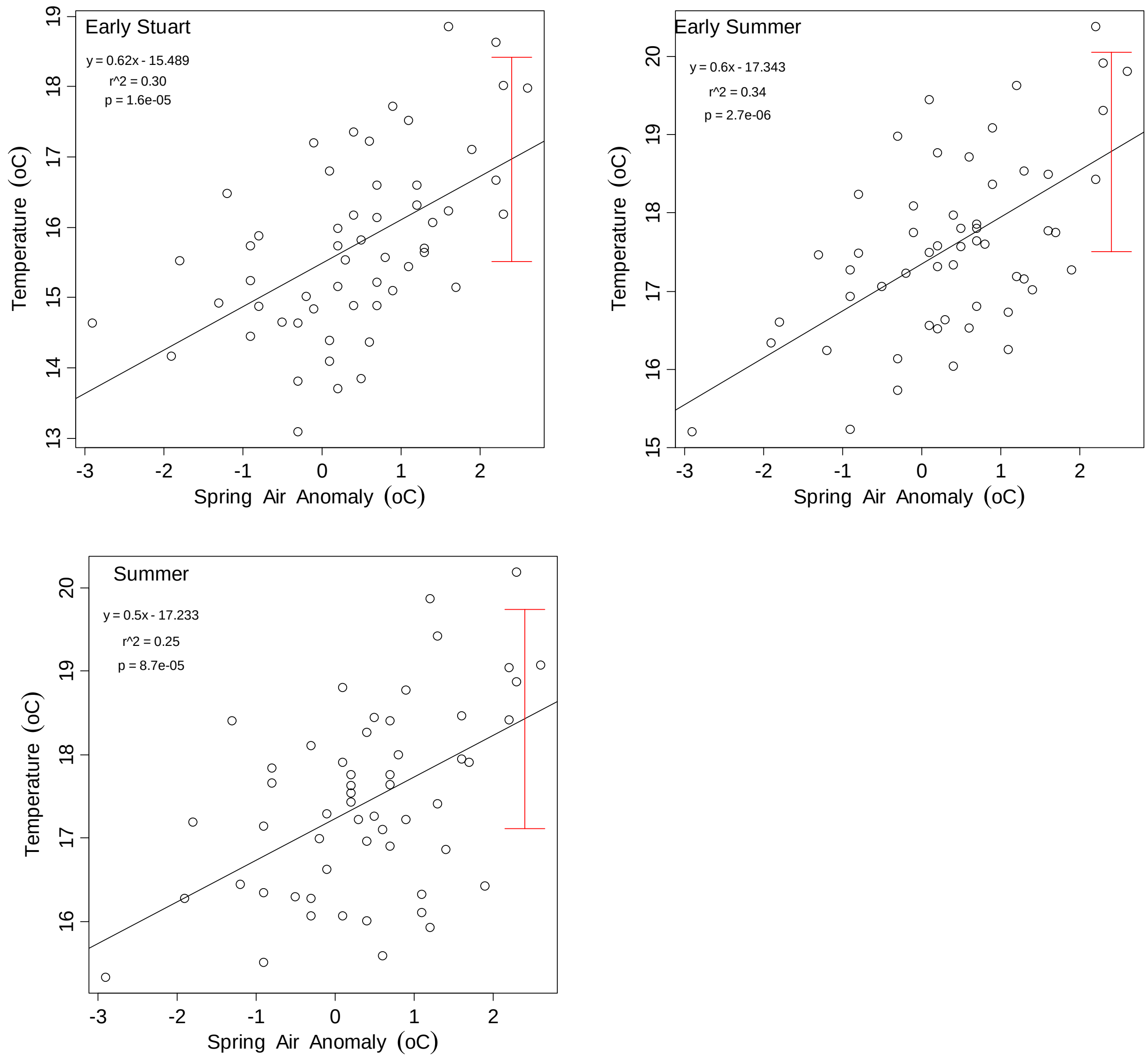

Figure 5. Spring air temperature anomalies vs. 19-day mean temperature trends with fitted regression line and 80% prediction intervals for the forecasted year (2005). Forecasts correspond to the 2005 2.4^{o}C spring anomaly.

Multiple regression analysis

Multiple regression linear models explained more variance in historic temperature than the single variable models using each component variable. The fit of all multiple regression models was measured using adjusted r^2. The summer air anomaly – snowpack water volume model explained the greatest level of variance in observed temperature trends (adjusted r^2 = 45-65%). The years – ensemble flow model (adjusted r^2 = 28-43%) and the years – snowpack water volume model (adjusted r^2 = 21- 47%), performed comparably well.

DISCHARGE FORECAST MODELS

Winter precipitation index

A significant positive linear fit was observed between winter precipitation anomalies and mean discharge for all run-timing groups (r^2 = 9-19%; Figure 6). However, the relationship between winter precipitation and temperature (not shown) was only significant for the Early Stuart run (r^2 = 20%), and became insignificant in early August. Due to its limited predictive capacity, the winter precipitation – temperature relationship was not considered further.

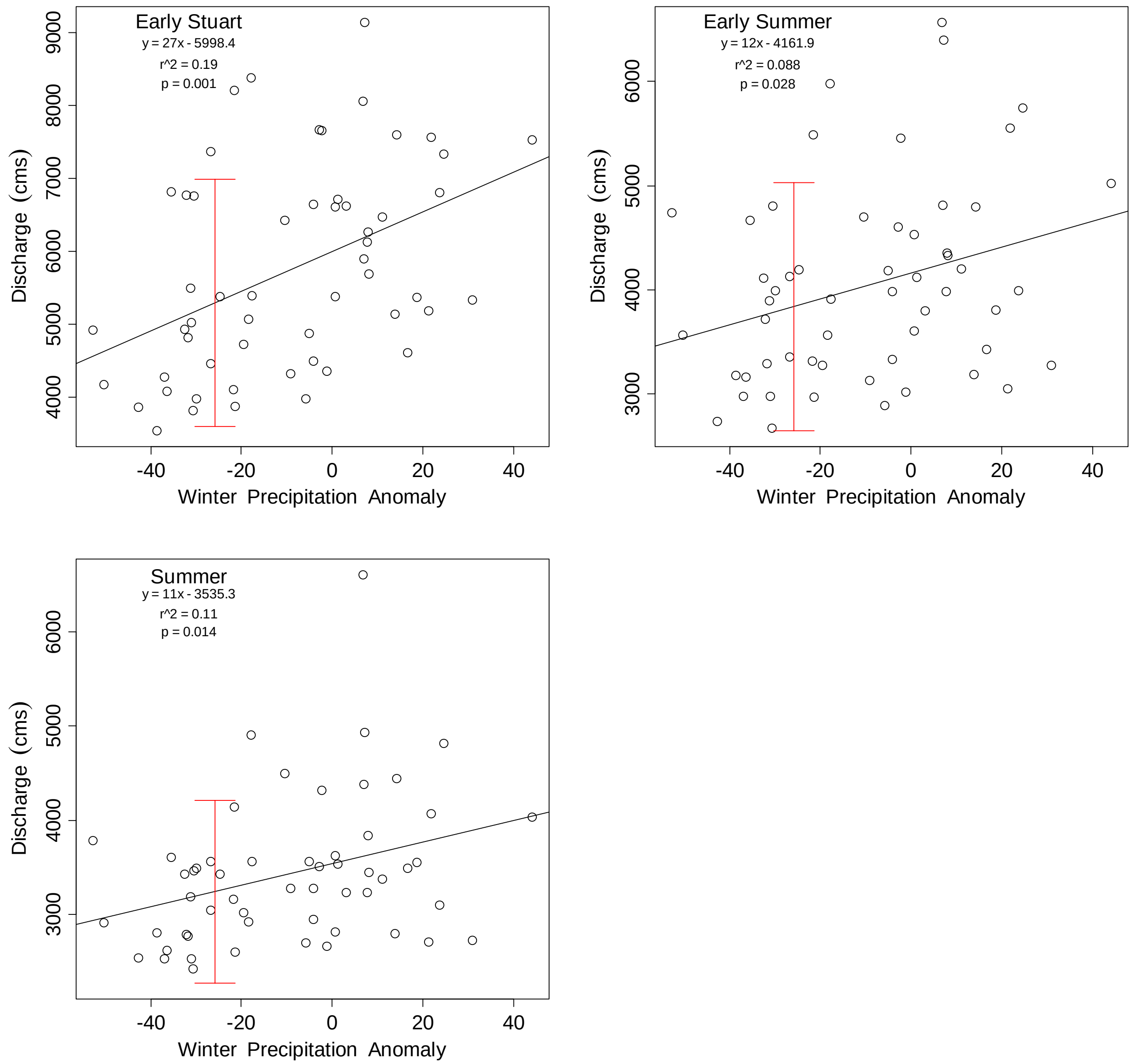

Figure 6. Winter precipitation index vs. 19-day mean discharge trends with fitted regression line and 80% prediction intervals for the forecasted year (2005). A precipitation anomaly of -25.9% was recorded for 2005, indicating a drier than average season.

