USING ENSEMBLE STREAMFLOW PREDICTION FOR WATER SUPPLY FORECASTING IN BRITISH COLUMBIA - PRACTICALITIES AND PITFALLS

Stephanie Smith and Frank Weber*

ABSTRACT

BC Hydro has just completed new watershed model calibrations for 20 of its reservoirs, which would enable the use of Ensemble Streamflow Prediction (ESP) techniques for preparing water supply forecasts. ESP methods have been used to forecast the seasonal water supply for BC Hydro's large interior snow-driven basins for a number of years with much success. But using the technique for smaller coastal rain- and rain/snow-driven basins has proven to be much more challenging. This paper will focus on two known issues: snowpack simulation in the model and month-to-month model calibration bias. This study showed that in coastal and interior basins, snowpack adjustments improved the forecast accuracy and slightly decreased the dispersion. The bias correction technique computed daily bias correction offsets as differences between simulated and observed flows during the calibration period. The offsets were then applied to the ensembles of the forecast. The technique produced mixed results.

INTRODUCTION

As one of the largest electric utilities in Canada, BC Hydro provides approximately 94% of British Columbia's population with electrical energy (47,500 GWh/year), 90% of which is hydroelectric energy. To help operate BC Hydro's reservoirs, assess resource capabilities, and determine pricing for the sale of energy outside of BC, water supply forecasts are issued between January and August. Water supply forecasts estimate the volume of water that is expected to flow into reservoirs during the February to September period.

BC Hydro's forecast team currently uses the UBC watershed model (Quick 1995) to forecast the seasonal inflow into four major reservoirs located on the main stems of the Columbia and Peace Rivers. It is planned to expand seasonal UBC watershed model forecasts to coastal basins in the near future. This study attempts to shed some light on known issues with seasonal water supply forecasting. The use of an updating technique for simulated snowpack in an interior basin and in a coastal basin with a relatively high percentage of rainfall-runoff and spatially distant snow course observations is evaluated. Furthermore, the effects of monthly biases in the calibration on water supply forecasts are investigated and a post-modeling correction procedure is tested.

The UBC watershed model is a conceptual, continuous hydrologic simulation model, developed initially in the mid-60's to calculate streamflow from mountainous watersheds. The UBC watershed model is a semi-distributed model. Runoff is not calculated for the watershed as a whole, but is calculated separately for predetermined, lumped elevation bands and then linearly combined to obtain total runoff. For a given watershed, the model simulates the various components of runoff using daily precipitation and daily minimum and maximum temperatures as input. Runoff components include surface runoff from rainfall, snowmelt, and glacier melt, interflow, and upper and lower groundwater flow. Together, these components represent the observed and expected inflow hydrograph for the reservoir. The model also simulates basin state conditions, such as snow water equivalent (SWE) and snow cover.

The UBC watershed model is calibrated for each basin using historical input and output data. The criteria for calibration of the basins are to match the annual runoff volume and then to match runoff peaks. Although calibrations are typically unbiased on an annual time frame, for some watersheds, some month-to-month bias remains after the calibration process is complete. Model bias is caused by systematic errors, such as timing errors in matching the model-simulated flows to the observed flows in the calibration process.

The ability to forecast the seasonal runoff lies in the fact that runoff from melting of the mountain snowpack is a major component of the seasonal water supply for many BC Hydro reservoirs. Consequently, water supply

---

Paper presented Western Snow Conference 2003

* Senior Hydrologists, BC Hydro, 6911 Southpoint Drive, Burnaby, British Columbia, V3N 4X8, Canada (Stephanie.Smith@BCHydro.com)

forecasts are more accurate for basins in which the greater proportion of seasonal runoff is attributable to snowmelt. The most accurate water supply forecasts are issued at the end of the accumulation season, when the snowpack storage is known best. Conversely, seasonal forecasts are relatively uncertain if the runoff volume is composed mainly of rainfall-runoff during the forecast period or from glacier melt, which responds to summer temperatures. For example, in the Columbia Region of British Columbia (BC), approximately 70% of annual runoff is comprised of snowmelt, while snowmelt contributes approximately 30-40% of annual runoff to south-coastal basins. Therefore, typically water supply forecasts for the Columbia Region are more accurate than for south-coastal basins.

