Seasonal prediction with error estimation of the Columbia River streamflow in British Columbia

William W. Hsieh[*], Yuval, Jingyang Li

Department of Earth and Ocean Sciences, University of British Columbia,

Vancouver, B.C. V6T 1Z4, Canada

Amir Shabbar

Meteorological Services of Canada, Downsview, Ont. M3H 5T4, Canada

Stephanie Smith

B.C. Hydro, Burnaby, B.C. V3N 4X8, Canada

Submitted to Journal of Water Resources Planning and Management

December 12, 2002

[*]Corresponding author, email: whsieh@eos.ubc.ca, tel: 604-822-2821, fax: 604-822-6088

Abstract

Large-scale climatological states [tropical Pacific sea surface temperatures (SST), Pacific-North American (PNA) atmospheric teleconnection and Pacific Decadal Oscillation (PDO)] and local precipitation data are used to predict the April–August Columbia River streamflow at Donald, British Columbia, Canada. Using predictors up to the end of November in the preceding year, forecasts of the April–August streamflow were made by multiple linear regression (MLR) under a jackknife scheme. A correlation skill of 0.52 is attained using PDO, PNA and SST as predictors, with PDO being the strongest and SST the weakest. When local precipitation is added among the predictors, PDO becomes redundant, and MLR with precipitation, PNA and SST as predictors attained a correlation skill of 0.70. Feedforward neural network models were used for nonlinear regression, but the results were essentially identical to the MLR predictions, implying that the detectable relationships in the short, 49-sample record are linear. A bootstrap process estimates the relative errors of the MLR predictions.

1 Introduction

About 40% of the hydro-electric power produced in British Columbia, Canada, is generated in the Columbia River basin. Advanced predictions of the volume of flow in the river and its tributaries, and assessments of their probability, are important for making decisions related to the economical management of the water system, and for environmental consideration.

The issue of streamflow predictability was dealt with in many papers [e.g., Redmond and Koch, 1991; Garen, 1992; Nijssen et al., 1997; Cayan et al., 1999]. It is currently well established [Redmond and Koch, 1991; Cayan et al. 1999] that large scale climatological states, especially the El Niño–Southern Oscillation (ENSO) phenomenon, significantly influence the streamflow in rivers in the Pacific Northwest. In particular, Hamlet and Lettenmaier [1999], using a macroscale hydrology model driven by ENSO and the Pacific Decadal Oscillation (PDO) concluded that the climatological attributes have strong impacts, and can be effectively used, for the prediction of the Columbia River streamflow above The Dalles, Oregon. Hsieh and Tang [2001] found that ENSO and the Pacific–North American (PNA) atmospheric teleconnection pattern influences the interannual variability of accumulated snow in the Columbia basin in British Columbia.

This study examines the role of ENSO, PNA and PDO indices, in addition to local precipitation, as long-range predictors of the upper Columbia River streamflow close to its source. The predictions are for the streamflow at Donald, British Columbia, in the April–August period, which consists of about 78% of the annual flow. Reliable data to carry out this study exist only for the last 50 years. Development of a statistical prediction model using such a short record is expected to include large errors, thus a bootstrap scheme [Efron and Tibshirani, 1993] was employed to estimate those errors and give some probabilistic assessment of the predictions.

2 Data

Sea Surface Temperature Anomalies (SSTA) in the tropical Pacific were chosen to represent the ENSO climatological state. The NOAA monthly SST fields [Smith et al., 1996] from 1950 to 1999 were regridded, smoothed and subjected to Principal Component Analysis (PCA) analysis. The first six principal components (PCs) were retained as candidate predictors. The mid–troposphere state over the Pacific Ocean and the North American continent is represented by the PNA index, which is the standardized amplitude from a rotated PCA of the 700 mb height anomalies [Barnston and Livezey, 1987]. The PDO index used is the leading PC of the monthly SST anomalies in the North Pacific Ocean, provided by Nathan Mantua, University of Washington. Local precipitation data came from the Meteorological Services of Canada. They include 72 gridded precipitation records in the area between 49°N to 52°N and 114°W to 121°W. A PCA analysis was carried out and the first six PCs were retained.

The April–August upper Columbia River total flow was calculated from the daily streamflow records of the B. C. Hydro station at Donald (51° 29′N, 117° 10′W).

3 Prediction methods

Multiple Linear Regression (MLR) and feedforward Neural Networks (NN) were considered as prediction models. The NN training was automatically controlled by the Generalized Cross Validation approach of Yuval [2000]. It was found that in spite of its nonlinear capability, the NN has no advantage over the MLR for this prediction problem involving a short record of only 49 samples. The NN model parameters degenerated into the MLR coefficients and the predictions were essentially identical. This points out that the detectable relationships between the predictors and the predictands are linear, although the possibility to detect nonlinearity in similar relationships using longer data sets cannot be ruled out. Hence the more straightforward and economical MLR was chosen for the following study.