Snowpack water volume forecast – discharge

There was a significant positive relationship between May volume and discharge (r^2 = 63-78%; Figure 7) and June volume and discharge (r^2 = 66-89%; Figure 8) for all run-timing groups. The significance of these results should be interpreted with caution because the calculation of historic volume estimates is dependent on total summer discharge (including the 19-day prediction interval). However, because we are interested in the relationship for its prediction capacity, and not for hypothesis testing, this partial inter-dependence should not influence our forecast conclusions.

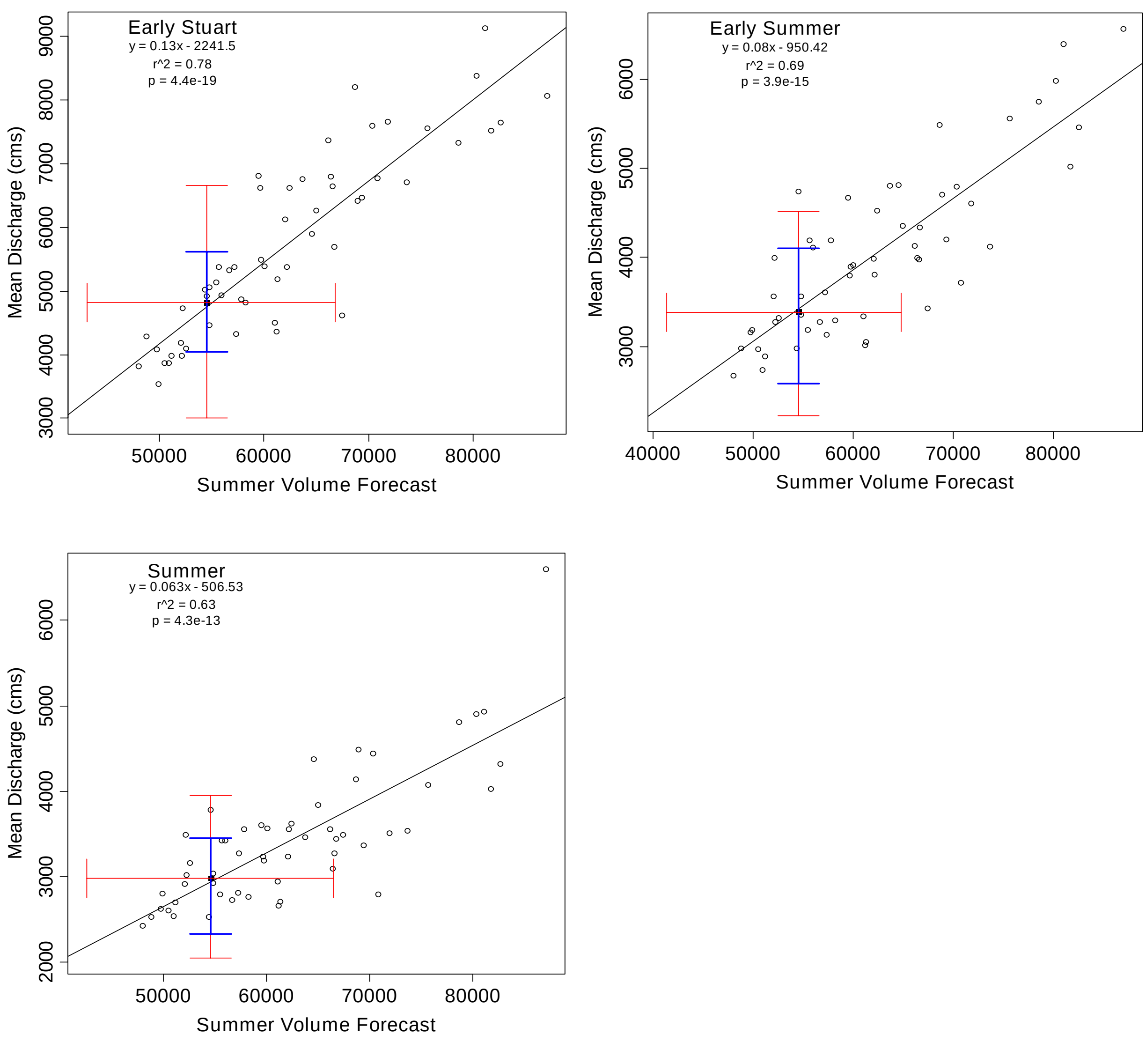

Figure 7. May snowpack water volume vs. 19-day mean discharge trends with fitted regression line. Bootstrapped 80% confidence limits for the 2005 volume forecast and predicted mean discharge are shown in red. The 80% prediction intervals generated by the linear regression are overlaid in blue. Forecasts correspond to the May 1 mean snowpack water volume prediction of 56 400 million m^3 (93% of the historic norm) by the BC River Forecast Centre.

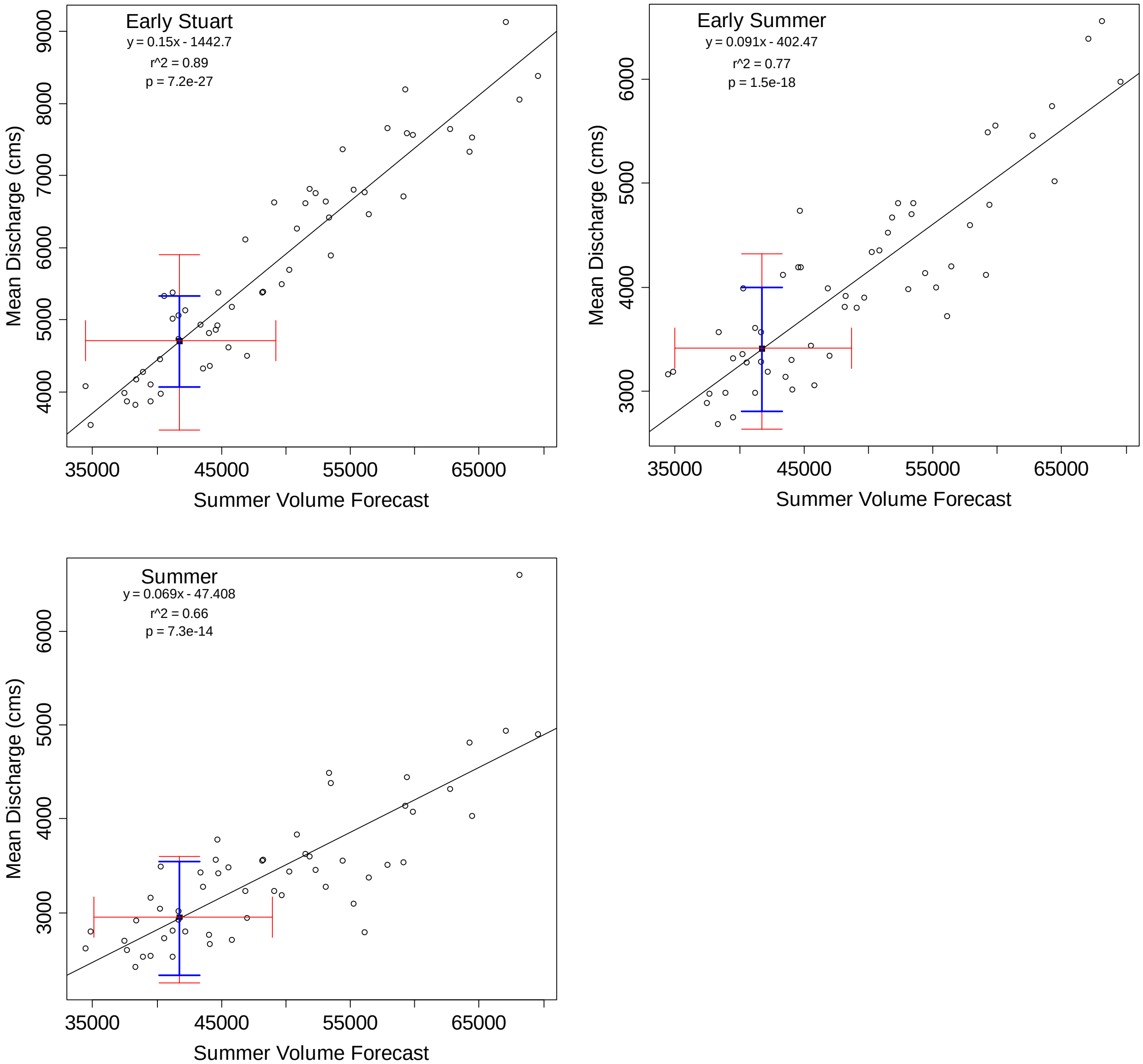

Figure 8. June snowpack water volume vs. 19-day mean discharge trends with fitted regression line. Bootstrapped 80% confidence limits for the 2005 volume forecast and predicted mean discharge are shown in red. The 80% prediction intervals generated by the linear regression are overlaid in blue. Forecasts correspond to the June 1 mean snowpack water volume prediction of 41762 million m^3, which was derived from the May 1 forecast of 56400 million m^3 provided by the BC River Forecast Centre and the total Hope discharge observed during May 2005.

