Statistical Multimodel Ensemble Forecasting Technique: Application to Spring Flows in the Gunnison River Basin

Presentation by
Satish Kumar Regonda [satish.regonda@noaa.gov]

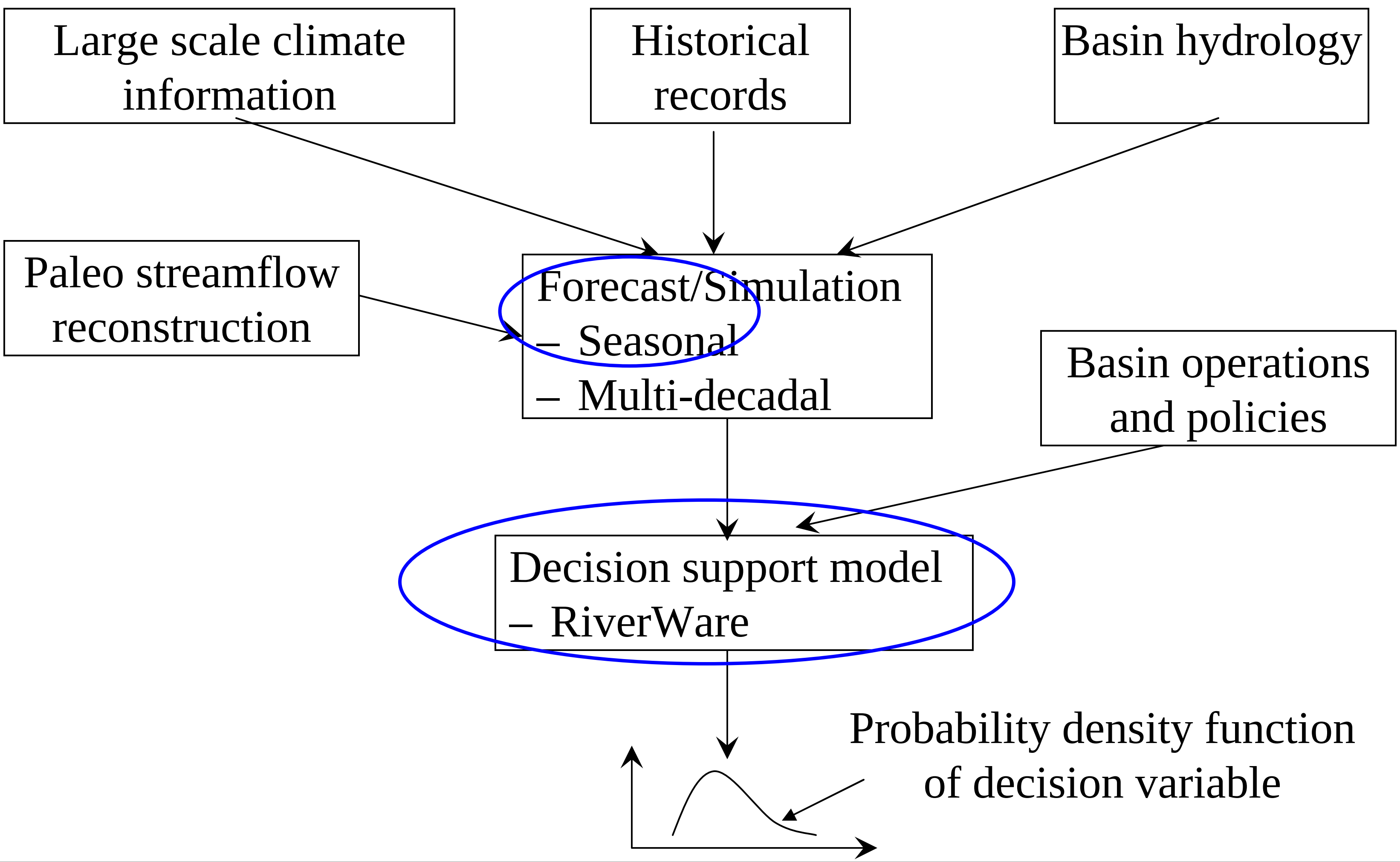
Advisors
Dr. Balaji Rajagopalan, Dr. Martyn Clark, Dr. Edith Zagona
University of Colorado at Boulder

CADSWES
The Center for Advanced Decision Support for Water and Environmental Systems

Western Water Assessment

CIRES

Research: Development of a Decision Support System With Improved Forecasting Tools



Regonda, S.K. (2006), Intra-annual to Inter-decadal Variability in the Upper Colorado Hydroclimatology: Diagnosis, Forecasting, and Implications for Water Resources Management, Ph.D. thesis, University of Colorado at Boulder, Colorado.

Highlights of this presentation

Multi-model Ensemble Streamflow Forecasting Framework

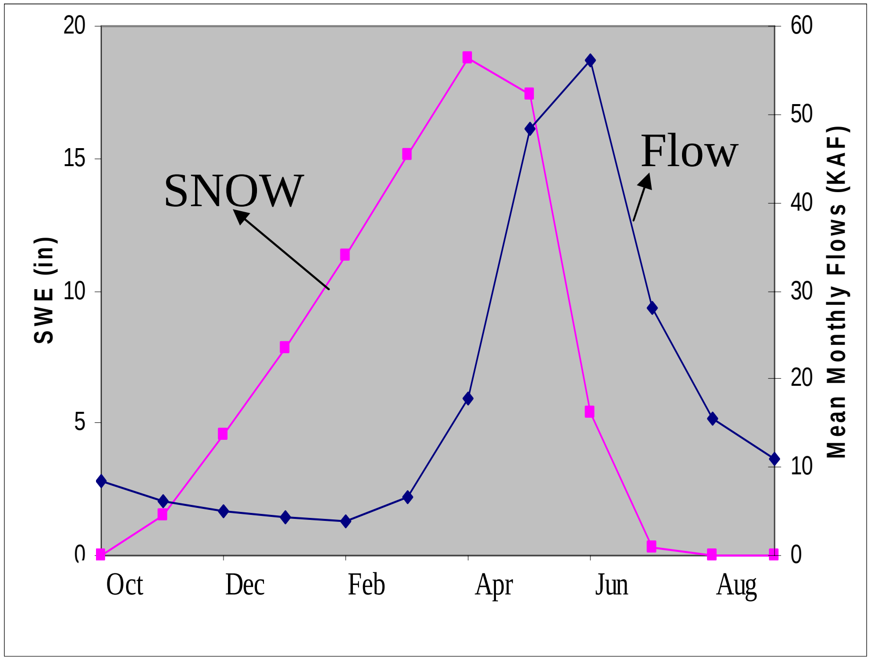
- A new framework that provides
  - Long lead forecasts i.e., starting from winter
  - Ensemble of forecasts that captures uncertainty
  - Multi-site ensemble forecast

Application: Gunnison River Basin and Aspinal Unit

- Decision Support Model
  - Value of the forecasts
  - Results of forecasting operations corresponding to January 1st and April 1st forecasts