In the Pacific Northwest, meteorological forecasts are more accurate than climatological averages for forecast lead times of only a few days. Therefore, the Ensemble Streamflow Prediction (ESP) procedure uses sequences of historic climatologic data as future UBC watershed model input. The ESP procedure assumes that meteorological events that occurred in the past are representative of events that may occur in the future. Each year of historical meteorological data is assumed to be a possible representation of the future and is used to simulate a streamflow trace (http://www.srh.noaa.gov/lmrfc/ahps/esp.shtml). In a first step, the UBC watershed model is run up to the forecast date. The purpose of the tracking run is to create the correct basin state conditions in the model at the start of the forecast period. The current simulated basin-state conditions, such as snowpack, soil moisture, and groundwater conditions are the conditions from which the ESP run will start. Unweighted historic weather sequences are then used as model input to simulate the runoff that would have occurred in these years given the current basin-state conditions. The model produces a number of possible future inflow hydrographs that are used directly for follow-up planning studies or are statistically analyzed to produce a volumetric forecast. At BC Hydro, the mean of the ensemble of forecasts is typically used.

The uncertainty in the forecast is due to modeling error and the uncertainty in the future weather. Unknown weather conditions during the forecast period are generally the largest source of forecasting error, as historical weather sequences are all assumed to be equally likely in the future. Model uncertainty arises because hydrologic processes and their spatial variability are too complex to be fully represented by a semi-distributed mathematical model. Model uncertainty is closely related to parameter uncertainty. Model parameters are estimated from limited historical samples. For example, some basins were calibrated with only 10-15 years of input data, and, therefore, possible future conditions may not be well represented by the short historic data record used for calibration. Modeling error is also caused by input uncertainties. Model inputs are based on a relatively small number of stations in and around the watershed. This may result in an incorrect representation of basin input parameters. Additionally, instrument errors cause the measurement devices to not perfectly record the variable they are meant to measure (e.g. undercatch, snowcapping). Last but not least, the calculated reservoir inflows used for calibration and to check the simulation may differ from the true reservoir inflows (output uncertainties).

The accurate simulation of the basin snowpack is critical for an accurate water supply forecast. However, due to model, parameter, and input uncertainties, as well as instrument errors, the simulated snowpack can differ from the observed snowpack. The model snowpack is calculated from temperature and precipitation data that are extrapolated from the station elevations to other elevation bands using a computed lapse rate. Spot measurements of snow water equivalent at snow courses and snow pillow stations may not be representative of the basin wide snowpack, although efforts are made to select representative locations on gently sloping, small, and protected mountain meadows and north-facing slopes. For water supply forecasts, it is assumed that the measured snowpack is typically a better representation of the snow conditions in the basin than the snowpack modeled using climate inputs.

To correct the simulated model snowpack at the forecast date, the model can either be rerun using modified input data or the observed SWE data can be used to estimate what the simulated SWE should be at the forecast date. Since it is typically a very time consuming and subjective process to assess when and by how much input parameters need to be adjusted, it is preferred to adjust the SWE at the forecast date. Assuming that a number of spatially distributed snow course and snow pillow measurements would satisfactorily represent the snowpack at different elevations, a technique has been developed to utilize snow course and snow pillow measurements to update the UBC watershed model estimates of snowpack (Assaf et al. 1992). To develop the snowpack adjustment tool, the model is run for each year available from the beginning of the snow accumulation season, i.e. from October 1, to a forecast date during the accumulation period for which snow course measurements are available, i.e.

February 1, March 1, April 1, and May 1. Using the simulated snowpack for the forecast date the model was then run to the end of that snowmelt season, i.e. September 30. If the observed and simulated hydrographs were in good agreement – as determined by the cumulative simulation error and visual analysis - no adjustment was made to the simulated snowpack for the forecast date. However, if the simulated hydrograph diverged from the observed one, adjustments were made to the simulated snowpack at the forecast date until the simulated hydrograph satisfactorily matched the observed one. Subsequently, regression relationships were developed between the optimized simulated snowpack and the observed snow course and snow pillow measurements. These linear models were developed for each snow course and snow pillow for the beginning of each month between February and May. Since the updating procedure assumes that runoff volume errors are proportional to errors in snowpack estimation, errors in the snowpack simulation can give an early indication of errors in seasonal forecasts. The linear models are then operationally used to calculate the optimal simulated snowpack at the forecast date from the observed data. The current year's simulated snowpack at the forecast date is adjusted using a single percentage adjustment factor across all elevation bands until a satisfactory match between the optimal simulated and current year's simulated snowpack is reached.

The study sought to identify and address two main issues with using the ESP method for water supply forecasting – (1) snowpack simulation in the model and (2) correcting for known monthly biases in the calibration results.

