INTRODUCTION
In this study we propose a post-processing procedure that assigns weights to streamflow ensemble members using the large scale climate signals. Analysis is performed over the rainfall dominated Leaf River basin and snow dominated East River basin to improve the spring ensemble streamflow volume forecast. We employ Fuzzy C-Means clustering, Formal Likelihood, Informal Likelihood and K-Nearest Neighbors methods for the weighting and it is found that Principle Component Analysis (PCA) improve the accuracy of the weighting scheme considerably. The presented objective method can be applied to enhance the final ESPs; nevertheless the user expertise may change any of the process steps. The current predictions based on simple average or the median of the ensemble members may come with the weighted ensemble forecasts to better provide possible ranges and uncertainty bounds.

STUDY AREAS
The snow dominated East River basin is a tributary of the Gunnison River in the Colorado River basin. It is 754 km² with the elevation ranges from 2445 to 3898m. The data are obtained from the NWS CBRFC and consists of the six hourly precipitation and temperature data for each of the three elevation bands along with the daily streamflow in the outlet. The duration of the streamflow data is 1975-2005 and the forcing data 1960-2005. The Leaf river basin is a rainfall dominated basin with an area of 1944km², which is located in the north of Collins, Mississippi, United States. The Precipitation, Temperature and Streamflow data were available from 1948-1988.

Models
We combine SNOW-17 model with the Sacramento Soil Moisture Accounting Model (SAC-SMA) to predict the streamflow in snow dominated basin. SNOW-17 model is driven by temperature and precipitation as the forcing data and it simulates the Snow Water Equivalent (SWE), Snow Cover Area (SCA), and Rain plus Melt (RM) at each time step. RM then becomes the input to the SAC-SMA along with the Potential EvapoTranspiration (PET) to estimate the runoff. For Leaf rainfall dominated basin it solely runs the SAC-SMA model.

Selecting the Climate Signals
We use the information content in the climate signals in weighting the ensemble trajectories for the spring (April, May, June) streamflow volume forecasts. For this purpose we select the climate signals using the spearman correlation between each monthly signal (starting from the January in the previous year until March in the forecast year) and the spring runoff volume. For the hypothesis test, a correlation with a P-value of less than 0.05 is considered significant. In order to perform the ESP weighting, this process is done for each forecast year separately.

Provision of Uncorrelated Climate Signals
PCA is a procedure that removes the co-linearity between the variables and minimizes the redundancy in the information by obtaining principal components (PCs) that are linear combinations of the original variables.

We make the signals uncorrelated using PCA to reduce this redundancy and to increase the efficiency of the weighting methods. In order to determine the number of PCs to retain, we consider two approaches of the so called “NRCs” and “PRESS” methods:

1. If the NRCs method, PCs are added to the regression model one-by-one. A statistical t-test is performed each time a PC is added. If the addition of the PC is significant compared with a critical value of 1.6, it would be retained, else the current number of PCs are chosen.

2. In the Prediction Residual Sum of Squares (PRESS) method, a leave-one-out cross validation is performed and the value of PRESS is calculated each time a value is extracted. The optimum number of PCs is obtained based on the minimum value of PRESS.

Ensemble Streamflow Prediction
Weights are assigned to each streamflow ensemble member based on the similarity of the climate signal PCs in the forecast year to the ones in the previous years using the following methods:
- Fuzzy C-Means clustering algorithm: (Bedeck and Ehrlich, 1984)
- Formal Likelihood: Involves a Multivariate Normal distribution with the selected climate signals in the forecast year as the mean and the covariance of the climate signals in the historical years as the covariance of the distribution.
- Informal Likelihood: Using the function presented by (Beven and Binley, 1992):

\[ I_{j} = \frac{1}{I_{max}^{j}} \]

where \( I_{j} \) is the variance of the residuals between the climate signals in the forecast year and the ones in the historical year.

- K-Nearest Neighbor (KNN): The K number of nearest climate signal PCs in the previous years to the ones in the forecast year is chosen and weighted based on their Euclidean distance.

- 1D-KNN: Similar approach to KNN is adopted by (Werner et al., 2004).

Results

\[ \%change = \frac{\sum_{i=1}^{n} \left( \bar{x}_i - \bar{x}_o \right)}{\bar{x}_o} \times 100 \]

where \( \bar{x}_i \) and \( \bar{x}_o \) represent the observed spring streamflow, simple mean and weighted mean of the ensemble streamflow predictions respectively.

CONCLUSIONS
- Climate Signals show significant Correlation with the spring streamflow in both basins.
- The Principle Component Analysis effectively reduces the redundancy in the selected signals and enhances the overall weighting scheme considerably.
- The post processing weighting schemes improves the accuracy of the spring runoff volume prediction.
- Care should be given to the choice of parameters of each method and number of PCs. An objective method to select the number of PCs is proposed in this study.
- The weighting methods are simple and efficient and the results are promising. The whole process can be automatically performed in an objective way.

REFERENCES