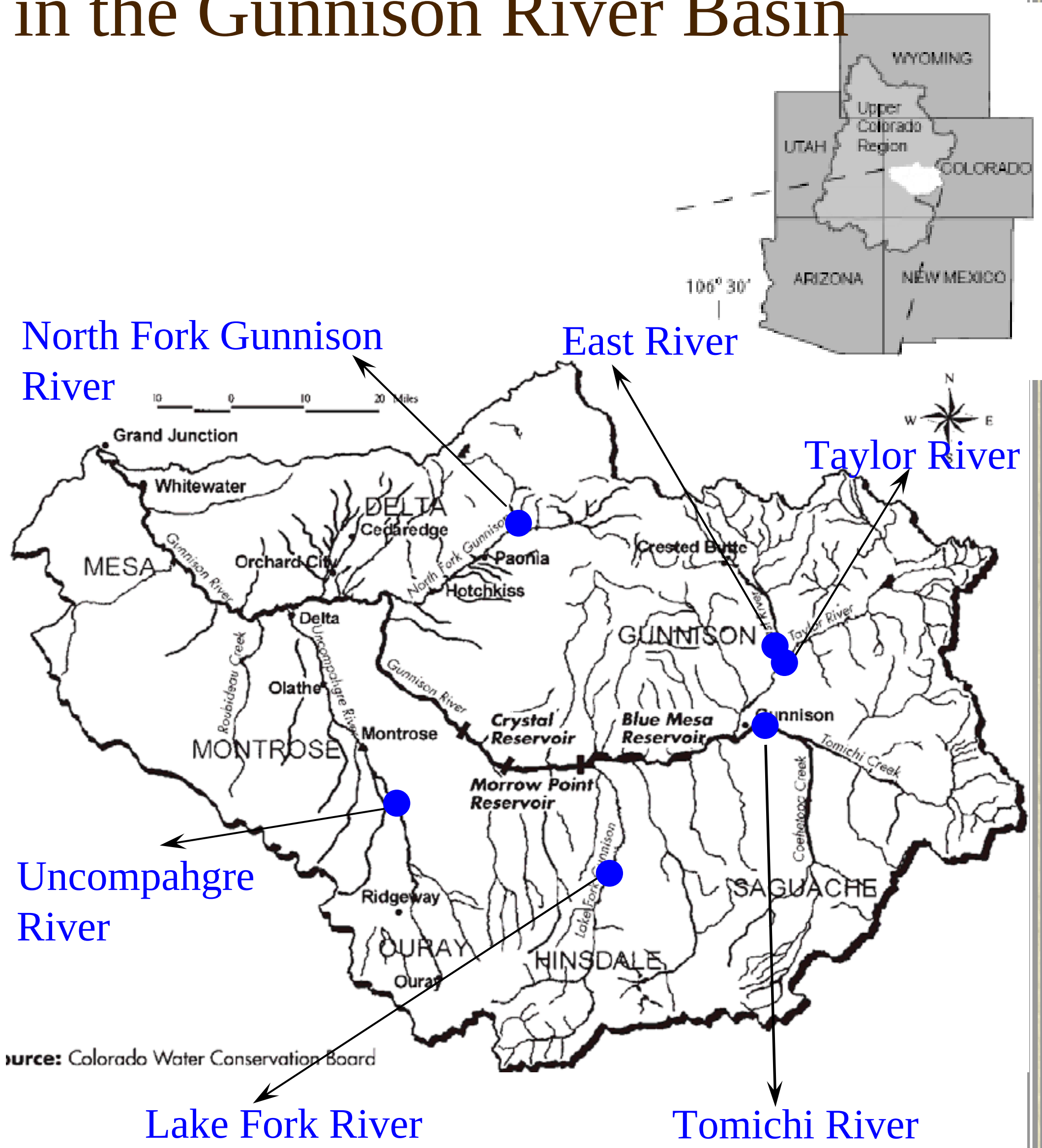
Gunnison River Basin: hydrologic characteristics

- Majority of annual runoff is snowmelt (70%)
- Seasonal (spring) forecast is critical to optimal basin management:
  - conservation and delivery to meet irrigation demands,
  - hydropower production
  - environmental releases
- Spring streamflow forecast needed early in winter



Six selected gauges in the Gunnison River Basin

- Drainage area:
  7930 mi2
- Elevation:
  4600 - 14530 ft
- Records:1949-2002
- Used to represent integrated hydrologic variability of the basin

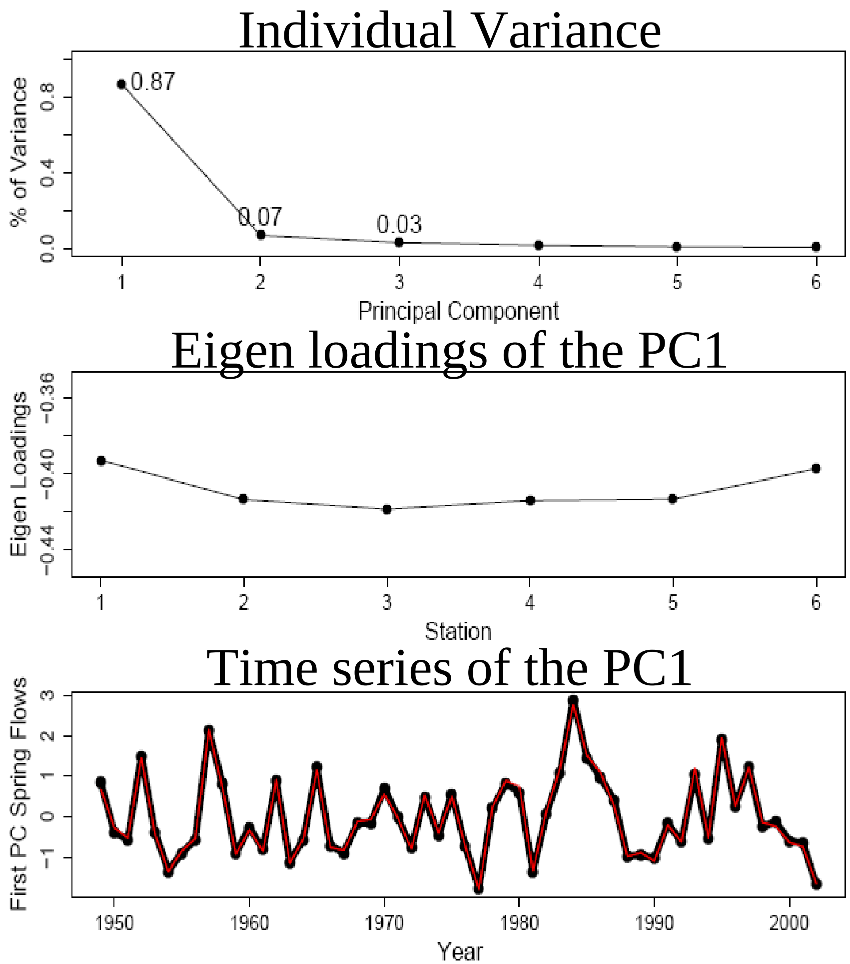


Methodology

- Diagnosis
  - Principal Component Analysis (PCA) on spring streamflow at the six locations (leading streamflow component)
  - Climate Diagnosis → identify predictors of the leading component
- Forecast (of leading component)
  - Local Polynomial Method (Multi-model ensemble)
  - Back transform → flows at all the six locations
- Forecast Skill Evaluation
  - BoxPlots
  - Ranked Probability Skill Score (RPSS)

PCA results

- First PC explained most of the variance ~ 87%
- Eigen loadings of the first are uniform
- Select the first PC, i.e., PC1 spring streamflows for further analysis
- Correlation between PC1 and average of spring flows is 0.99
  - Red: Average of spring flows
  - Black: PC1 spring flows



Hydrologic variability of the basin is represented by PC1 spring streamflows

Correlate PC1 spring flows with potential predictors

- Teleconnection indices

  e.g., ENSO index, PNA index, PDO index, etc.

- Large scale climate variables

  e.g., Pressure anomalies, Sea surface temperatures, Surface air temperature, Meridional and Zonal winds

- Snow

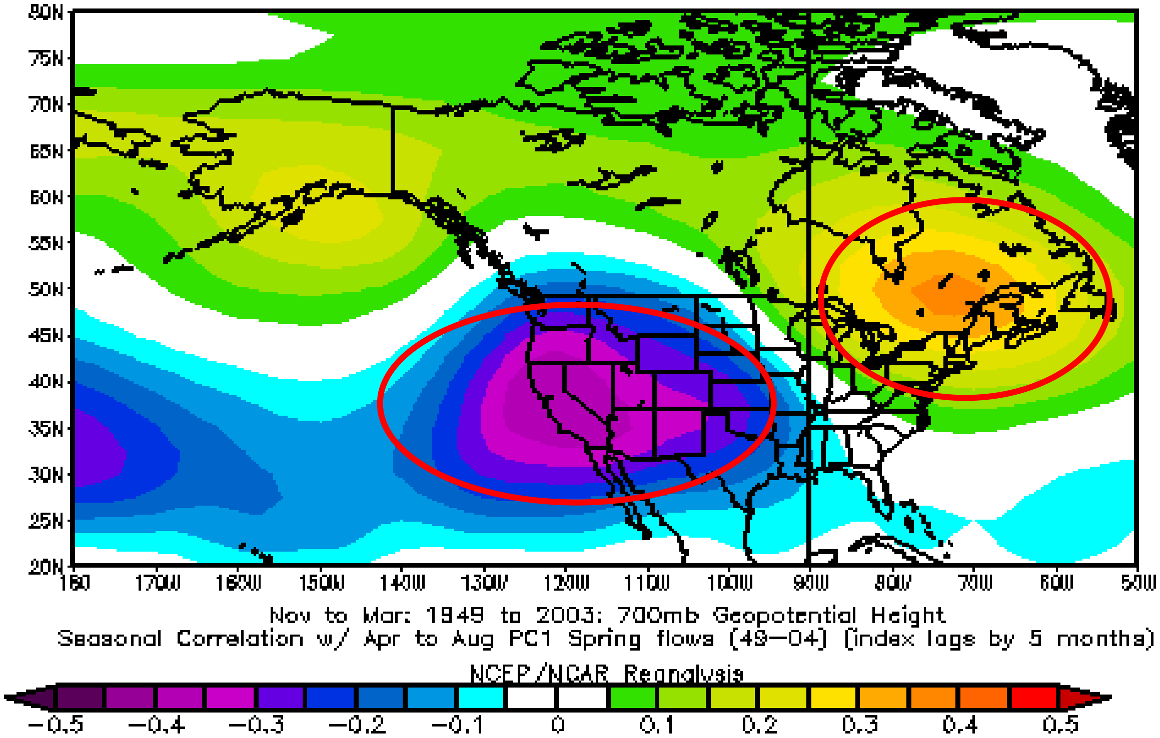
  Snow Water Equivalent (SWE)