THE METHOD

In order to explore some of the possible issues that can arise from the ESP forecasting method, the study focussed on two watersheds in different regions of BC. The two watersheds were chosen for their location and also for some known deficiencies in the model calibrations for these particular basins. The location of the basins is shown in **Figure 1**. Table 1 summarizes the basin characteristics for the Clowhom and Duncan basins.

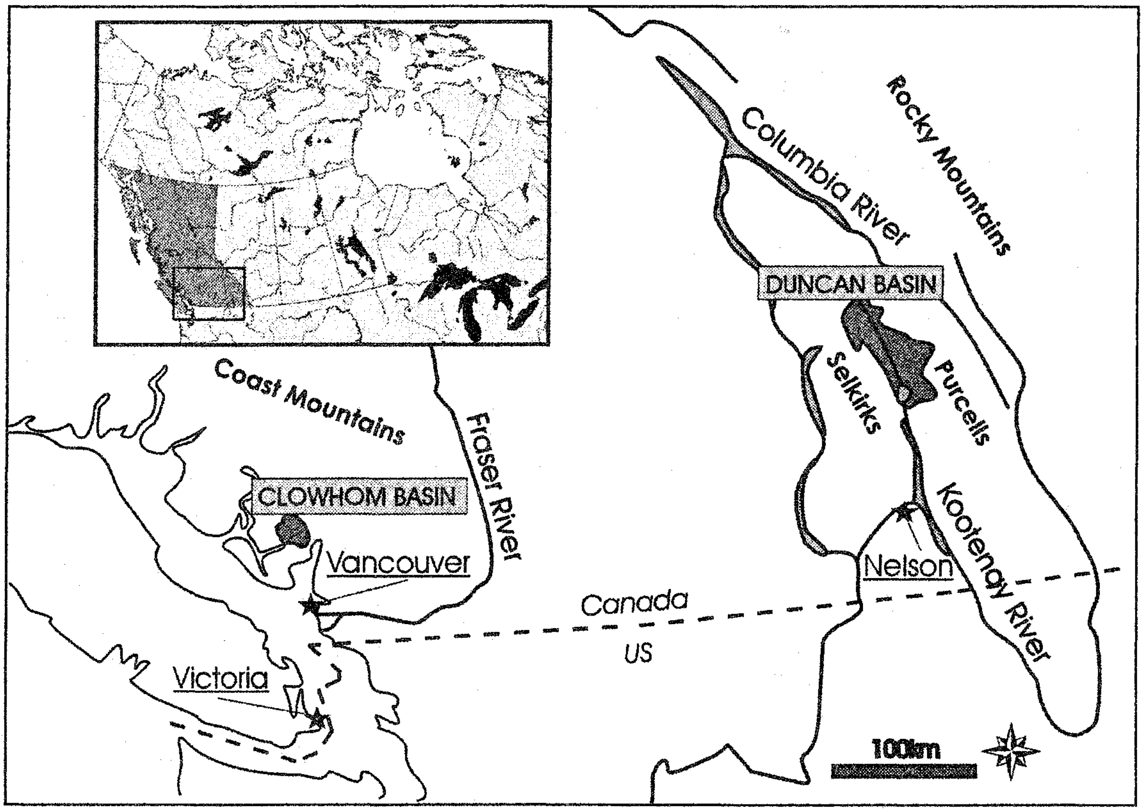


Figure 1. Location of the Clowhom and Duncan basins.

The Clowhom River basin is situated in the southwestern coastal mountains of BC. It is located approximately 55 km northwest of Vancouver and discharges into the Strait of Georgia. It has a basin area of 382 km^2 and ranges in elevation from 54 m to 2593 m. As a coastal watershed, the hydrology of the Clowhom basin is dominated by fall and winter storms coming off the Pacific, resulting in a largely rainfall-driven runoff. Approximately 30 percent of the annual inflow volume is from snowmelt. The Duncan watershed is a subbasin of the Kootenay River basin and is located in southeastern BC between the Selkirk and the Purcell mountain ranges. The basin area is 2443 km^2 and covers an elevation range of 550 m to 3170 m. The Duncan River below the dam flows south into Kootenay Lake. The hydrology for this interior watershed is approximately 70% snowmelt-driven. The climate of the region is generally dry throughout the summer months, with periodic heavy rainfall events. During the fall and winter, a deep snowpack accumulates. Approximately 4% of the basin is glaciated.

Table 1. Summary of basin characteristics.