The first six PCs of the SSTA data, the PNA index, the PDO index and the first six PCs of the local precipitation data at various lead times were considered as possible predictors of the April–August streamflow. The time series were standardized and their statistical significance as predictors was tested using stepwise regression [Wilks 1995]. Only the predictors with significant contributions were retained for the purpose of actual prediction. Among the six PCs of the SSTA, only the first PC, representing ENSO, is retained in the stepwise MLR— this first PC will be referred to as SST1.

The final performance of the prediction models was tested by a Leave-One-Out (Jackknife) cross-testing, where the datum at each year was set aside in turn, then a prediction model was developed using the rest of the data set. The predicted values for the left-out years were collected together, yielding a full record for testing predictions.

4 Results

To test long-range forecasting, we only used predictors up to the end of November to predict the April–August streamflow. Table 1 shows the various predictors and the cross-tested streamflow prediction skills from stepwise MLR. It is clear that among the three large-scale climate indices in November, PDO has the strongest apparent influence, followed by PNA, then SST1. Using only the November PNA and PDO indices as predictors yielded a forecast correlation skill of 0.527. In fact, adding SST1 as an extra predictor besides PNA and PDO lead to a marginal decline in skill.

However, when the local precipitation PCs are added, then PDO is discarded as redundant by stepwise regression, and the strongest apparent influence comes from the fourth PC of the local precipitation in October (Prec4). With November SST1, PNA, and the October Prec4 as predictors, the highest forecast correlation skill of 0.699 was attained (Fig.1). Prec4 is correlated with PDO at -0.493, hence Prec4 contains local PDO information. Adding PDO as an extra predictor, actually lowers the cross-tested correlation skill (Table 1). The selection of the fourth precipitation mode, explaining only 5% of the total variance, is intriguing, though one must bear in mind that the leading PCs may have been eliminated during stepwise regression as they contain similar information as the PNA and SST1 predictors. Also the October Prec4 did better than the November Prec4. This may be because that the November precipitation data are noisier than the October data due to the advent of winter storms, hence detrimental to long-term signals such as the local PDO contained in the precipitation data. Persistence forecast is very poor, yielding only a correlation of 0.16.

5 Bootstrap error estimation

The bootstrap error estimation process is based on the idea of bootstrap resampling of the data [Davison and Hinkley ,1997; Efron and Tibshirani,1993]. Blocks of data 3 years in length are sampled from the original record and are assembled together to form a new training dataset equal in length to the original one. This is repeated B times to give B new training datasets. An MLR model is developed from each of these new datasets and used to predict replicas P_i^*, $(i = 1, 2, \cdots, B)$ of each original prediction P.

The set of differences $P_i^* - P$, $(i = 1, 2, \cdots, B)$ is called here a bootstrap deviation set. Distribution of the deviations is not always symmetrical and thus we consider separately $P_i^{*+} - P$ and

$P_i^{*-} - P$, the positive and negative deviation sets. Following Yuval [2001], the statistic considered for the error estimation is the square root of the means of squared bootstrap deviations, i.e.

$$\sigma^+ = \left(\frac{1}{B^+} \sum_{i=1}^{B^+} (P_i^{*+} - P)^2 \right)^{1/2} ; \quad \sigma^- = \left(\frac{1}{B^-} \sum_{i=1}^{B^-} (P_i^{*-} - P)^2 \right)^{1/2} , \tag{1}$$

where B^+ and B^- are the corresponding sizes of the sets such that $B^+ + B^- = B$.

The values of σ^+ and σ^- give a measure of the relative accuracy expected for each prediction. Error bars can be scaled by a factor Q, which is arbitrary and should be chosen to achieve a certain goal. Motivated by economical water management considerations, a reasonable goal can be minimizing the cost function ϕ with respect to the parameter Q:

$$\phi = \sum_{j=1}^{N-H} \left| O_j - (P_j + Q\sigma_j^{\ddagger}) \right| + \alpha \sum_{i=1}^{N} Q \left(\sigma_i^+ + \sigma_i^- \right) \ , \tag{2}$$

where O is an observed predictand value, N is the number of data points in the dataset, H is the number of 'hits' which are the cases where $P - Q\sigma^- \leq O \leq P + Q\sigma^+$, (the hit cases being excluded from the first summation), $\sigma^{\ddagger}$ is $-\sigma^-$ if $O < P - Q\sigma^-$, or is σ^+ if $O > P + Q\sigma^+$, and α weighs the relative importance between the two terms on the right hand side.