- Antecedent soil moisture conditions
  - Palmer Drought Severity Index (PDSI)

Correlation of PC1 with teleconnection indices is 0!

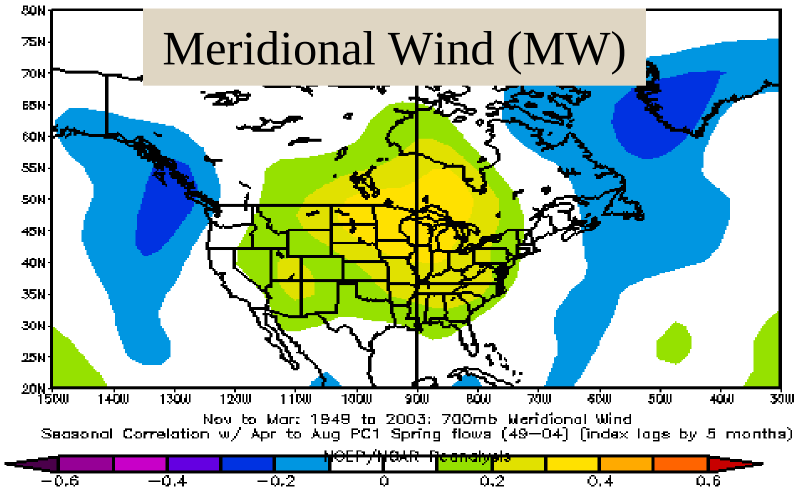
PC1 spring flows correlated with large scale climate variables

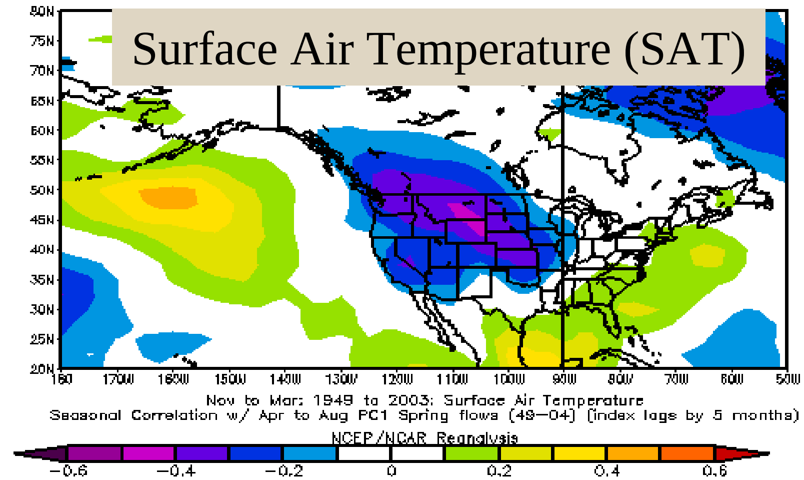
Pressure anomalies (GPH) for winter (Nov-Mar)

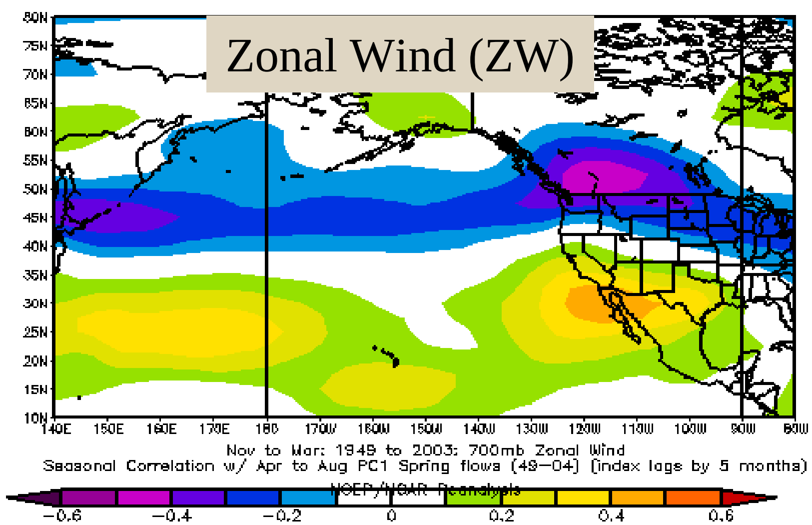


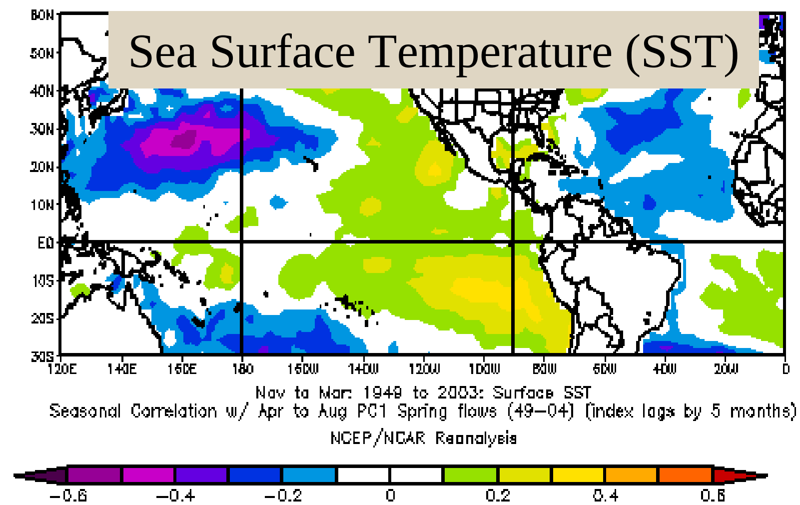
1. Identify regions of high correlation
2. Prepare corresponding time series, i.e., area averaged value