	Clowhom	**Duncan**
Hydrologic regime	Pluvio-nival	Nivo-glacial
Snowmelt runoff	Approximately 30%	Approximately 70%
Drainage area (km^2)	382	2443
Minimum elevation (m)	54	547
Maximum elevation (m)	2593	3171
Reservoir area (ha)	800	7150
Reservoir storage (Mm^3)	105	1700

The Clowhom and Duncan watersheds were calibrated in the UBC watershed model using 12 years of historic meteorological and inflow data. Table 2 summarizes the model calibration results. The model performances were evaluated using several goodness-of-fit parameters. The coefficient of determination is a criterion for the shape and timing of the hydrograph (Quick 1995) with 1 indicating a perfect match. It is 0.77 and 0.87 for Clowhom and Duncan, respectively. To assess the model's volumetric and temporal accuracy, the coefficient of efficiency of Nash and Sutcliffe (1970) is used. It calculates the ratio of explained to total variance of observed and calculated flows (Micovic and Quick 1999). A perfect simulation yields a coefficient of 1. The Nash-Sutcliffe coefficients of efficiency for Clowhom and Duncan are 0.76 and 0.87, respectively. The coefficients indicate a good fit for the coastal basin and a very good fit for the interior basin. These numbers are typical of the two respective climates. Due to the flashy nature of coastal basins, it is more difficult to calibrate a basin with a pluvio-nival regime than a basin with a nivo-glacial regime.

Table 2. Summary of calibration statistics.

	Clowhom	**Duncan**
Calibration period	Oct '92 – Sep '99 (7 yrs.)	Oct 1987 - Sep 1999 (12 yrs.)
Verification period	Oct '87 – Sep '92 (5 yrs.)	Oct 1987 to Sep 1999 (12 yrs.)
Coefficient of efficiency	0.76	0.87
Coefficient of determination	0.77	0.87
Mean annual volume error (% average)	1%	-4%
Mean seasonal volume error (February – September % of average)	2%	-2%

Figure 2 shows the simulation error for the February through September runoff from the calibration of the Duncan watershed, (a) for each year and (b) sorted from smallest to largest error. The UBCWM was calibrated to match annual runoff volumes, and the error plots for the seasonal runoff show that the errors are normally distributed as would be expected, with an average seasonal runoff volume error of about –2% during the calibration. But the monthly distribution of simulation errors for Duncan indicates some problems with the calibration results.

Figure 2. Duncan Feb-Sep simulation error in Mm^3 (a) sorted by year and (b) sorted by volume error.

Figure 3(a) and (b) show the monthly volume errors in the simulation for the months of May and June for Duncan. Note that May is nearly always under-simulated, while June is nearly always over-simulated during the calibration.

April shows similar results to May and July is similar to June. This suggests that the timing of the modeled snowmelt-runoff is off. The seasonal volumes are essentially unbiased at the end of the season, but the modeled runoff for any given month is off. Operationally, a problem arises when producing a forecast in any of the problem months (April – July), as the runoff to the forecast date will be computed from the observed runoff, while

Figure 2. Duncan Feb-Sep simulation error in Mm^3 (a) sorted by year and (b) sorted by volume error.