Snowpack water volume forecast – temperature

Although we were primarily interested in the relationship between snowpack water volume and discharge, snowpack water volume also demonstrated proficiency for forecasting river temperatures. There was a significant negative relationship between May snowpack water volume and river temperature (r^2 = 13-46%; Figure 9) and June volume and temperature (r^2 = 17-52%; Figure 10) for the three run-timing groups.

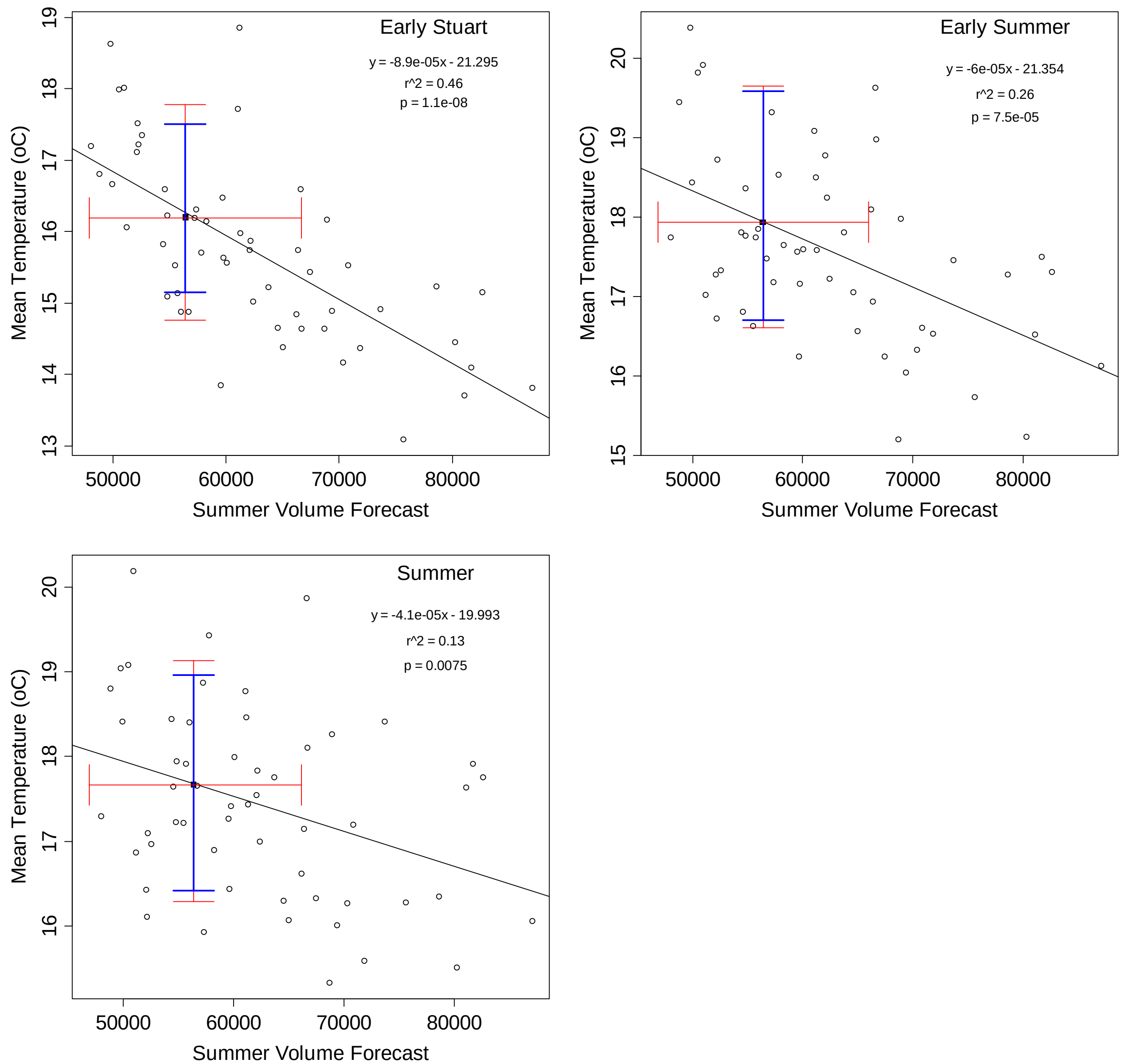

Figure 9. Snowpack water volume vs. 19-day mean temperature trends with fitted regression line. Bootstrapped 80% confidence limits for the 2005 volume forecast and predicted mean discharge are shown in red. The 80% prediction intervals generated by the linear regression are overlaid in blue. Forecasts correspond to the May 1 mean snowpack water volume prediction of 56 400 million m^3 (93% of the historic norm) by the BC River Forecast Centre.