PC1 spring flows correlated with large scale climate variables





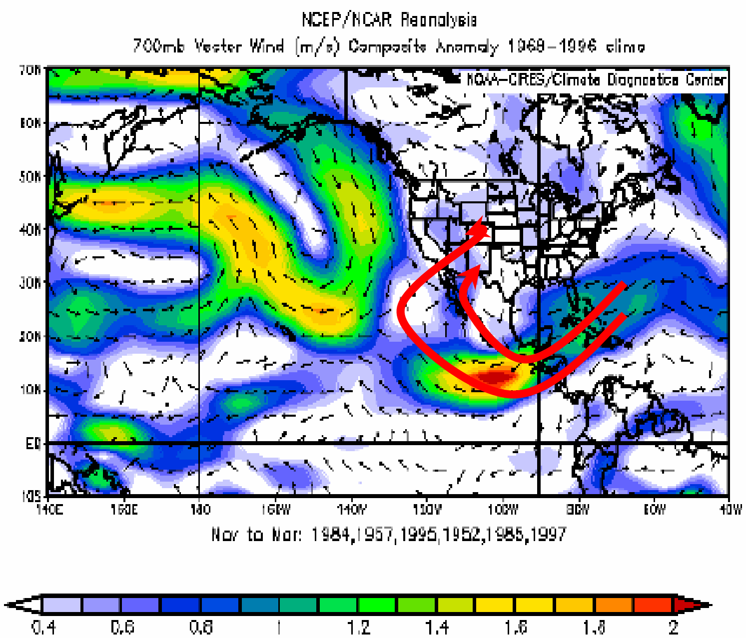




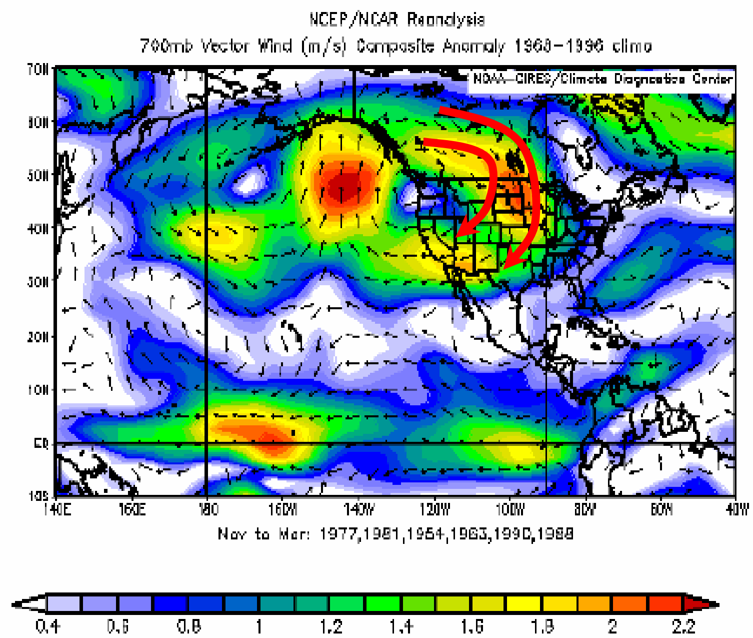
Understanding Physical Processes

winter (Nov-Mar) vector wind composites

Wet years



Dry years

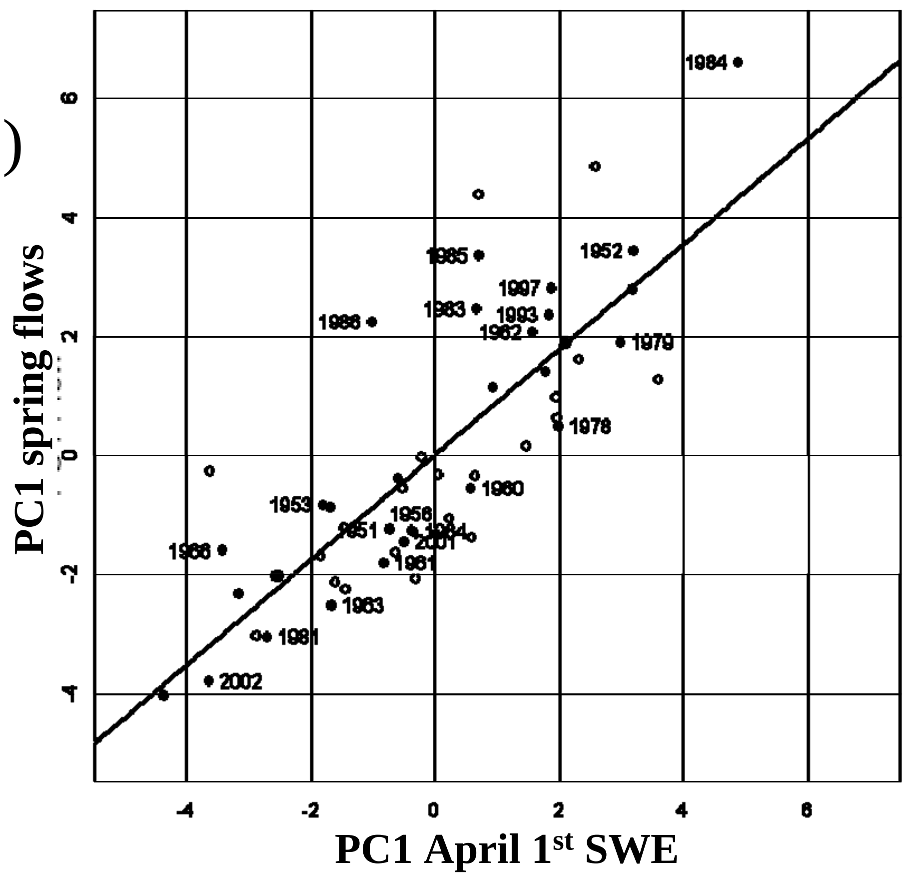


source: cdc.noaa.gov

PC1 spring flows correlated with April 1st SWE

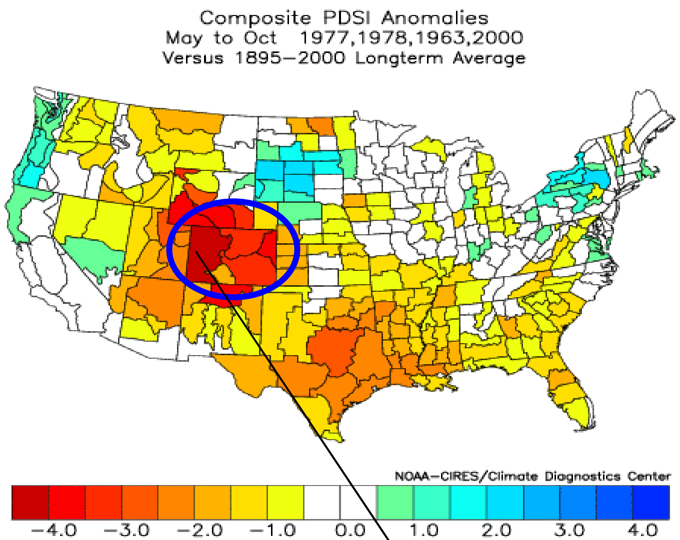
- deviations from linear relationship (solid circles)
- role of antecedent land conditions?

PC1 Flows Vs. PC1 SWE	Feb.	Mar.	Apr.
Correlation Coefficient	0.67	0.72	0.82

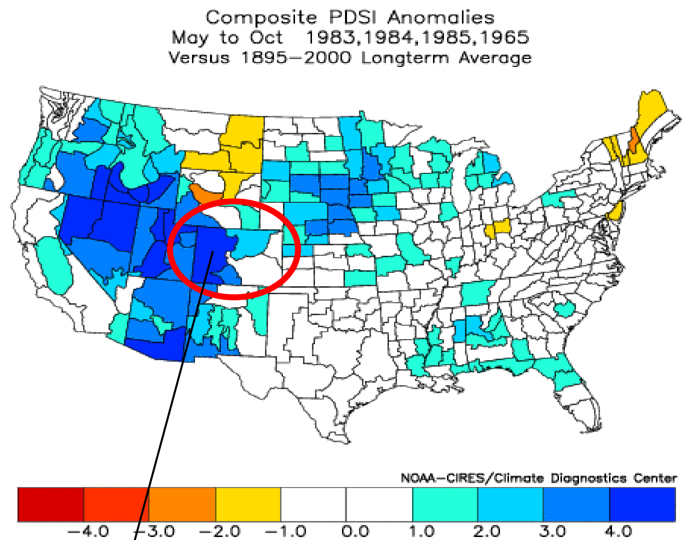


Role of antecedent land conditions

Years with high snow and proportional low flows



Years with low snow and proportional high flows



Gunnison River Basin [Colorado Region 2]

Palmer Drought Severity Index (dry ←------→wet)

Model selection

- Develop a model for each combination of predictors, test model with historical data, then select the best model
- Several predictor subset models have similar goodness of fit (GCV) values!