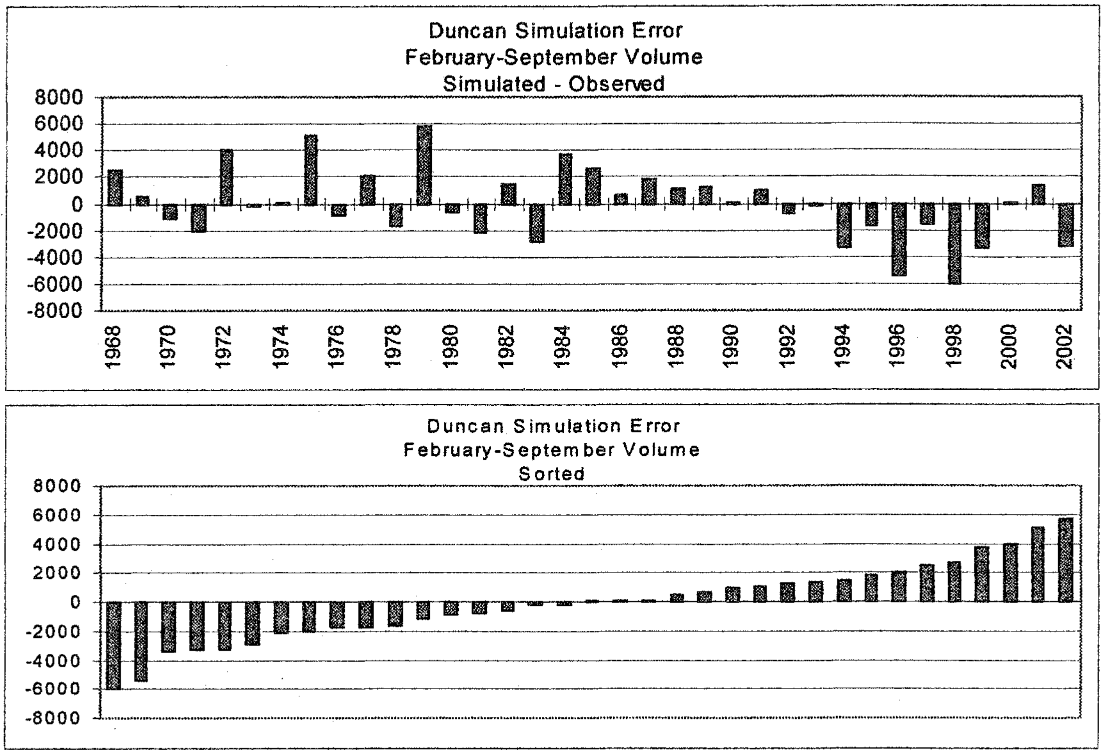
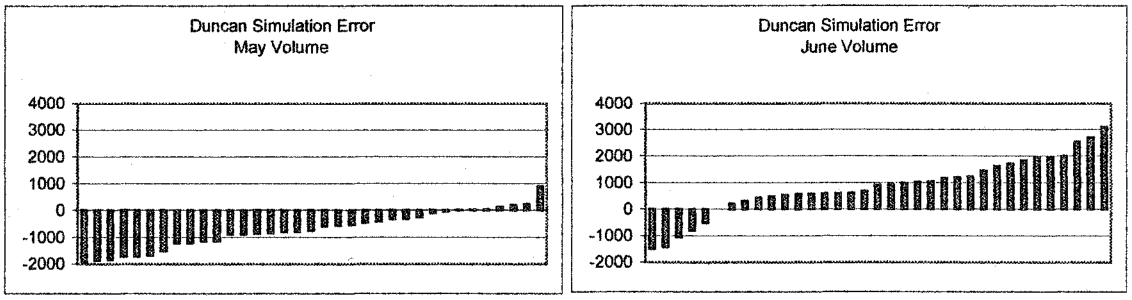


Figure 3. Duncan monthly volume simulation error in Mm^3 (a) for May and (b) for June.



the forecast will be generated from the model ensemble runs. When the next month's forecast is generated, the change in the seasonal forecast may be simply caused by the shift in the modeling bias, rather than due to real changes in the basin state conditions.

To explore the possible issues across a broad range of climatic conditions, the study generated ESP water supply forecasts for April 1st of each year from 1988 through 2002. The April 1st water supply forecast has historically been the most critical water supply forecast date in BC as the snowpack is near its peak by this date and is the best indicator of what runoff volume to expect for the spring freshet. Due to the limited number of years of data, the ESP runs used all available years of data (1988-2002) for each of the forecasts generated for April 1st for each year in the record, including the data for the year that was being forecast.

Generating a forecast for April 1 starts with a tracking run from October 1 of the previous year up to March 31. This creates the initial basin state conditions for the ESP runs. The simulated SWE is compared to the optimally simulated SWE, which is derived from the observed SWE via linear regression equations. Since the observed inflows are available for the tracking period, it is also possible to review the performance of the model to date prior to starting the ESP forecast. Operationally, both the simulated SWE and the tracking period would be used to assess model performance. To test the snowpack adjustment procedures, two forecasts were generated: one with no adjustments, and one, if necessary, with adjustments to the simulated snowpack at the forecast date. The ensemble forecasts were then generated from each of these initial basin state conditions.

To address the month-to-month bias in the calibration for Duncan, a very simple technique was applied to attempt to remove the bias from the model simulation. For each year of the calibration, the observed runoff and the corresponding simulated runoff from the model were generated. Subtracting the simulated from the observed value for any given day gave the model bias on that day for that year. This model bias was then applied to each of the forecast ensembles to "correct" the forecast value for each day of each ensemble and remove the bias from the forecast.

RESULTS AND DISCUSSION

Snowpack adjustment

Figure 4 and Figure 5 compare ESP water supply forecasts produced with and without snowpack adjustments for the Clowhom and Duncan basins, respectively. **Table 3** summarizes the average forecast error and RMSE of the April-September ESP water supply residual forecasts using unadjusted and adjusted snowpack.

