SEASONAL STREAMFLOW FORECASTING USING CLIMATE INFORMATION

Connely K. Baldwin[1]

ABSTRACT

Note: This paper is a companion to "INCORPORATING SEASONAL STREAMFLOW FORECASTS INTO OPERATIONAL DECISIONMAKING" (Waage et al, this volume). It represents the text of a forecast provided to Denver Water in early November 2000.

We present forecasts for 2001 April-July streamflow volume at four gages: Dillon Reservoir (DL), Williams Fork (WF), Fraser River at Winter Park (FR), and the South Platte at South Platte (SP). The first three gages are aggregated since they are highly correlated. We include recent information on the PDO and NINO3 indices to provide tercile probabilistic forecast and associated analog years. We rely on the nearest-neighbor analog (NNA) methodology. There is increased chance of upper tercile volume at DL, WF, FR and increased chance of lower tercile volume at PL.

INTRODUCTION

We use the NINO3 (September-November) El Nino sea surface temperature (SST) index, previous winter Pacific-Interdecadal Oscillation (PDO) index conditions (Mantua et al, 1997), and average June conditions in two sea surface temperature regions—one in the Eastern Atlantic and one in the Pacific—as predictors of April-July runoff at Dillon Reservoir (DL), Williams Fork (WF), Fraser River at Winter Park (FR), and the South Platte at South Platte (SP). The PDO index has been negative in recent years and has recently become more negative. Simultaneously, the NINO3 El Niño index indicates a weak La Niña to near-normal condition. For the specific SST regions in the East Atlantic (EA) and North Pacific (NP) (Figure 1), the maximum correlation appears in the month of June prior to the water year. We use a regional average of April 1 SWE as a surrogate for April-July flow volume. We found that forecasts were better because the longer period of record captured the previous impact of long-time scale climatic oscillations that mimic recent conditions. The correlation is high between DL, WF, FR and its surrogate SWE index ($R^2 = 0.9$), but lower between PL and its surrogate SWE index ($R^2 = 0.35$), which is representative of the greater importance of lower elevation precipitation for the PL gage.

---

[1] Paper presented at the 69th Annual Meeting of the Western Snow Conference, 2001. Connely K. Baldwin is a research engineer at Utah State University, Logan, Utah.