Model #	SAT	GPH		MW		ZW		SST		PDSI	PC1 SWE	# Variables	GCV
1	0	0	0	0	0	0	0	0	1	0	1	2	2.03
2	0	0	0	0	0	0	0	1	0	0	1	2	2.04
4	0	0	0	0	0	0	0	0	0	0	1	1	2.07
5	0	0	0	1	0	0	0	1	0	0	1	3	2.08
10	0	0	0	0	0	0	0	0	0	1	1	2	2.14

Multimodels

- All models with GCV values less than a fixed threshold (i.e., 1.2 times the minimum GCV) are selected.
- Eliminate models in which predictors are significantly correlated amongst each other (i.e. multicollinearity)

Model #	SAT	GPH		MW		ZW		SST		PDSI	PC1 SWE	# Variables	GCV
1	0	0	0	0	0	0	0	0	**1**	0	**1**	**2**	2.03
2	0	0	0	0	0	0	0	**1**	0	0	**1**	**2**	2.04
4	0	0	0	0	0	0	0	0	0	0	**1**	**1**	2.07
5	0	0	0	**1**	0	0	0	**1**	0	0	**1**	**3**	2.08
10	0	0	0	0	0	0	0	0	0	**1**	**1**	**2**	2.14

Predictors for the forecast issued on January 1st & April 1st

- January 1st forecast predictors:
  - Pressure anomalies, Winds, Sea Surface Temperatures, Soil Moisture Conditions
  - "12" different combinations of above predictors
- April 1st forecast predictors
  - Winds, Soil Moisture Conditions, SWE
  - "6" different combinations of above predictors

Number Of Predictors	SAT-N1	(GPH-P1) – (GPH-N1)	(GPH-P2) – (GPH-N1)	(MW-P1) - (MW-N1)	(MW-P1) – (MW-N2)	(ZW-P1) – (ZW-N1)	(ZW-P2) – (ZW-N1)	(SST-P1) – (SST-N1)	(SST-P2) – (SST-N1)	PDSI	SWE (First PC)	GCV
1	0	0	0	0	0	0	0	0	0	0	1	2.07
2	0	0	0	0	0	0	0	0	0	1	1	2.14
2	0	0	0	1	0	0	0	0	0	0	1	2.15
2	0	0	0	0	1	0	0	0	0	0	1	2.32
3	0	0	0	1	0	0	0	0	0	1	1	2.34
3	0	0	0	0	1	0	0	0	0	1	1	2.51

Generate ensemble of flows from best models for each lead time

Use best models (weights are function of goodness of fit)

Generate an ensemble of estimated flows (traces) from each model as a function of explained and unexplained model variance

Final ensemble = weighted combination of traces

Model 1 (0.6)
- Esti. flow $_{1,1}$
- -------------
- Esti. flow $_{1,100}$

Model 2 (0.3)
- Esti. flow $_{2,1}$
- …………..
- Esti. flow $_{2,100}$

Model 3 (0.1)
- Esti. flow $_{3,1}$
- ……………
- Esti. flow $_{3,100}$

Esti. flow $_{1,a}$
Esti. flow $_{1,b}$
Esti. flow $_{1,c}$
Esti. flow $_{1,d}$
Esti. flow $_{1,e}$
Esti. flow $_{1,f}$
Esti. flow $_{2,a}$
Esti. flow $_{2,b}$
Esti. flow $_{2,c}$
Esti. flow $_{3,a}$

Multimodel ensemble forecasts

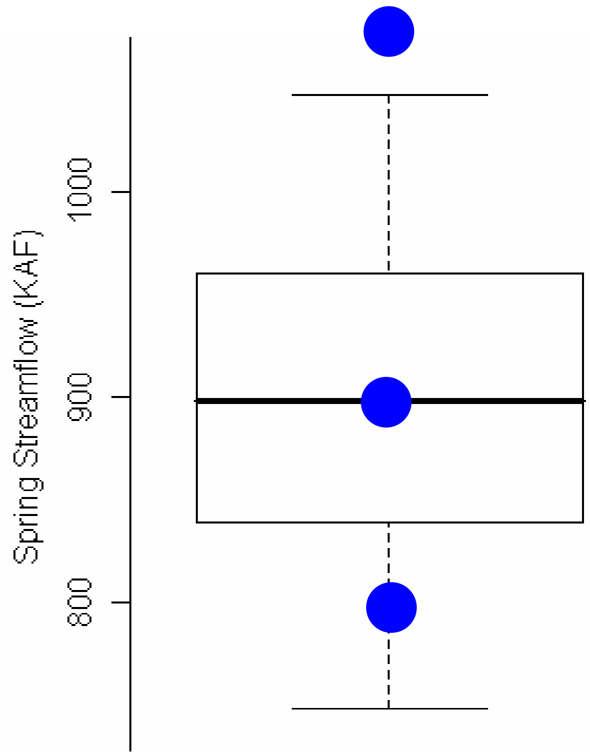
Model forecasts

Multimodel ensemble spring streamflows are forecasted for each year, in cross-validate mode, at different lead times, i.e., on the 1st of each month from December through April, at all six locations simultaneously

BoxPlots show probability distribution of ensemble forecast

Forecasted spring streamflows = {896,795.65, 936, 1056, 891.76,…… }

Actual spring streamflows



95th percentile

75th percentile

50th percentile

25th percentile

5th percentile

Evaluate forecast skill

Ranked Probability Skill Score (RPSS) is percent of improvement over the reference forecast e.g., climatology

- Categorical skill score
- Three categories are defined

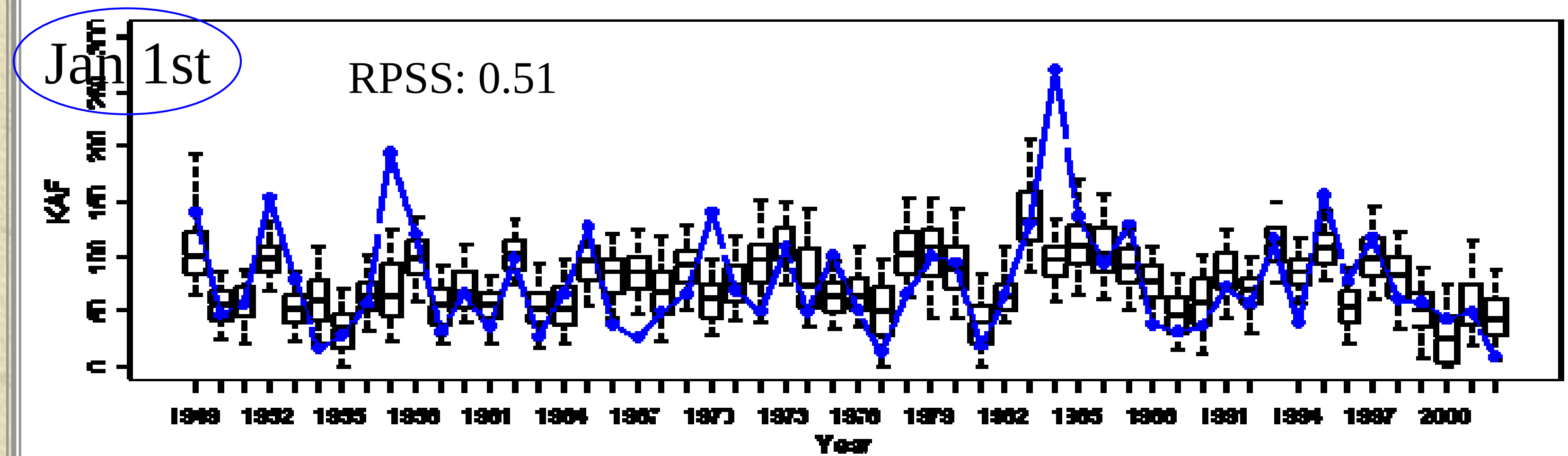
$$RPS(p,d) = \frac{1}{k-1}\left[\sum_{j=1}^{k}\left(\sum_{n=1}^{i} P_n - \sum_{n=1}^{i} d_n\right)\right] \qquad \mathrm{RPSS} = 1 - \frac{\mathrm{RPS(forecast)}}{\mathrm{RPS(climatology)}}$$