The first component of ϕ is the sum of distances between error limits and the observations exceeding them. It quantifies failures by penalizing against error bars too small to provide adequate safeguards. The second component is the sum of error bars lengths in the whole testing set. This component measures the lack of confidence in the predictions. The relative weight of each of these components is determined by the parameter α. What that relative weight should be depends on the tolerance of failures, and the permissible level of uncertainty. A cautious manager tends to require small number of failures and thus a small value for α. This results in large error bars and limits the flexibility in the process of decision making. A more risky approach permits larger values of α, resulting in smaller error bars, more maneuvering space in the management of the water resources, but higher chances for failures. It is for the users of the prediction to select a value of α which meets their economical and environmental constraints.

The bootstrap error bars in Fig. 1 were produced with $B = 500$ and $\alpha = 0.1$, yielding $Q = 5.75$, and a mean error bar length of 1.05×10^9 m^3— about 24% of the mean seasonal streamflow—which includes 37 of the 49 observations inside the error bar limits.

6 Summary and Conclusions

Large-scale climatological states [tropical Pacific sea surface temperatures (SST), Pacific-North American (PNA) atmospheric teleconnection and Pacific Decadal Oscillation (PDO)] and local precipitation data were used for long-range forecasting of the April–August Columbia River streamflow at Donald, British Columbia, Canada, by multiple linear regression. When only the three large-scale climatological states were used as predictors, a correlation skill of 0.52 was attained under a jackknife scheme, with the PDO having the strongest regression coefficient, the PNA, the second strongest, and the tropical Pacific SST (representing the El Niño–Southern Oscillation), the weakest coefficient. In fact, leaving out the SST index results in marginally higher skills. When principal components of local precipitation is added among the predictors, PDO becomes redundant, and MLR with precipitation, PNA and SST attained a correlation skill of 0.70. Since none of the predictors used were beyond the end of November in the preceding year, potentially useful predictions of the Columbia River streamflow in British Columbia can be made 4 months before the April–August period.

A bootstrap scheme is used to estimate the prediction errors. The errors are scaled such that they minimize a cost function that combine measures of lack of confidence in the predictions, and their failures. The relative weighting of these two components must be decided by a user of the predictions. Feedforward neural network models were also used for nonlinear regression, but the results were essentially identical to the MLR predictions, implying that the detectable relationships in the short, 49-sample record are linear.

Acknowledgments

We benefitted from helpful discussions with B.C. Hydro staff, Brian Fast and Eric Weiss. This work was supported by research and strategic grants to William Hsieh from the Natural Sciences and Engineering Research Council of Canada, and a contract from B.C. Hydro.

References

Barnston, A. G., and Livezey, R. E. (1987). "Classification, seasonality and persistence of low–frequency circulation patterns." *Mon. Weather Rev.,* 115, 1083-1126.

Cayan, D. R., Redmond, K. T., and Riddle, L. G. (1999). "ENSO and hydrological extremes in the western United States." *J. of Climate,* 12, 2881-2893.

Davison, A. C., and Hinkley, D. V. (1997). *Bootstrap Methods and their Application*, Cambridge University Press, Cambridge, 582 pp.

Efron, B., and R. J. Tibshirani, (1993). *An Introduction to the Bootstrap*, Chapman & Hall, New York, 436 pp.

Garen, D. C. (1992). "Improved techniques in regression-based streamflow volume forecasting." *J. Water Resources Planning and Management,* 118, 654-670.

Hamlet, A. F., and Lettenmaier, D. P. (1999). "Columbia River streamflow forecasting based on ENSO and PDO climate signals." *Water Resources Planning and Management,* 25, 333-341.

Hsieh, W. W., and Tang, B. (2001) "Interannual variability of accumulated snow in the Columbia basin, British Columbia." *Water Resour. Res.,* 37, 1753-1759.

Nijssen, B., Lettenmaier, D. P., Liang, X., Wetzel, S. W., and Wood, E. F. (1997). "Streamflow simulation for continental–scale river basins." *Water Resour. Res., 33*, 711-724.

Redmond, K. T., and Koch, R. W. (1991). "Surface climate and streamflow variability in the western United States and their relationship to large–scale circulation indices." *Water Resour. Res.,* 27, 2381-2399.

Smith, T. M., Reynolds, R. W., Livezey, R. E., and Stokes, D. C. (1996). "Reconstruction of historical sea surface temperatures using orthogonal functions." *J. Climate,* 9, 1403-1420.

Wilks, D. S. (1995). *Statistical Methods in the Atmospheric Sciences.* Academic Press, San Diego, 467 pp.

Yuval. (2000). "Neural network training for prediction of climatological time series, regularized by minimization of the Generalized Cross Validation function." *Mon. Wea. Rev.,* 128, 1456-1473.