For Clowhom, the 1988-2002 April-September runoff volume tended to be over-forecast. The positive bias of 24 Mm^3 was reduced to 7 Mm^3 by adjusting the simulated snowpack at the forecast date. The unadjusted and adjusted biases corresponded to 4% and 1% of the average 1988-2002 April-September runoff volume. In individual years, for example 1997, but in particular 1993, 1999, and 2001, snowpack updating decreased forecasting accuracy. For three of these years, updating resulted in even more extreme under-forecasts than with the unadjusted forecasts, which may indicate that in these years unusual weather patterns may have prevailed. The RMSE of the unadjusted and adjusted Clowhom forecast decreased slightly from 14% to 13% of the 1988-2002 April-September runoff volume. Overall, snowpack updating improved the forecasts.

The 1988-2002 Duncan April-September runoff volume tended to be under-simulated with a bias of -104 Mm^3. The snowpack updating procedure reduced the overall forecast bias to -13 Mm^3. The unadjusted and adjusted biases corresponded to –4% and 0% of the average 1988-2002 April-September runoff volume. In 1989, 1992, and in particular in 1993 and 2001, updating decreased forecasting accuracy. With snowpack updating, the RMSE of the forecast decreased from at 8% to 6% of the 1988-2002 April-September runoff volume. For Duncan, snowpack updating improved the forecast.

To assess whether extreme snowpack years affect the forecast error, representative April 1 snow course readings were correlated with the forecast error of the unadjusted model run. The Grouse Mountain (BC snow course no. 3A01 at 1100m asl) and Mount Templeman (BC snow course no. 2D09 at 1860 m asl) snow courses were selected as representative snow courses for the Clowhom and Duncan basins, respectively (Grouse Mountain $R^2 = 0.50$; Mount Templeman $R^2 = 0.67$). It was determined, that the forecast error is basically uncorrelated with SWE (Grouse Mountain $R^2 = 0.04$; Mount Templeman $R^2 = 0.02$). Consequently, extremely high or low snowpack years do not significantly affect forecast errors.

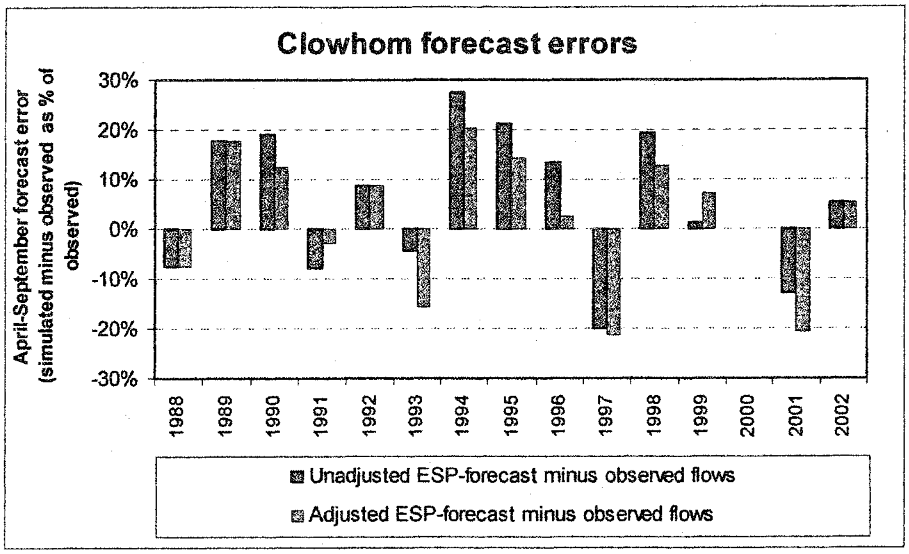


Figure 4. Comparison of Clowhom April-September ESP water supply forecast errors using unadjusted and adjusted snowpack.