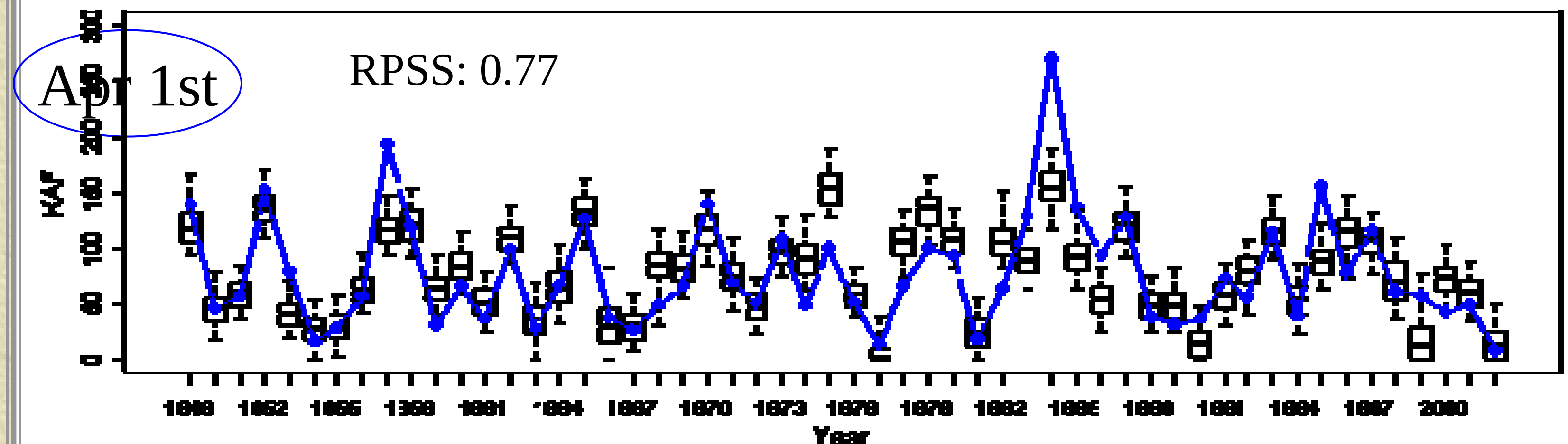
RPSS = 0, as good as climatology

RPSS = 1, high skill – much better than climatology

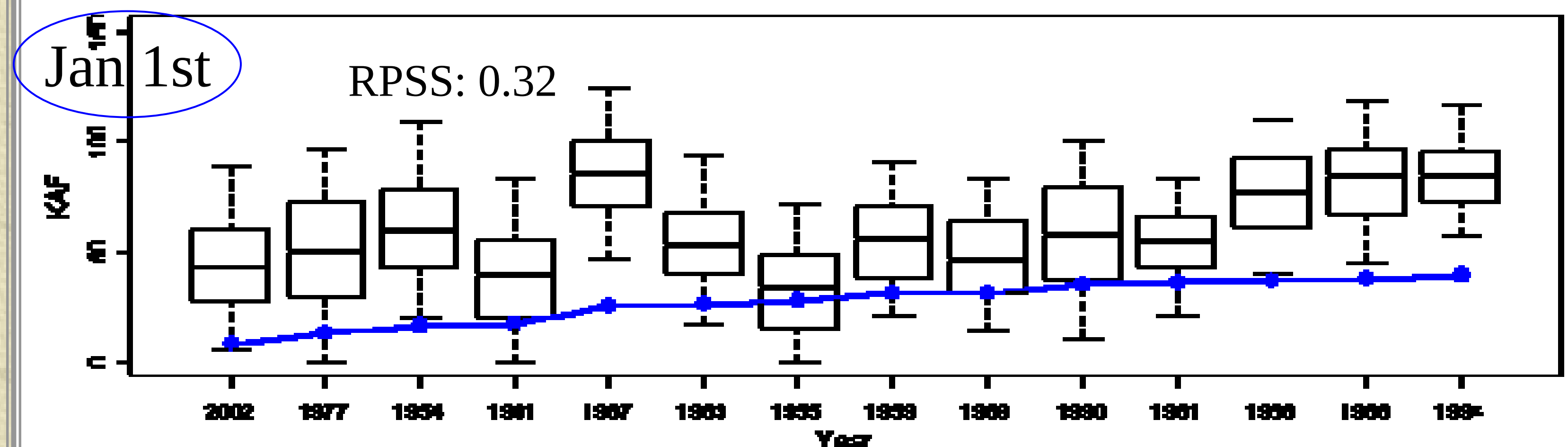
RPSS < 0, forecast worse than climatology

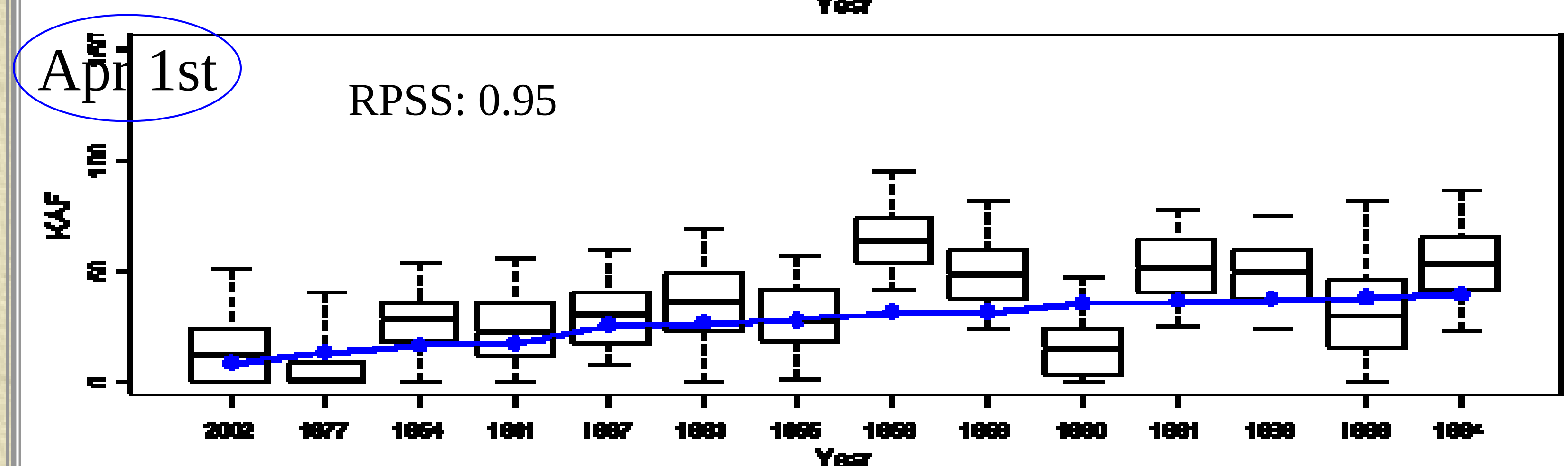
Model validation for Tomichi River (1949-2002)



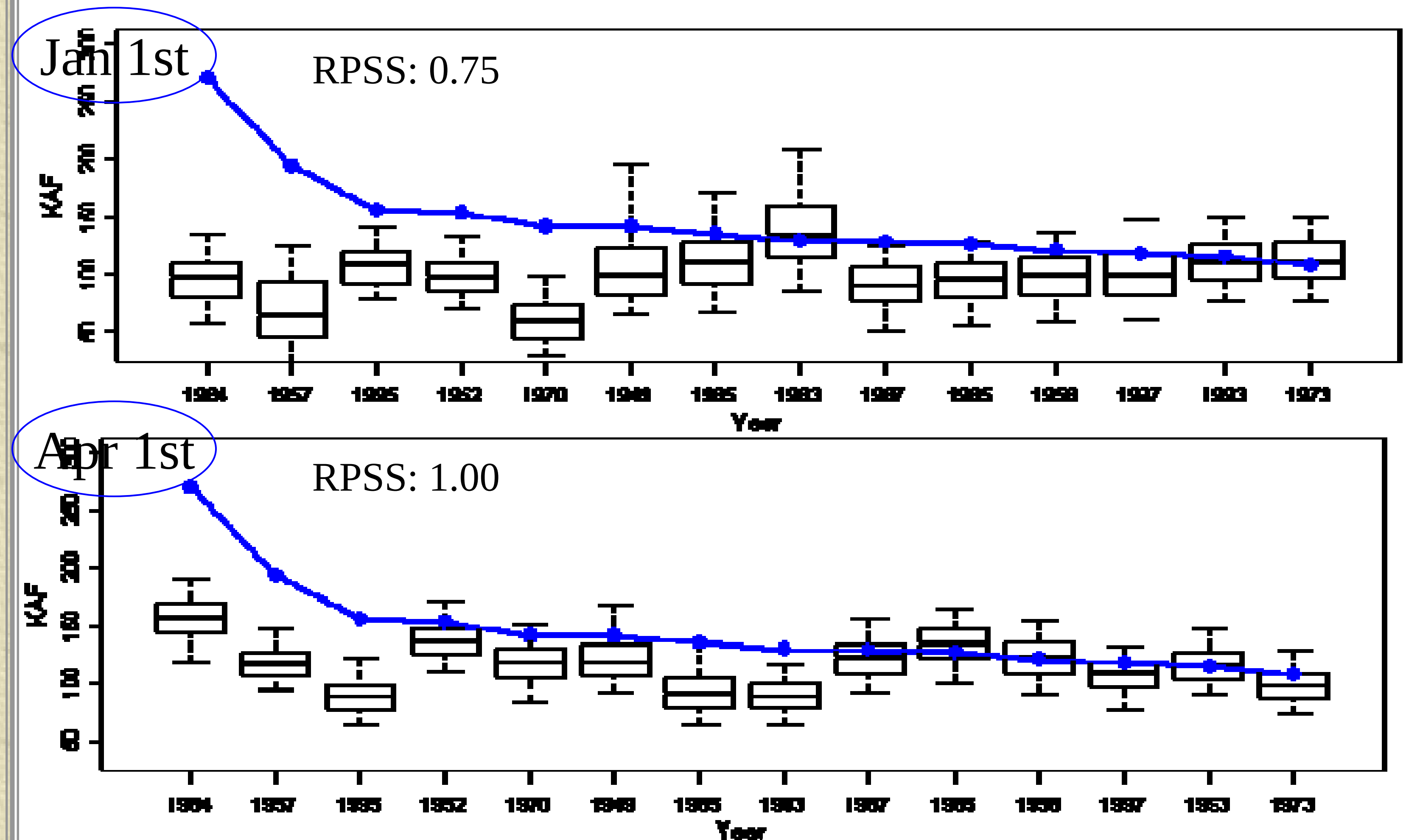


Model validation for Tomichi River (dry years)