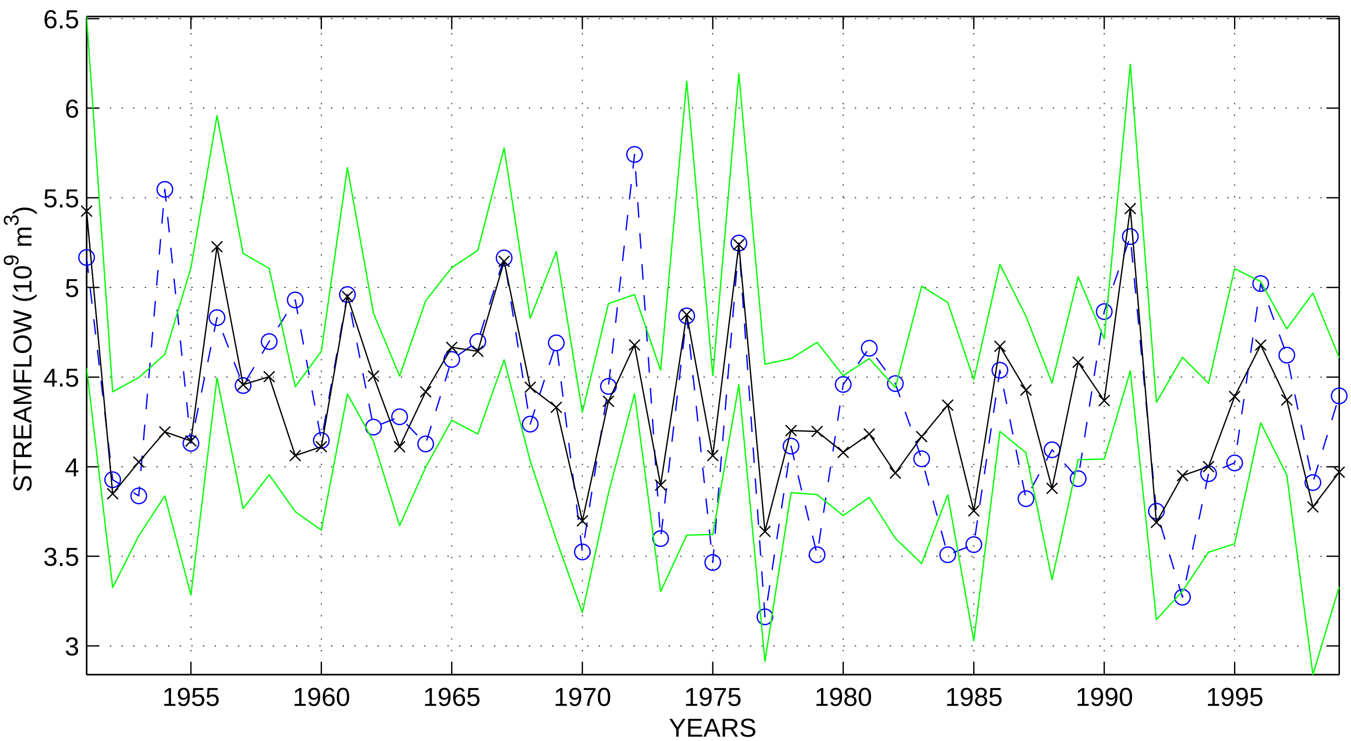
Yuval. (2001). "Enhancement and error estimation of neural network prediction of Niño–3.4 SST anomalies." *J. of Climate,* 14, 2150-2163.

Table 1: The cross-tested prediction correlation skills for the April–August streamflow at Donald using stepwise MLR. The predictors are the November SST1 (the first PC of the tropical Pacific SSTA), PNA, PDO, and the October Prec4 (the 4th PC of the local precipitation). The regression coefficients are given in parenthesis after each predictor.

correl.	predictors
0.467	SST1 (0.282), PNA (−0.412)
0.490	SST1 (0.138), PDO (−0.493)
0.527	PNA (−0.258), PDO (−0.427)
0.523	SST1(0.150), PNA (−0.266), PDO (−0.348)
0.699	SST1(0.193), PNA (−0.342), Prec4 (0.514)
0.686	SST1(0.179), PNA (−0.325), Prec4 (0.496), PDO(−0.047)

Figure captions

Fig. 1 The observed April–August Columbia River streamflow (dashed line with circles) and its corresponding MLR prediction (solid line with crosses). The predictions were made using predictors up to the end of November. The pale solid lines around the prediction are the positive and negative bootstrap error estimates.



6.5
6
5.5
5
4.5
4
3.5
3
STREAMFLOW (10^9 m^3)
1955
1960
1965
1970
1975
1980
1985
1990
1995
YEARS