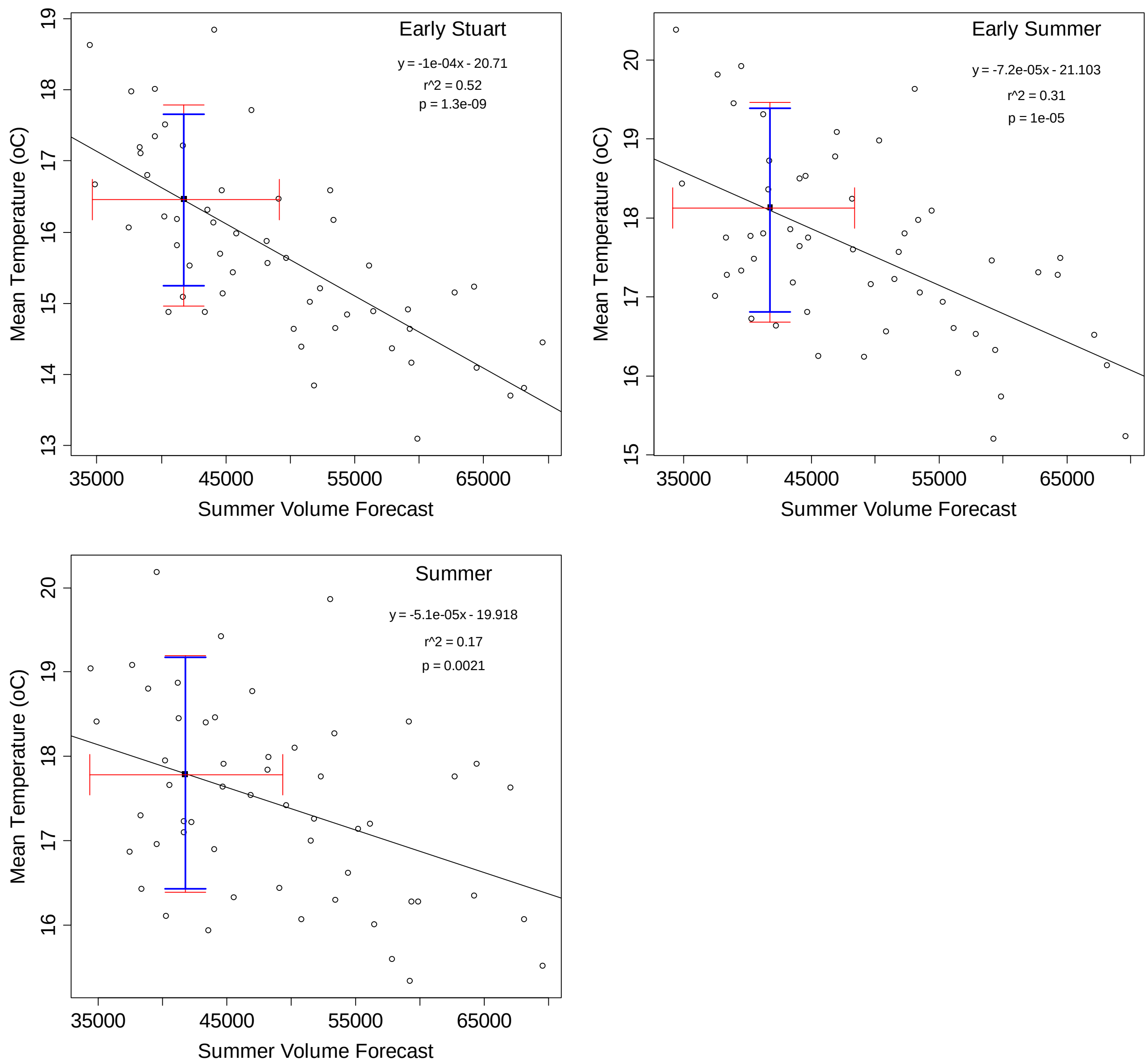

Figure 10. June snowpack water volume vs. 19-day mean temperature trends with fitted regression line. Bootstrapped 80% confidence limits for the 2005 volume forecast and predicted mean temperature are shown in red. The 80% prediction intervals generated by the linear regression are overlaid in blue. Forecasts correspond to the June 1 mean snowpack water volume prediction of 41762 million m^3, which was derived from the May 1 forecast of 56400 million m^3 provided by the BC River Forecast Centre and the total Hope discharge observed during May 2005.

Ensemble flow technique

The discharge estimates from ensemble flow models produced by the RFC in May and June 2005 are presented in Figures 11 and 12 (open points). Uncertainty in forecasted discharge showed a tendency to increase with longer-range predictions. Limited information was available on the exact structure of the RFC ensemble flow model, or the potential for error in the modelled values.

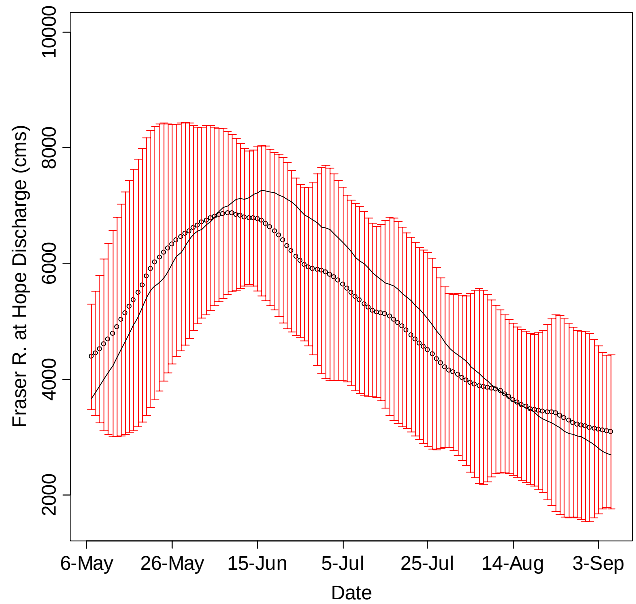

Figure 11. Forecasted daily river discharge using the May 2005 ensemble flow model (open circles) plotted with +/- 2 standard deviations (red error bars), and the historic mean discharge (black line).

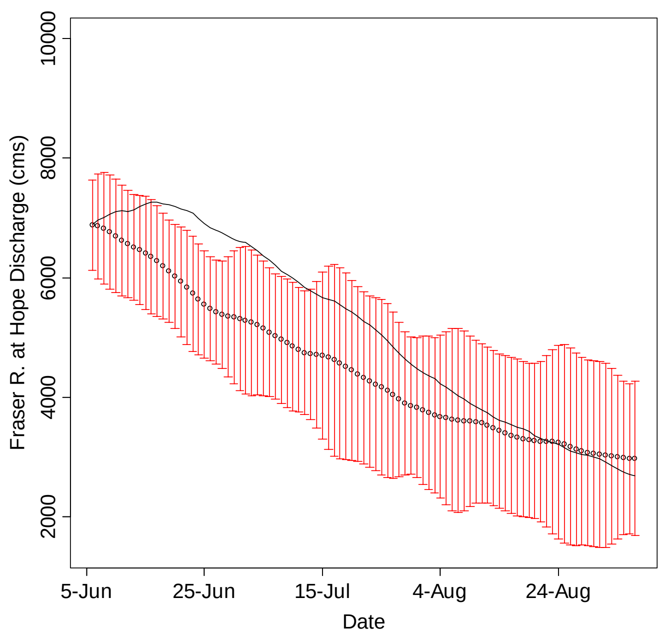

Figure 12. Forecasted daily river discharge using the June 2005 ensemble flow model (open circles) plotted with +/- 2 standard deviations (red error bars), and the historic mean discharge (black line).

SENSITIVITY ANALYSIS

Seasonality

Multiple regression analysis

There were considerable seasonal trends in the adjusted r^2 value, the p-value (significance of the regression slope), the coefficient of variation and standard deviation in predicted temperature, the percent bias compared to the measured 2005 temperature, and the predicted mean. Figure 13 clearly illustrates the effect of 50% date estimates on statistical output using the years - snowpack water volume multiple regression model.

The amount of variation in mean temperature explained by years and snowpack water volume remained high (adjusted $r^2 > 0.45$) from approximately July 10 – July 25 (Figure 13A). The adjusted r^2 value then declined over time (tracked until September 6), with the exception of another plateau in mid-August. The plateaus may occur because the year – temperature relationship improves from early July until late August, while the volume – temperature relationship weakens, occasionally balancing each other out. The declining adjusted r^2 trend observed in the temperature – volume relationship ultimately dominates the overall trend observed for the multiple regression because volume explains a greater percentage of the variability in historic temperature. Despite the decline in explained variability, the regression slope remained significant throughout the date range examined (Figure 13B).

Standard deviation (SD) and coefficient of variation (CV) represent absolute and relative error measures, respectively (Figure 13C and D). The CV in mean temperature declined until early August, as river temperature (Figure 13F) continued to rise but SD remained roughly constant. Both the CV and predicted temperature plateaued during mid-August, roughly coinciding with the plateau in r^2. During the same period, 2005 predictions were also relatively unbiased (percent bias ~ -1 to 1%; Figure 13E). The model plateau is of particular interest, as it suggests that the model is less sensitive to uncertainties in run-timing dates during mid-August than variability arising earlier or later in the season.

The decline in model performance after August may signify the approximate time when the majority of the snowmelt has passed through Hells Gate, and river temperature becomes more dependent on other proximate factors. The rapid model deterioration illustrates the time span limitations of this pre-season forecasting approach. For example, the model is not significant for median dates historically associated with the Late run-timing group (historic median date = September 13). If the unusually late run-timings observed in 2005 become the norm (July 27th, September 3rd, and September 4th, for the Early Stuart, Early Summer and Summer groups, respectively), most of the pre-season forecast methods will no longer be suitable for predicting river conditions for Early Summer and Summer groups. In contrast, if Late run fish continue their current trend of entering the Fraser River from mid-August to early September, (as opposed to their historic timing of late September to early October; Lapointe et al. 2003; Cooke et al. 2004), then it may become possible to generate pre-season forecasts for this run-timing group.