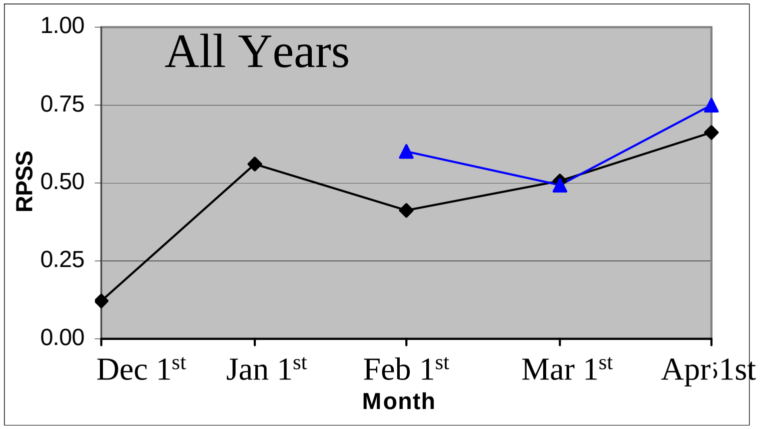
Model validation for Tomichi River (wet years)

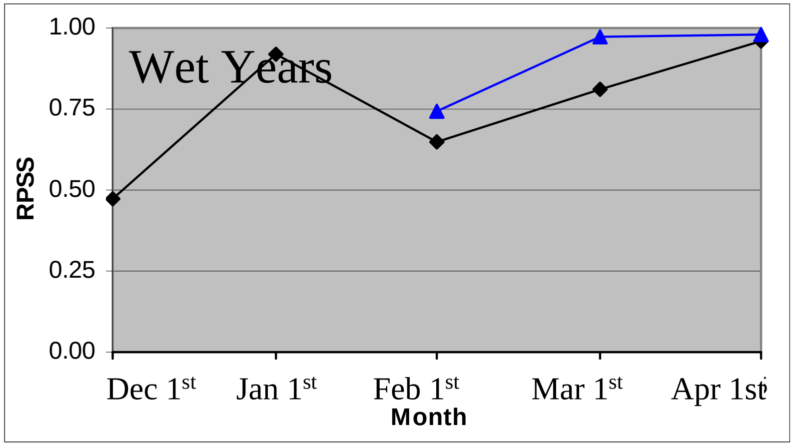


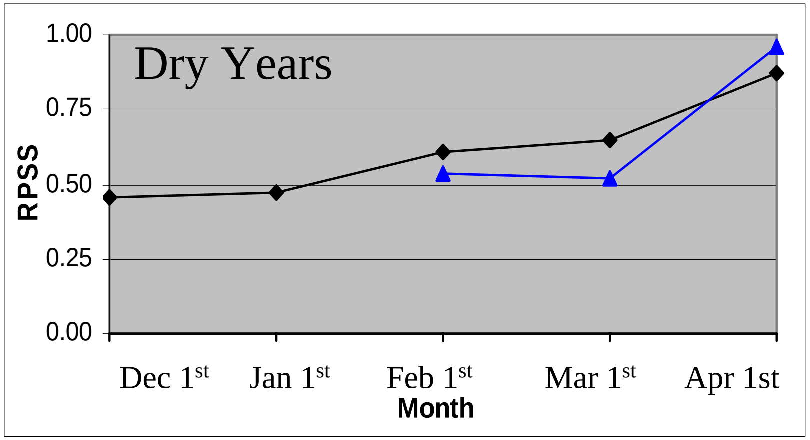
Forecast skill of spring flows at different lead times