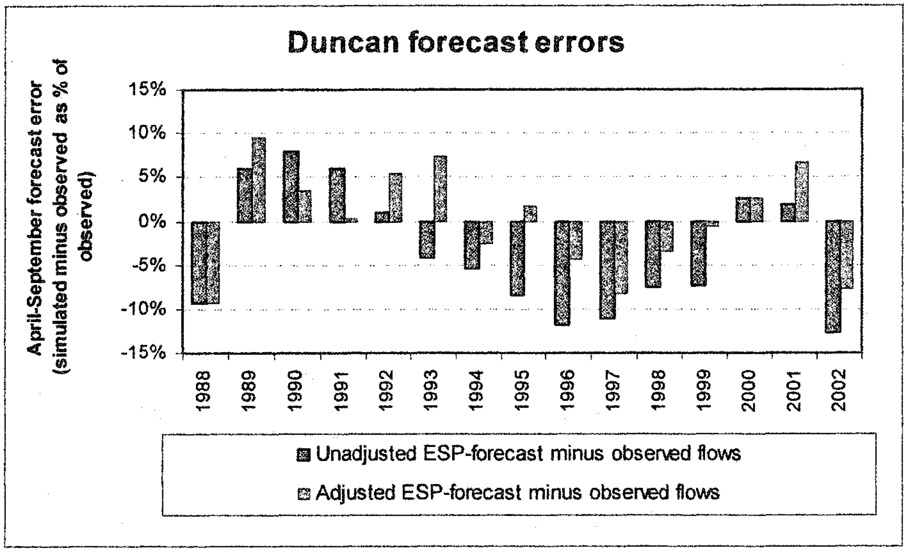


Figure 5. Comparison of Duncan April-September ESP water supply forecast errors using unadjusted and adjusted snowpack.

Table 3. Summary statistics of April-September ESP water supply forecasts using unadjusted and adjusted snowpack.

	Unadjusted April-September forecast minus observed		**Adjusted April-September forecast minus observed**		**RMSE of the unadjusted April-September forecast minus observed**		**RMSE of the adjusted April-September forecast minus observed**	
	Mm^3	% of 88-02	Mm^3	% of 88-02	Mm^3	% of 88-02	Mm^3	% of 88-02
Clowhom	24	**4**	7	**1**	97	**14**	92	**13**
Duncan	-104	**-4**	-13	**0**	227	**8**	155	**6**

Bias correction

To remove the calibration bias from the forecast results for the Duncan watershed, a bias correction was applied to each of the ensembles. The bias correction offsets were computed as the difference between the simulation and the observed flows during the calibration. They were applied to each day of each of the ensembles that made up the forecast. The corrected ensemble members were then analyzed to produce the seasonal runoff volume forecast. Figure 6(a) shows the mean, 90th percentile and 10th percentile of the forecast ensembles for the April 1994 forecast for Duncan before any corrections for bias were made. Figure 6(b) shows the resulting traces after the application of the bias correction to all years in the ensemble. Note that in the bias-adjusted forecast, the freshet starts earlier and the peak flow in June and July are significantly reduced. This seemed to compensate for the problem in the model calibration with under-simulation of the runoff early in the spring and over-simulation of the peak flows in June and July. Table 4 reports the April-September volume forecast results for the uncorrected and corrected ensembles. Note that there was little change in the mean seasonal volume forecast, but that the variability, reported as standard deviation, of the forecast ensembles increased from 6% to 9% of the April-September average with the bias correction.