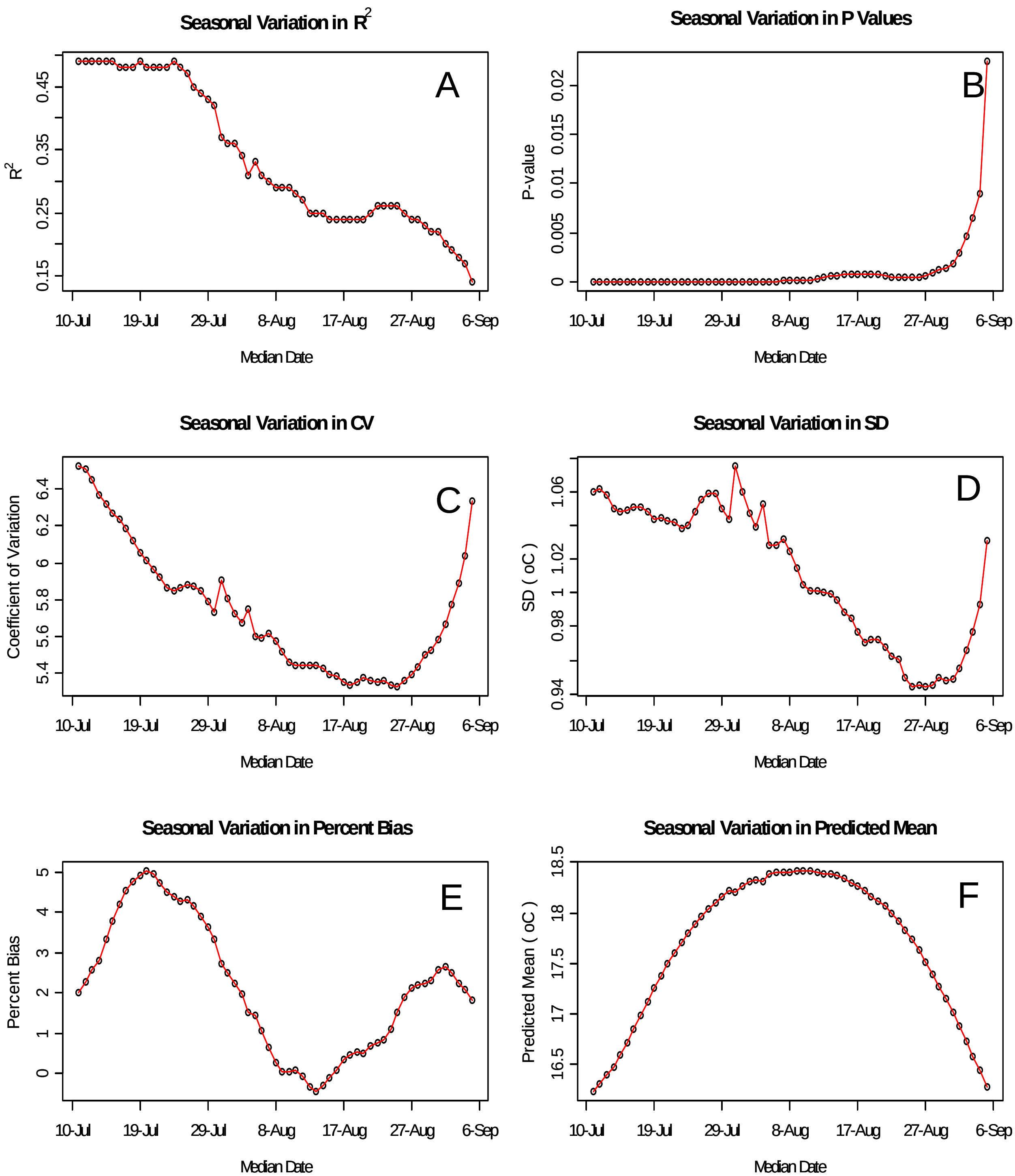

Figure 13. Time series analysis of the years - snowpack water volume multiple regression temperature prediction model illustrating the seasonal variability in 19-day mean temperature predictions and associated statistics. Seasonal variation in percent bias refers to 2005 predictions.

June ensemble flow

Figure 14 illustrates a deteriorating trend in the performance of the ensemble flow model over time. Again, the decline in model power may reflect the decreasing influence of snowpack water volume (used to create the ensemble) on river discharge as the season progresses (r^2 and p-values were not generated because the ensemble model does not use linear regression). A comparison between Tables 6 and 8 illustrates the larger CVs associated with discharge forecasts as opposed to temperature forecasts. There are also no plateau regions, as observed with the multiple regression model. These results highlight the limitations in using long range ensemble flow discharge predictions later in the summer season.

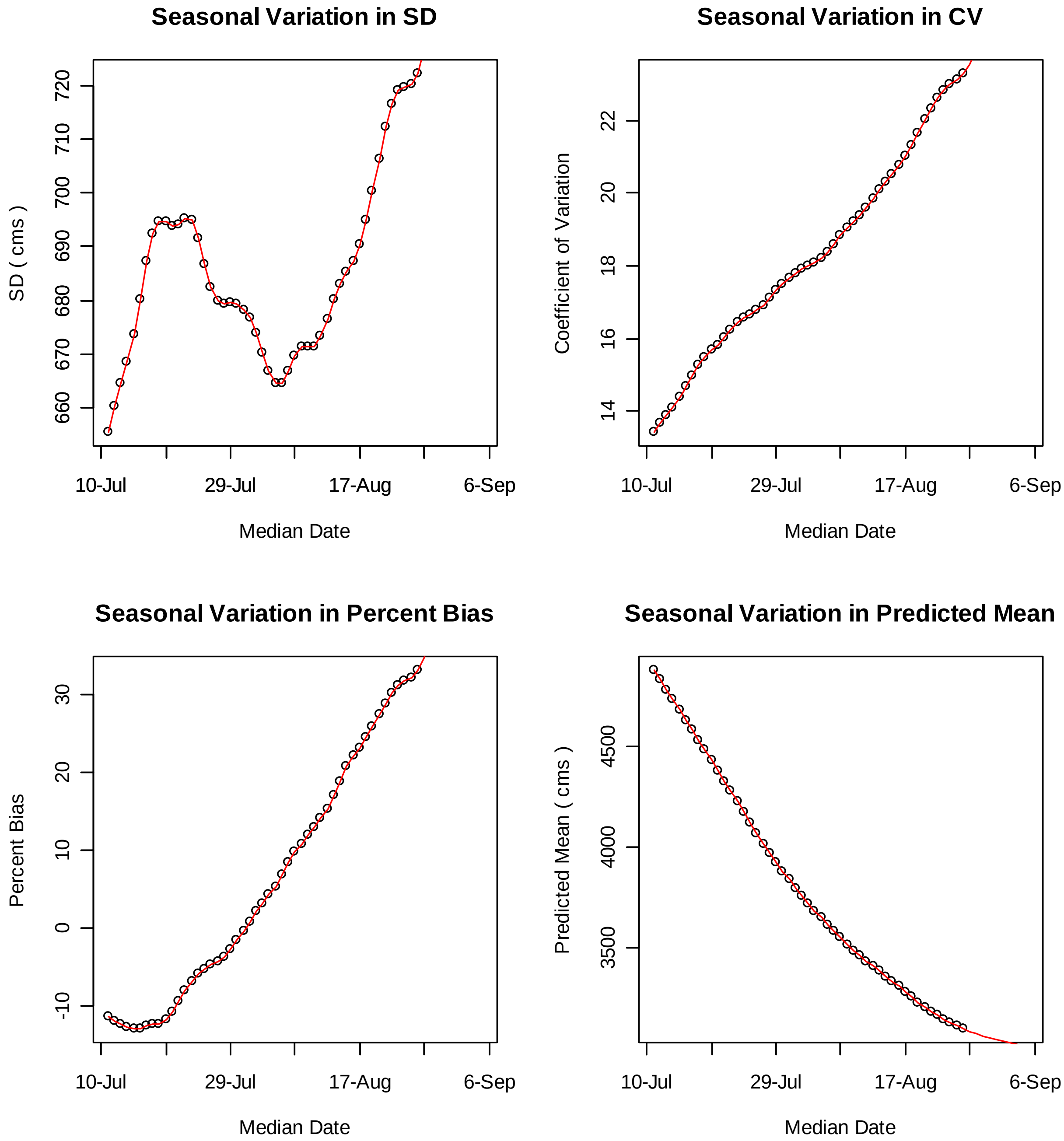

Figure 14. Time series analysis of the June 2005 ensemble flow model illustrating the variability in 19-day mean discharge predictions and associated statistics over time. Seasonal variation in percent bias refers to 2005 predictions.

Mean days

The examples provided thus far analysed symmetric 19-day mean temperature and discharge averages centred on the Hells Gate 50% date. The 19-day period was selected for the sockeye salmon examples, based on the assumption that this range captures most of the conditions experienced by each run-timing group as they pass Hells Gate. For example, previous DBE models for Fraser River sockeye salmon management have used both 19-day and 31-day means (D. Patterson, DFO, pers. comm. 2006; I. Guthrie, PSC, pers. comm. 2006). The sensitivity of predictions to the number of days used to generate mean temperature and discharge is illustrated in Tables 4 and 5.

Increasing the number of days used to calculate the mean does lead to small improvements in the CV of predicted temperature and the adjusted r^2 calculated from the multiple regression model (Table 4). The improvements are most noticeable moving from a 3-day to 11-day mean. Similar trends were observed for the ensemble flow discharge predictions (Table 5). Increasing the number of days likely improves model performance by smoothing out sources of higher frequency variation in the data that do not contribute to the covariation between the environmental time series (Pyper and Peterman 1998).

Table 4. Sensitivity of mean temperature prediction statistics (CV = coefficient of variation in predicted temperature; r^2 = coefficient of determination for best fit linear regression) from the years – snowpack water volume multiple regression model to the number of days used to calculate the mean. The historic Early Stuart peak run-timing date (July 14) is used as an example.