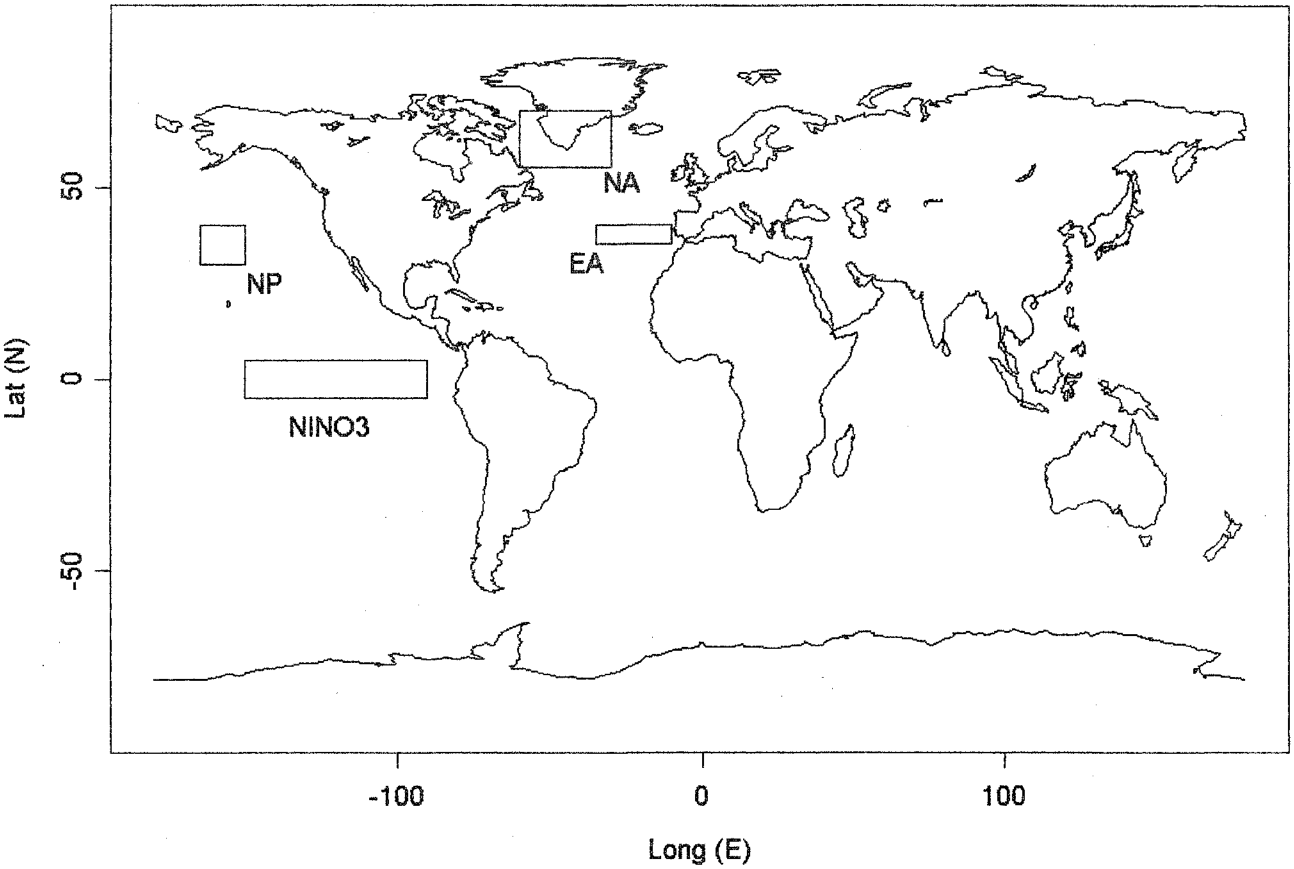

Figure 1. Location of sea surface temperatre areas used as predictors. PDO is the result of a statistical analysis of all Pacific SSTs north of 20 degrees North.

NEAREST-NEIGHBOR ANALOG METHODOLOGY

The end result of the nearest-neighbor analog (NNA) methodology is a number of "analog" years drawn from the historical record. Years are similar to the forecast year as defined by the predictors used. A complementary approach is in common use among water managers through the Extended Streamflow Prediction (ESP) system forecasts. However, unlike ESP, the NNA method ranks years in order of similarity to the forecast year. We give a more detailed description of the method and predictors used below.

Given a set of predictors (noted above) and a predictand (April-July flow volume in this case), we find the "k" years similar to the current state, as defined by the predictors. The distance between the current state and the state of historical years determines the similarity. Technically, we use the Euclidean distance metric,

$$D_i = \{(x1_m - x1_i)^2 + (x2_m - x2_i)^2 + \ldots + (xN_m - xN_i)^2\}^{0.5},$$

where D_i is the distance, or similarity measure, between the current state, m, and all other points, i, for all $i \neq m$. xN_m is the N^{th} predictor of the point of interest and xN_i is the nth predictor for all other points. The "k" historical years that have the smallest distance are the most similar and are chosen as the analog years. An illustrative example is presented in Figure 2.

The analog years define what may be reasonably expected for the upcoming runoff season. The range of volumes is a guide to the certainty of the forecast. If the range is large, we have little predictability and make decisions accordingly. However, if the range is narrow we bias decisions in favor of this condition to improve decision-making.