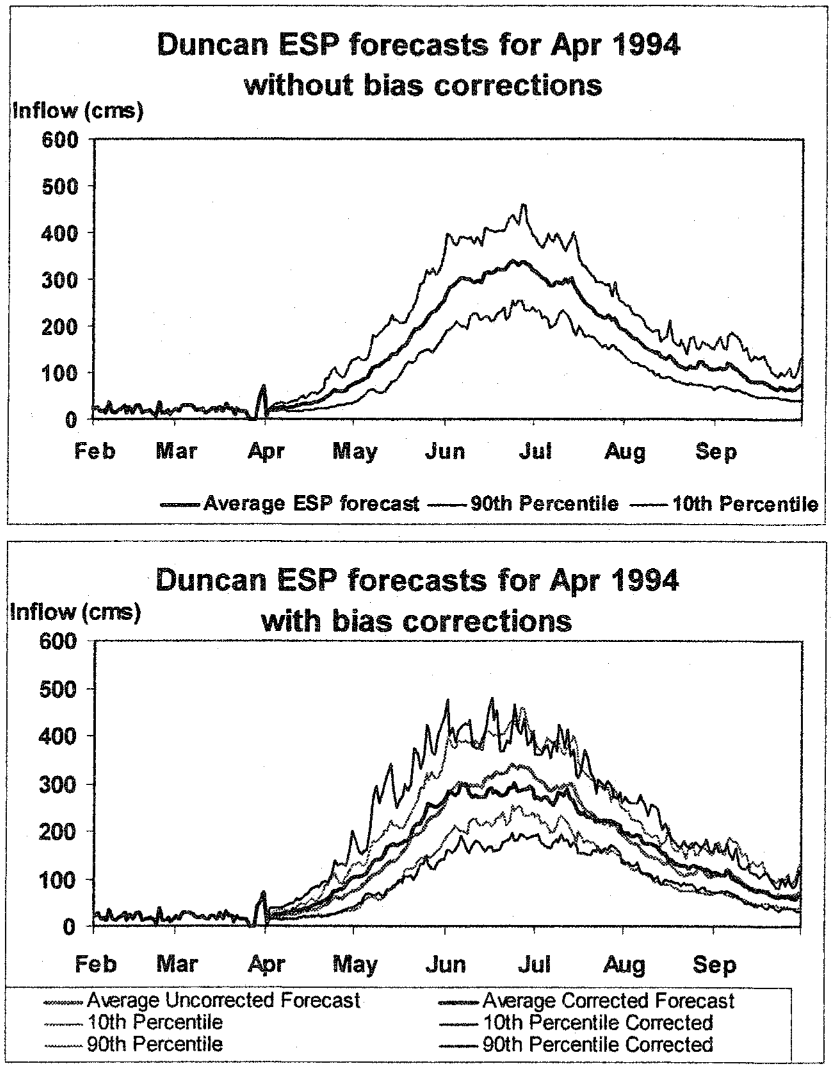


Figure 6. Comparison of April 1994 ESP forecast (a) without and (b) with bias correction.

Table 4. Comparison of Forecast Seasonal Runoff Using Bias Correction on Ensembles in ESP Forecast.

Duncan April –September 1994 Forecast	**Uncorrected**	**Bias-Corrected**
April – September Volume (% of average)	94.9	95.2
Standard deviation (% of average)	6	9

One of the problems with the bias correction method was that the method assumed that the errors in the model simulation were entirely due to systematic error in the model for a particular day during the calibration, when the error may have been due to both systematic and random errors unique to each year's simulation, such as poor snowpack simulation over the previous winter. In trying to compensate for systematic errors in the model simulation, the bias correction also removed the random error from the model simulation. Random errors vary from year to year due to temporary problems with parameter representation in the model. Given that the simulated snowpack was adjusted prior to producing the ensemble forecast, the bias corrections may have over-compensated for errors in the model that result from poor snowpack simulation. In some of the years tested, the bias correction produced negative daily flows in some of the ensemble members, which did not make physical sense, and it created complications in producing summary statistics for the forecast. Clearly, more research is required to refine the methods to compensate for the deficiencies in the calibrations for the Duncan watershed.

CONCLUSIONS AND RECOMMENDATIONS

Ensemble streamflow prediction can be a practical method for predicting seasonal runoff in British Columbia, but it is not without its pitfalls. It is most important to start with a good calibration of the conceptual model for the watershed, as all other errors will only compound if there are problems with the original calibration.

As expected, runoff forecasts are generally more accurate in the snowmelt-dominated basin than in the rainfall-snowmelt hybrid basin. In the coastal as well as in the interior basin, snowpack adjustments improved the forecast accuracy and slightly decreased the dispersion. For the interior basin, the snowpack adjustments improved the forecasts slightly more than on the coast. Updating of the simulated snowpack using snow course and snow pillow data should, therefore, be an integral part of ESP water supply forecasts for coastal and interior basins. Future research should investigate variations on the snowpack adjustment tool to allow for individual adjustments for different elevations as opposed to a single percentage adjustment across all elevation bands.

There does not appear to be a simple method to correct for month-to-month bias in the calibration after the fact when creating an ensemble forecast, at least for the Duncan watershed as presented in the study. Month-to-month bias causes inconsistencies in the water supply forecasts from one month to the next. These inconsistencies are difficult to explain and can be misleading to inexperienced or unwary users of the forecast product. A solution for this problem must be found before ESP forecasts can be used in the basins that are affected by month-to-month bias. Fortunately monthly bias is an issue only in a very small number of the BC Hydro watersheds calibrated.

LITERATURE CITED

Assaf, H., A. Loukas and M. Quick. 1995. Updating Seasonal Snowmelt Estimates Using Snowcourse Measurements. Western Snow Conference April 17-20, 1995 pp.21-28.

http://www.srh.noaa.gov/lmrfc/ahps/esp.shtml: accessed June 26, 2003.

Micovic, Z. and M. Quick. 1999. A rainfall and snowmelt runoff modelling approach to flow estimation at ungauged sites in British Columbia. Journal of Hydrology 226 (1999) pp.101-120.

Quick, M. 1995. U.B.C. Watershed Model Manual. Mountain Hydrology Group, Civil Engineering, University of British Columbia, Vancouver , Canada.

ACKNOWLEDGEMENTS

The authors want to thank Z. Micovic and E. Yu for providing the snowpack adjustment routines and K. Lee and Z. Micovic for calibrating the UBC watershed model for the Clowhom and Duncan basins. We would also like to thank E. Weiss for his guidance and B.Samson for his valuable comments.