	3-day	**11-day**	**19-day**	**31-day**
Predicted mean (^{o}C)	16.61	16.59	16.58	16.77
CV (%)	7.28	6.57	6.39	6.26
r^2	0.39	0.47	0.49	0.52

Table 5. Sensitivity of mean discharge prediction statistics (CV = coefficient of variation in predicted temperature) from the June 2005 ensemble flow model to the number of days used to calculate the mean. The historic Early Stuart peak run-timing date (July 14) is used as an example.

	3-day	**11-day**	**19-day**	**31-day**
Predicted mean (cms)	4719	4698	4685	4668
CV (%)	15.15	13.58	12.27	11.31

MODEL COMPARISON

Temperature forecasts

A retrospective analysis of 2005 pre-season temperature forecasts is presented in Table 6. Both summer air anomaly forecasting methods performed similarly, so only the historic trend technique (bootstrap method II) is provided for illustrative purposes. Performance metrics include the mean predicted temperature, the r^2 value corresponding to the best-fit linear regression, and the coefficient of variation (CV) of the predicted value. The percent bias presented is only associated with the 2005 prediction and may not be representative of bias trends in alternate years.

Table 6. Summary of pre-season temperature forecasting methods and the precision and bias of 2005 forecasts. Predicted means are presented in bold. r^2 = coefficient of determination corresponding to best fit linear regression (adjusted r^2 for the multiple regression models); CV = coefficient of variation (CV = (forecast standard deviation/forecast mean)*100); B = percent bias for 2005 (B = ((forecast mean – measured value)/measured value)*100).

Method	Availability	Early Stuart	Early Summer	Summer
Historic temperature trend	Anytime	**16.38°C** *r^2 = 9.4%* *CV = 7.45%* *B = 1.74%*	**18.26°C** *r^2 = 11%* *CV = 6.00%* *B = 0.33%*	**18.22 °C** *r^2 = 17%* *CV = 5.71%* *B = -0.44%*
Snowpack water volume (May)	May	**16.26°C** *r^2 = 46%* *CV = 6.89%* *B = 0.99%*	**17.97°C** *r^2 = 26%* *CV = 5.73%* *B = -1.26%*	**17.66°C** *r^2 = 13%* *CV = 6.17%* *B = -3.50%*
Snowpack water volume (June)	June	**16.44°C** *r^2 = 52%* *CV = 6.67%* *B = 2.11%*	**18.13°C** *r^2 = 31%* *CV = 5.68%* *B = -0.38%*	**17.80°C** *r^2 = 17%* *CV = 5.90%* *B = 5.31%*
Years – snowpack water volume multiple regression analysis	April	**16.58°C** *r^2 = 47%* *CV = 6.39%* *B = 2.98%*	**18.41°C** *r^2 = 29%* *CV = 5.54%* *B = 1.15%*	**18.31°C** *r^2 = 21%* *CV = 5.68%* *B = 0.05%*
Historic trend – ensemble multiple regression analysis (June data)	June	**16.71°C** *r^2 = 28%* *CV = 7.36%* *B = 3.79%*	**18.41°C** *r^2 = 38%* *CV = 6.19%* *B = 1.15%*	**18.13°C** *r^2 = 43%* *CV = 6.01%* *B = -0.09%*
Summer air anomaly – snowpack water volume multiple regression analysis	June	**16.43°C** *r^2 = 65%* *CV = 6.57%* *B = 2.05%*	**18.18°C** *r^2 = 56%* *CV = 5.33%* *B = -0.11%*	**17.94°C** *r^2 = 45%* *CV = 6.02%* *B = -1.97%*
Temperature-discharge correlation (June ensemble)	June	**16.50°C** *r^2 = 62%* *CV = 6.24%* *B = 2.48%*	**17.85°C** *r^2 = 29%* *CV = 6.33%* *B = -1.92%*	**17.51°C** *r^2 = 21%* *CV = 6.62%* *B = -4.32%*
Summer air anomaly (historic trend bootstrap)	Anytime	**16.23°C** *r^2 = 45%* *CV = 7.27%* *B = 0.80%*	**18.06°C** *r^2 = 51%* *CV = 6.20%* *B = -0.77%*	**17.86°C** *r^2 = 45%* *CV = 5.71%* *B = -2.40%*
Spring air anomaly	June	**16.96°C** *r^2 = 29%* *CV = 5.90%* *B = 5.34%*	**18.78°C** *r^2 = 34%* *CV = 5.06%* *B = 3.19%*	**18.43°C** *r^2 = 26%* *CV = 5.32%* *B = 0.71%*
2005 Measured		**16.1°C**	**18.2°C**	**18.3°C**

The temperature forecasting methods performed comparably well. Percent bias for 2005 forecasts was typically <5%, and CVs did not exceed 8%. In all cases, the measured 2005 river temperature fell within the 80% prediction (or bootstrapped confidence) intervals produced by all models. There was no general temporal trend in model performance over the time range evaluated; however, we noted that almost all relationships became insignificant by late August/early September. The fit of the historic trend model improved over the season until mid-August, while discharge and snowpack water volume explained less temperature variability as the season progressed. Finally, spring and summer air anomaly data were most strongly related to river temperature during early August (Early Summer run-timing).

The main trade-offs to consider when selecting the most appropriate temperature forecast method are the seasonal availability of the data, model performance, and predictor variable uncertainty. For example, although the historic temperature trend model explains a small portion of the variance in summer water temperatures, and so exhibits poor model fit, the historic trend model is available year round and there is no bias or uncertainty in the predictor variable. The models that utilised the close association between river temperature and discharge demonstrated reasonable model fit but were limited by the substantial uncertainty in the forecasted predictor variable. The same issue arises using any model relying on a forecasted predictor variable; the uncertainty in the predictor variable combines with the uncertainty in the model, and may result in an imprecise prediction even from models with reasonably high r^2 values.

The addition of a second explanatory variable in the three multiple regression models typically improved both model fit and prediction precision compared to models using a single predictor. In most cases a step-wise AIC procedure selected the multiple regression over single variable models. The single exception was the summer air anomaly – snowpack water volume model for the Summer run, which was best-fit using only the summer air anomaly data. Although the AIC procedure selected the historic trend – ensemble flow multiple regression, the precision in the predicted river temperature actually decreased compared to the historic temperature trend analysis alone. As discussed above, the decrease in precision could be due to the additional uncertainty introduced by the ensemble forecast. The general trend towards improvement in temperature forecasts using multiple regression models is an area for future research, as this report evaluated only a few potential options.

Table 7 provides a general summary of the performance of each temperature forecast model with respect to model availability, and average CV and r^2 values over all run-timing groups. Percent bias was not used as a model selection tool because 2005 results are unlikely to be indicative of performance in other years. In general, the strength of the model relationship improved when the predictor variable(s) was a value measured closer to the summer months (e.g. historic discharge, summer air anomaly); however, the late nature of these data necessitates the use of a forecasted predictor variable, which decreases the precision of the temperature estimate. Given the results of the 2005 retrospective analysis, the years – volume and summer air anomaly - volume multiple regression models both performed well. These two multiple regression models can be performed as early as April (if April snowpack water volume forecasts are available from the RFC and the historic trend method is used to predict the summer air anomaly). Ultimately, the individual requirements of a given research project or management objective will determine which forecast method is most appropriate.

Table 7. Rank comparison of temperature forecasting methods. A lower rank indicates improved performance. In the case of a tie, equal ranks are applied to each model. Means represent averages over results for all three run-timing groups.

Method	**Availability** *Rank*	**Mean CV (%)** *Rank*	**Mean r^2 (%)** *Rank*
Historic temperature trend	**Anytime** *1*	**6.38** *6*	**13** *9*
Snowpack water volume (May)*	**May** *2*	**6.26** *5*	**28** *8*
Snowpack water volume (June)*	**June** *3*	**6.08** *4*	**33** *4*
Years – snowpack water volume multiple regression analysis*	**May** *2*	**5.87** *2*	**32** *6*
Years – ensemble multiple regression analysis (June data)	**June** *3*	**6.52** *9*	**36** *5*
Summer air anomaly – snowpack volume multiple regression analysis*	**May** *2*	**5.97** *3*	**55** *1*
Temperature-discharge correlation (June ensemble)*	**June** *3*	**6.40** *8*	**37** *3*
Summer air anomaly (historic trend bootstrap)*	**Anytime** *1*	**6.39** *7*	**47** *2*
Spring air anomaly	**June** *3*	**5.43** *1*	**30** *7*

*Indicates model includes uncertainty in predictor variable i.e. CV reflects error in both model structure and predictor variable.