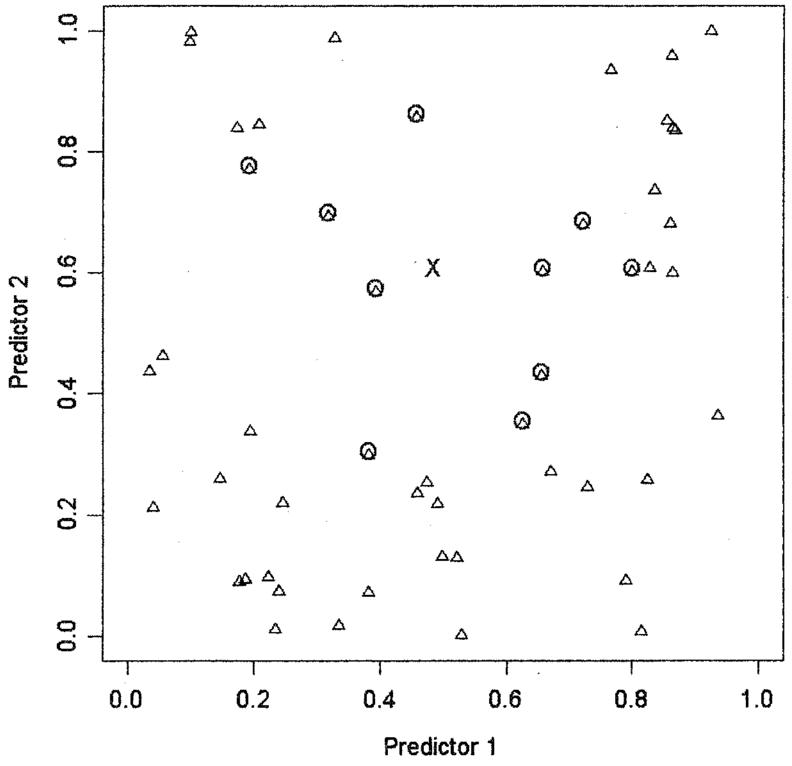

Figure 2. Illustrative example of choosing the *k* most historical analogs based on 2 predictors, where $k = 10$. The X represents current conditions, the triangles, historical conditions, and the open circles the 10 previous years most similar to the current conditions.

The analog years can also guide the use of existing ESP forecasts. The streamflow traces associated with good analog years for the upcoming season can be emphasized.

We also use the results of the analog years to produce a probabilistic forecast. For example, if 5 of the 10 analog years fell in the "Above Normal" category, the forecast would be 50% chance of observing above normal runoff. However, eue to the high variability of using a small number of analog years to produce the forecast, we use a method developed by Rajagopalan et al. (in preparation) to make the forecast more robust. In essence they reduce the probabilities closer to that observed on average (33% in each tercile) according to how well the forecast performed in the historical record.

FORECAST RESULTS

We present the robust probabilistic forecast and corresponding analog years. The analog years can be used in conjunction with the ESP forecasts to focus attention on a few traces, indicating which of the *possible* traces are most *probable*.

The conditions that dominated in the analog years, and the probabilistic predictions for these years are shown in Table 1. The analog years are presented in the order of similarity to the upcoming year. Note that we present 7 analog years for the DL, WF, FR gages and 15 for PL. This is the number of analog years that produced the best probabilistic forecast over the historical record. The jackknife (leave-one-out) blind probabilistic predictions for the analog years are also presented in Table 1 to reinforce the probabilistic nature of the forecast. As we compare the probabilistic forecast and outcome for each analog year, we see that in most years the event with the largest probability occurs (1962, 1974, 1971, 1950, 1952), however some years the event with lower probabilities occur (1989, 1995). This is consistent with the probabilistic nature of the forecast.

The probabilistic forecast for 2001 is shown in Table 2. This forecast indicates greater likelihood of seing upper tercile April-July volume in DL, WF, FR and greater likelihood of seeing lower tercile April-July volume in PL.

Table 1. Analog years and associated jackknifed probability forecasts for regional April 1 SWE for the analog years identified by the NNA tool. The number in the Actual Category column refers to the same categories defined in the Forecast Probabilities columns (i.e., a 1 represents Lower Tercile, etc.).

Year	Actual Cat.	Forecast Probabilities		
		Lower (1)	Middle (2)	Upper (3)
Dillon, Williams Fork, Fraser				
1962	3	0.21	0.30	0.48
1974	2	0.30	0.39	0.30
1971	3	0.21	0.30	0.48
1989	1	0.12	0.48	0.39
1950	2	0.21	0.39	0.39
1952	3	0.21	0.39	0.39
1995	1	0.30	0.21	0.48
South Platte				
1959	1	0.28	0.38	0.33
1966	1	0.28	0.38	0.33
1962	3	0.33	0.38	0.28
1960	2	0.43	0.28	0.28
1984	3	0.13	0.38	0.48
1974	2	0.38	0.23	0.38
1948	3	0.18	0.38	0.43
1950	1	0.38	0.33	0.28
1958	3	0.28	0.28	0.43
1946	1	0.48	0.13	0.38
1956	1	0.38	0.28	0.33
1989	1	0.38	0.33	0.28
1964	2	0.38	0.23	0.38
1995	2	0.18	0.38	0.43
1990	1	0.48	0.23	0.28

NOTE: In the three-state case, where we have divided the historical data into terciles (driest 1/3 (LT), middle 1/3 (MT), wettest 1/3(UT)).

Table 2. Probabilistic forecasts for WY 2001 April-July volume.

Gage	Tercile		
	Lower	Middle	Upper
DL, WF, FR	30%	30%	40%
SP	47%	27%	26%

REFERENCES

Rajagopalan, B., U. Lall, and S. Zebiak, Submitted to the Bulletin of the American Meteorological Society. Categorical Climate Forecasts through Regularization and Optimal Combination of Multiple GCM Ensembles.

Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M., and Francis, R.C. 1997. A Pacific Inter-decadal Climate Oscillation with Impacts on Salmon Production, Bull. of the Amer. Meteor. Soc., 78, 1069-1079.