DISCHARGE FORECASTS

A retrospective analysis of 2005 pre-season discharge forecasts is presented in Table 8. In 2005, the performance of discharge models was poor compared to the temperature forecasts, with CVs often greater than 20% and a consistent trend towards increasing bias for late season predictions.

Table 8. Summaries of pre-season discharge forecasting methods and their precision and bias in 2005. Predicted means are presented in bold. r^2 = coefficient of determination corresponding to best fit linear regression; CV = coefficient of variation (CV = (forecast standard deviation/forecast mean)*100); B = percent bias for 2005 (B = ((forecast mean – measured value)/measured value)*100).

Method	Availability	Early Stuart	Early Summer	Summer
Winter precipitation index	March	**5296cms** *r^2 = 19%* *CV = 24.30%* *B = -1.65%*	**3841cms** *r^2 = 8.8%* *CV = 23.56%* *B = 13.47%*	**3241cms** *r^2 = 11%* *CV = 22.74%* *B = 19.51%*
Volume to discharge (May)	May	**5021cms** *r^2 = 78%* *CV = 25.63%* *B = -6.76%*	**3562cms** *r^2 = 69%* *CV = 22.15%* *B = 5.23%*	**3021cms** *r^2 = 63%* *CV = 22.38%* *B = 11.39%*
Volume to discharge (June)	June	**4657cms** *r^2 = 89%* *CV = 18.87%* *B = -13.52%*	**3369cms** *r^2 = 77%* *CV = 20.87%* *B = -0.47%*	**2932cms** *r^2 = 66%* *CV = 21.73%* *B = 8.11%*
Ensemble technique (May)	May	**5079 cms** *CV = 14.39%* *B = -5.68%*	**3877 cms** *CV = 16.48%* *B = 14.53%*	**3542 cms** *CV = 18.38%* *B = 30.60%*
Ensemble technique (June)	June	**4685cms** *CV = 12.27%* *B = -13.00%*	**3618cms** *CV = 16.03%* *B = 6.88%*	**3312cms** *CV = 18.54%* *B = 18.12%*
2005 Measured		**5385 cms**	**3385 cms**	**2712 cms**

In general, there was an inverse trend between discharge model availability and performance, with models available later in the season showing improved precision and model fit (see Table 9). As a result, the June ensemble flow model performed the best of the models evaluated. Ensemble flow models are updated throughout the spring with improved snowpack data, and 2005 results indicate a trend towards improved performance with later model runs. The evaluation of discharge forecast methods also involves trade-offs between model performance and predictor variable forecast uncertainty. For example, variability in the snowpack water volume forecast resulted in similar CVs predicted by both the snowpack water volume and winter precipitation models, even though snowpack water volume explains a much larger percent of the variability in historic discharge.

Table 9. Rank comparison of discharge forecasting methods. A lower rank indicates improved performance. In the case of a tie, equal ranks are applied to each model. Means represent averages over results for all three run-timing groups.

Method	**Availability** *Rank*	**Mean CV (%)** *Rank*	**Mean r^2 (%)** *Rank*
Winter precipitation index	**March** *1*	**23.53** *4*	**13** *4*
Volume to discharge (May)*	**May** *2*	**23.68** *5*	**70** *3*
Volume to discharge (June)*	**June** *2*	**20.49** *3*	**77** *2*
Ensemble technique (May)	**May** *2*	**16.41** *2*	**n/a**** *1*
Ensemble technique (June)	**June** *3*	**15.61** *1*	**n/a**** *1*

* Indicates model includes uncertainty in predictor variable i.e. CV reflects error in both model structure and predictor variable.
**Discharge is derived directly from the ensemble forecast, i.e. there is no linear regression.

CONCLUSIONS

Our analyses quantified prediction uncertainty arising from model structure and environmental forecast variability for several long range Fraser River summer temperature and discharge prediction models. However, our results are dependent on several underlying assumptions. Failure to meet these assumptions could potentially lead to further biases in the long range forecasts and an underestimation of forecast error. Simulation analyses may be a useful tool to evaluate the potential effect of varying degrees of deviation from the following assumptions.

First, analyses assumed no measurement error in the historic data. Given the length of the historic record for both lower Fraser temperature, there are several instances in which data gaps and data quality issues exist (Patterson et al. 2007b). Plots of historic trends indicate some temperature and discharge outliers, and further examination of the data is required to determine whether these outliers are simply extremes within the natural range, or whether they are due to measurement error. On the same note, future research should also consider sensitivity analyses looking at variability in regression statistics with respect to the number of years used to fit the model. Foreman et al. (2001) and Patterson et al. (2007b) found that the statistical significance of historic temperature trends in the Fraser varied depending on the time series analysed. For example, one may want to subset the years of data pertaining to the Pacific Decadal Oscillation.

The second assumption was that current trends continue beyond the scope of known observations. This is of particular importance for the models that use the historic trend, in which case every year's prediction is an extrapolation from the model. Given the observed increase in river temperature over time (Morrison et al. 2002), future trends may deviate from those demonstrated by historic data. Researchers should interpret extrapolated values with caution, as one cannot be certain that current trends will hold true at extreme ranges (Zar 1996).

Our third assumption was that the forecasted environmental conditions were synonymous with the conditions experienced by the species of interest. However, this assumption is too simplistic for migrating sockeye salmon, as they are known to adjust swimming behaviour by seeking out optimal temperature and flow conditions within the water column (Hinch and Rand 2000; Goniea et al. 2006; Salinger and Anderson 2006).

We identified several variables that help explain the variability in historic Fraser River temperature and discharge trends. However, our analysis was not completely comprehensive, and several additional environmental variables are known to contribute to observed river conditions (e.g. summer precipitation, percent cloud cover; Foreman et al. 2001). Model recommendations will vary depending on the suite of predictor variables examined, the length of the forecast, the response variable (i.e. temperature or discharge) and the relative trade-off between model fit, forecast precision and model availability. Ultimately, researchers may want to evaluate models on a case-by-case basis and select the most appropriate set of predictor variables that match their forecasting needs. A more extensive retrospective analysis evaluating historic model performance would also assist with the selection of a robust long range forecasting model. Future research should include the exploration of additional models, particularly multiple regressions, and a more detailed evaluation of historic measurement error and uncertainty in run timing forecasts. Researchers need to acknowledge and communicate the limitations of long range environmental forecasts, and quantify these uncertainties where possible.

ACKNOWLEDGEMENTS

Funding for this report was provided by the Pacific Salmon Treaty Research Fund, the Southern Endowment Fund (Project: Improvements to Environmental Management Adjustments), and the Fraser River Environmental Watch Program (http://www-sci.pac.dfo-mpo.gc.ca/fwh/index_e.htm) at DFO. For providing data and valuable input, we thank Mike Lapointe and Ian Guthrie from the Pacific Salmon Commission, John Morrison and Steve Macdonald from DFO, and Allan Chapman of the BC provincial River Forecast center for providing us ensemble flow forecasts. We would also like to thank Ian Guthrie, Herb Herunter and Gildas Toullec for reviewing earlier versions of this report and providing helpful comments.

EXECUTIVE SUMMARY

- **Purpose:** evaluate different long range forecast models for predicting Fraser River summer water temperature and discharge
- **Report objectives:**
  - o evaluate a series of long range forecasting models for summer temperature and discharge conditions in the lower Fraser River basin
  - o determine whether these methods can be appropriately fit using simple linear regression techniques
  - o quantify model and predictor variable uncertainty, and the impact of these uncertainties on the precision of forecasts
- Adult run timing for Fraser sockeye salmon are used as a specific case study
- Relationship between 19-day mean river temperature and discharge and several different predictor variables were explored:
  - o historic trends
  - o south BC mountains winter precipitation index
  - o snowpack water volume equivalent
  - o ensemble flow model
  - o south BC mountains summer air temperature anomalies
  - o south BC mountains spring air temperature anomalies

- **Result summary:**
    - most models could be fit using simple linear regressions
    - model parameters and performance varied over the course of the season
    - incorporation of both model and predictor variable uncertainty produced bootstrapped confidence intervals that were wider than the prediction intervals produced by the model
    - different temperature forecast models produced comparable results
    - multiple regression models (two variables) explained more of the historic variation in water temperature than single variable models
    - performance of historic trend – snowpack water volume multiple regression model for predicting temperature decreased over the course of the season; but mean and precision of forecasted values remained relatively constant during mid-August
    - ensemble flow models produced the most precise discharge predictions
    - ensemble flow model performance declined consistently over the summer
    - regression r^2 values increased as the number of days used to calculate mean temperature/discharge increased; changing from a 19-day to 31-day mean produced only minor improvements in model fit
- **Key conclusions:**
    - there are trade-offs between the timing of data availability, fit of the historic data, and uncertainty in the predictor variable
    - ignoring the uncertainty in environmental data could lead to overconfidence in results derived from temperature or discharge forecasts
- **Future research:**
    - explore other multiple regression options
    - quantify historic measurement error

LITERATURE CITED

Blackbourn, D.J. 1987. Sea surface temperature and pre-season prediction of return timing in Fraser River sockeye salmon (*Oncorhynchus nerka*). Can. Spec. Pub. Fish. Aquat. Sc. 96: 296 – 306.

Brett, J.R. 1995. Energetics. In Physiological ecology of Pacific salmon. *In:* Physiological ecology of Pacific salmon. *Eds.* Groot, C., Margolis, L, and Clarke W.C. UBC Press. Vancouver. Pp. 1-68.

Chernick, M.R. 1999. Bootstrap Methods: A Practitioner's Guide. John Wiley and Sons, Inc., Toronto.

Cooke, S.J., Hinch, S.G., Farrell, A.P., Lapointe, M.F., Jones, S.R.M., Macdonald, J.S., Patterson, D.A., Healey, M.C., and Van Der Kraak, G. 2004. Abnormal migration timing and high en route mortality of sockeye salmon in the Fraser River, BC Fish. Vol 29 (2): 22 - 33.

Fagerlund, U.H.M., McBride, J.R. and Williams, I.V. 1995. Stress and tolerance. *In:* Physiological ecology of Pacific salmon. *Eds.* Groot, C., Margolis, L, and Clarke W.C. UBC Press. Vancouver. Pp.461-503.

Foreman, M.G.G., Lee, D.K., Morrison, J., Macdonald, S., Barnes, D., and Williams, I.V. 2001. Simulations and retrospective analyses of Fraser watershed flows and temperatures. Atmosphere-Ocean. Vol 39 (2): 89-105.

Goniea, T.M., Keefer, M.L., Bjornn, T.C., Peery, C.A., Bennett, D.H., and Stuehrenberg, L.C. 2006. Behavioural thermoregulation and slowed migration by adult fall Chinook salmon in response to high Columbia River water temperatures. Trans. Am. Fish. Soc. Vol 135: 408-419.

Hinch, S.G., and Rand, P.S. 2000. Optimal swimming speeds and forward assisted propulsion: energy-conserving behaviours of upriver-migration adult salmon. Can. Jour. Fish. Aquat. Sc. Vol 57: 2470-2478.

Lapointe, M., Cooke, S.J., Hinch, S.G., Farrell, A.P., Jones, S., Macdonald, S., Patterson, D., Healey, M.C., and Van Der Kraak, G. 2003. Late-run sockeye salmon in the Fraser River, British Columbia, are experiencing early upstream migration and unusually high rates of mortality – what is going on? Proceedings of the 2003 Georgia Basin/Puget Sound Research Conference, Vancouver, B.C.

Lee, C.G., Farrell, A.P., Lotto, A., MacNutt, M.J., Hinch, S.G., and Healey, M.C. 2003. The effect of temperature swimming performance and oxygen consumption in adult sockeye (*Oncorhynchus nerka*) and coho (*O. kisutch*) salmon stocks. Jour. Exper. Bio. Vol (206): 3239-3251.

Macdonald, J.S. 2000. Mortality during the migration of Fraser River sockeye salmon (*Oncorhynchus nerka*): a study of the effect of ocean and river environmental conditions in 1997. Can. Tech. Rep. Fish. Aquat. Sc. 2315: 120 p.

Macdonald, J.S., Foreman, M.G.G, Farrell, T., Williams, I.V., Grout, J., Cass, A., Woodey, J.C., Enzenhofer, H., Clarke, W.C., Houtman, R., Donaldson, E.M., and Barnes, D. 2000. The influence of extreme water temperatures on migrating Fraser River sockeye salmon (*Oncorhynchus nerka*) during the 1998 spawning season. Can. Tech. Rep. Fish. Aquat. Sc. 2326: 117 p.

Maindonald, J., and Braun, J. 2003. Data Analysis and Graphics Using R: An Example-based Approach. Cambridge University Press, NY, NY.

Moore, R.D. 2006. Stream temperature patterns in British Columbia, Canada, based on routine spot measurements. Can. Water Res. Jourl. Vol 31 (1): 41-56.

Morrison, J., Quick, M.C., and Foreman, M.G.C. 2002. Climate change in the Fraser River watershed: flow and temperature predictions. Jour. Hydro. Vol 263: 230-244.

Morrison, J. 2005. Fraser River temperature and discharge forecasting: 2004 review. Can. Tech. Rep. Fish. Aquat. Sc. 2594: iv + 16 p.

Naughton, G.P., Caudill, C.C., Keefer, M.L., Bjornn, T.C., Stuehrenberg, L.C., and Peery, C.A. 2005. Late-season mortality during migration of radio-tagged adult sockeye salmon (*Oncorhynchus nerka*) in the Columbia River. Can. Jour. Fish. Aquat. Sc. Vol 62: 30-47.

Patterson, D.A., Skibo, K.M., Barnes, D.P., Hills, J.A., and Macdonald, J.S.. 2007a. The influence of water temperature on time to surface for adult sockeye salmon carcasses and the limitations in estimating salmon carcasses in Fraser River, British Columbia. N. Am. Jour. Fish. Mgmt. Vol 27: 878 – 884.

Patterson, D.A., Macdonald, J.S., Skibo, K.M., Barnes, D., Guthrie, I., and Hills, J.A. 2007b. Reconstructing the summer thermal history for the lower Fraser River, 1941 to 2006, and implications for adult sockeye salmon (*Oncorhynchus nerka*) spawning migration. Can. Tech. Rep. Fish. Aquat. Sc. 2724: 33 p.

Pyper, B.J., and Peterman, R.M. 1998. Comparison of methods to account for autocorrelation in correlation analyses of fish data. Can. Jour. Fish. Aquat. Sc. Vol 55: 2127-2140.

Quinn, T.P., Hodgson, S., and Peven, C. 1997. Temperature, flow, and the migration of adult sockeye salmon (*Oncorhynchus nerka*) in the Columbia River. Can. Jour. Fish. Aquat. Sc. Vol 54: 1349-1360.

Rand, P.S. and Hinch, S.G. 1998. Swim speeds and every use of upriver-migration sockeye salmon (*Oncorhynchus nerka*): simulating metabolic power and assessing risk of energy depletion. Can. Jour. Fish. Aquat. Sc. Vol 55: 1832-1841.

Salinger, D.H. and Anderson, J.J. 2006. Effects of water temperature and flow on adult salmon migration swim speed and delay. Trans. Am. Fish. Soc. Vol 135(1): 188-199.

Stine, R.A. 1985. Bootstrap prediction intervals for regression. Jour. Am. Stat. Asso. Vol 80 (392): 1026-1031.

Thomson, R., Ingraham, W.J., Healey, M.C., LeBlond, P.H., Groot, C., and Healey, C.G. 1994. Computer simulations of the influence of ocean currents on Fraser River sockeye salmon (*Oncorhynchus nerka*) return times. Can. Jour. Fish. Aquat. Sc. 51: 441 – 449.

Venables, W.N., and Ripley, B.D. 2002. Modern Applied Statistics with S 4th Ed. Springer-Verlag New York Inc., NY, NY.

Wagner, G.N., Hinch, S.G., Kuchel, L.J., Lotto, A., Jones, S.R.M., Patterson, D.A., Macdonald, J.S., Van Der Kraak, G., Shrimpton, M., English, K.K., Larsson, S., Cooke, S.J., Healey, M.C., and Farrell, A.P. 2005. Metabolic rates and swimming performance of adult Fraser River sockeye salmon (*Oncorhynchus nerka*) after a controlled infection with *Parvicapsula minibicornis*. Can. Jour. Fish. Aquat. Sc. Vol 62: 2124-2133.

Woodey, J.C. 1987. In-season management of Fraser River sockeye salmon (*Oncorhynchus nerka*): meeting multiple objectives. *In*: Sockeye salmon (*Oncorhynchus nerka*) population biology and future management. Eds. Smith, H.D., Margolis, L., and Wood, C.C. Can. Sp. Pub. Fish. Aquat. Sc. 96, pp. 367-374.

Zar, J.H. 1996. Biostatistical Analysis 3rd Ed. Prentice Hall, Inc., Upper Saddle River, New Jersey.