Climate indices + Soil moisture

Climate indices + Soil moisture + SWE

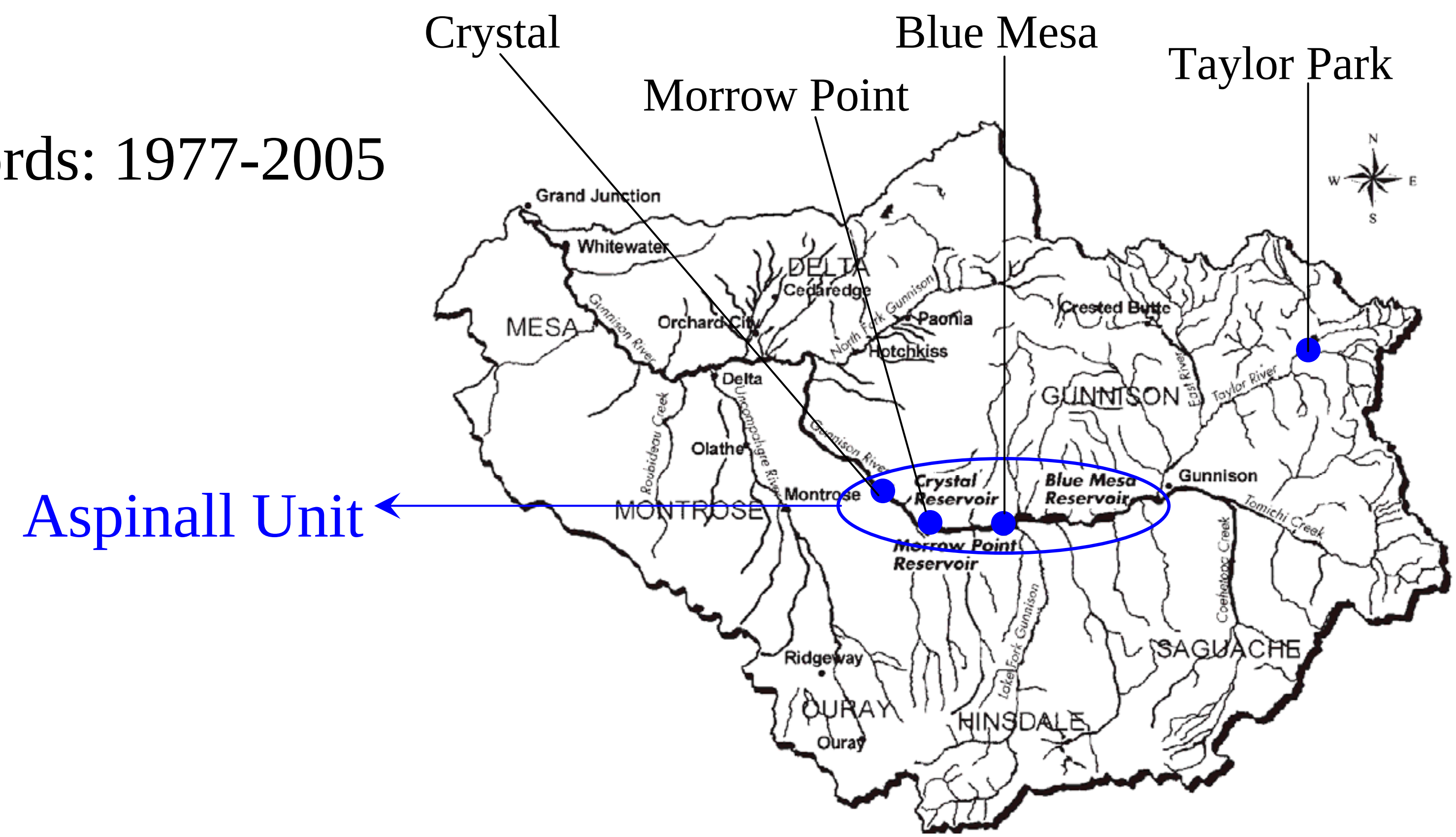






Apply forecast models (same predictors) to reservoir inflow

Records: 1977-2005

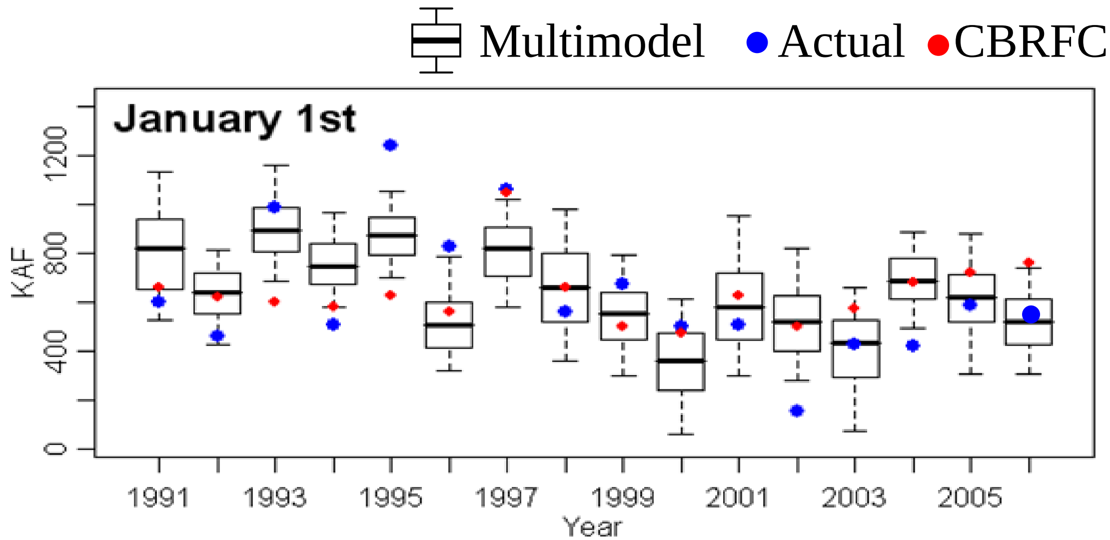


● Inflow locations required by decision model (28 years of data is available)

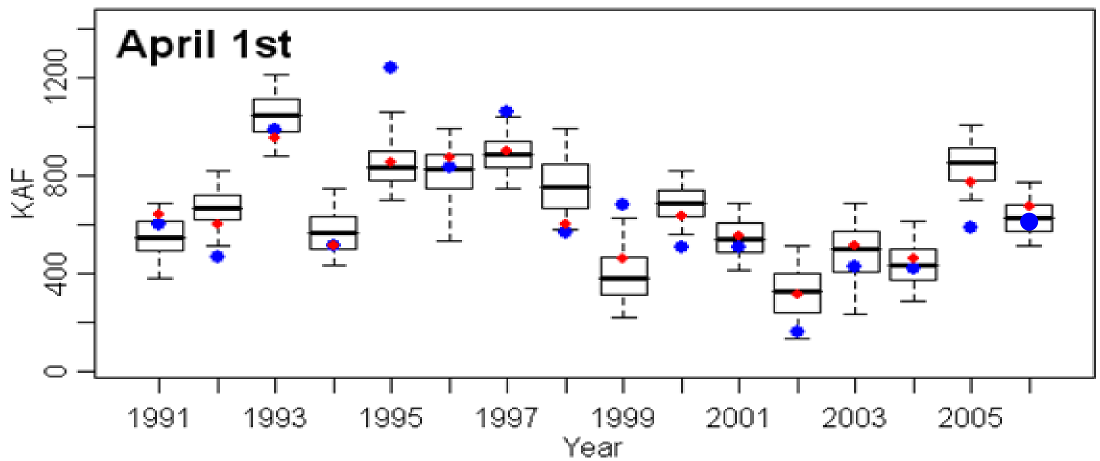
Blue Mesa reservoir forecasted inflows (unregulated)

Correlation

January 1st Forecast
Multimodel: 0.621
CBRFC forecast: 0.434



April 1st Forecast
Multimode : 0.743
CBRFC forecast: 0.860



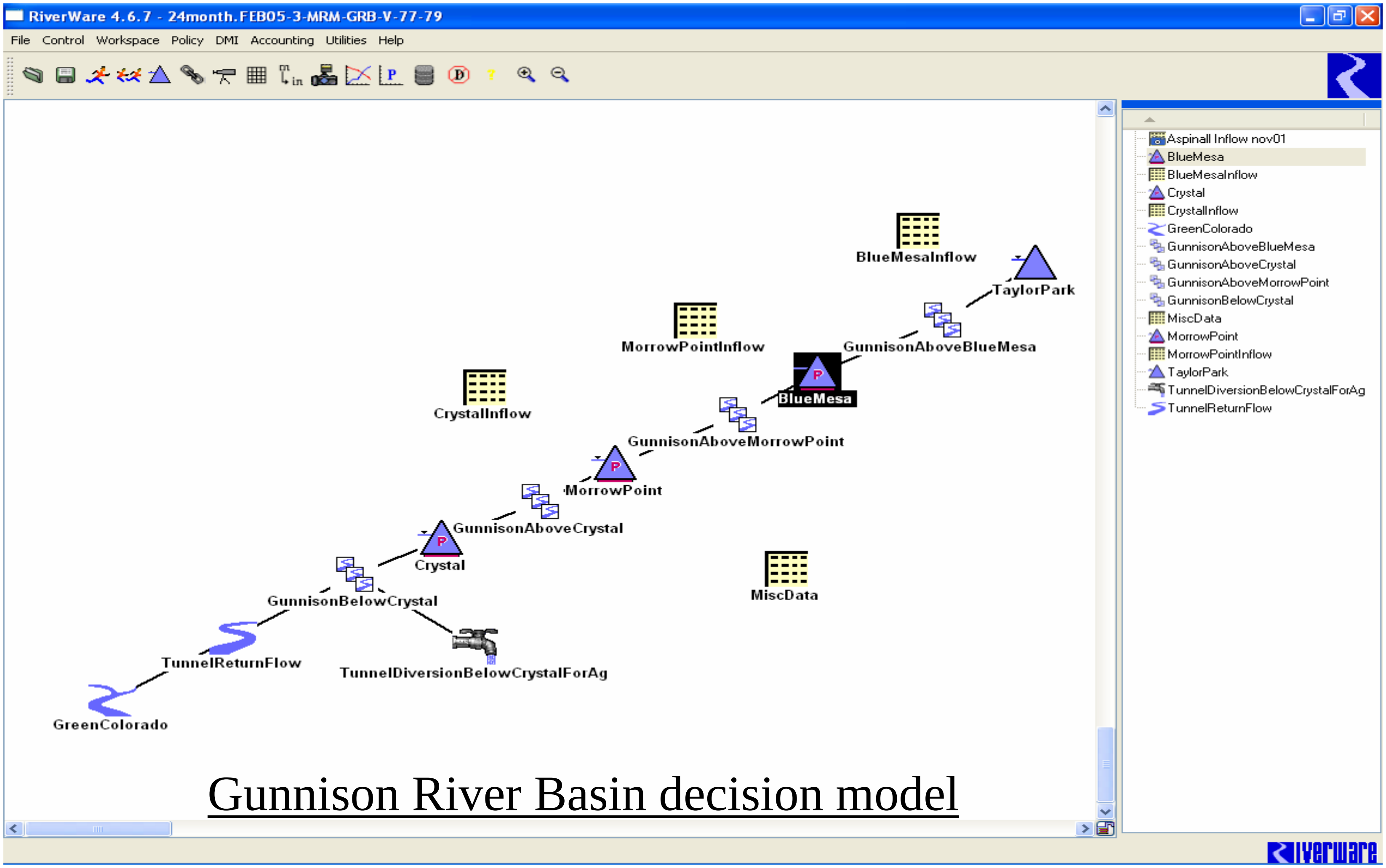
Summary

- Climate indices improve skill of long lead forecasts of spring flows starting from December 1st
- Probabilistic forecast quantifies uncertainty
- Flexible technique to issue forecast at different lead times and at multiple locations

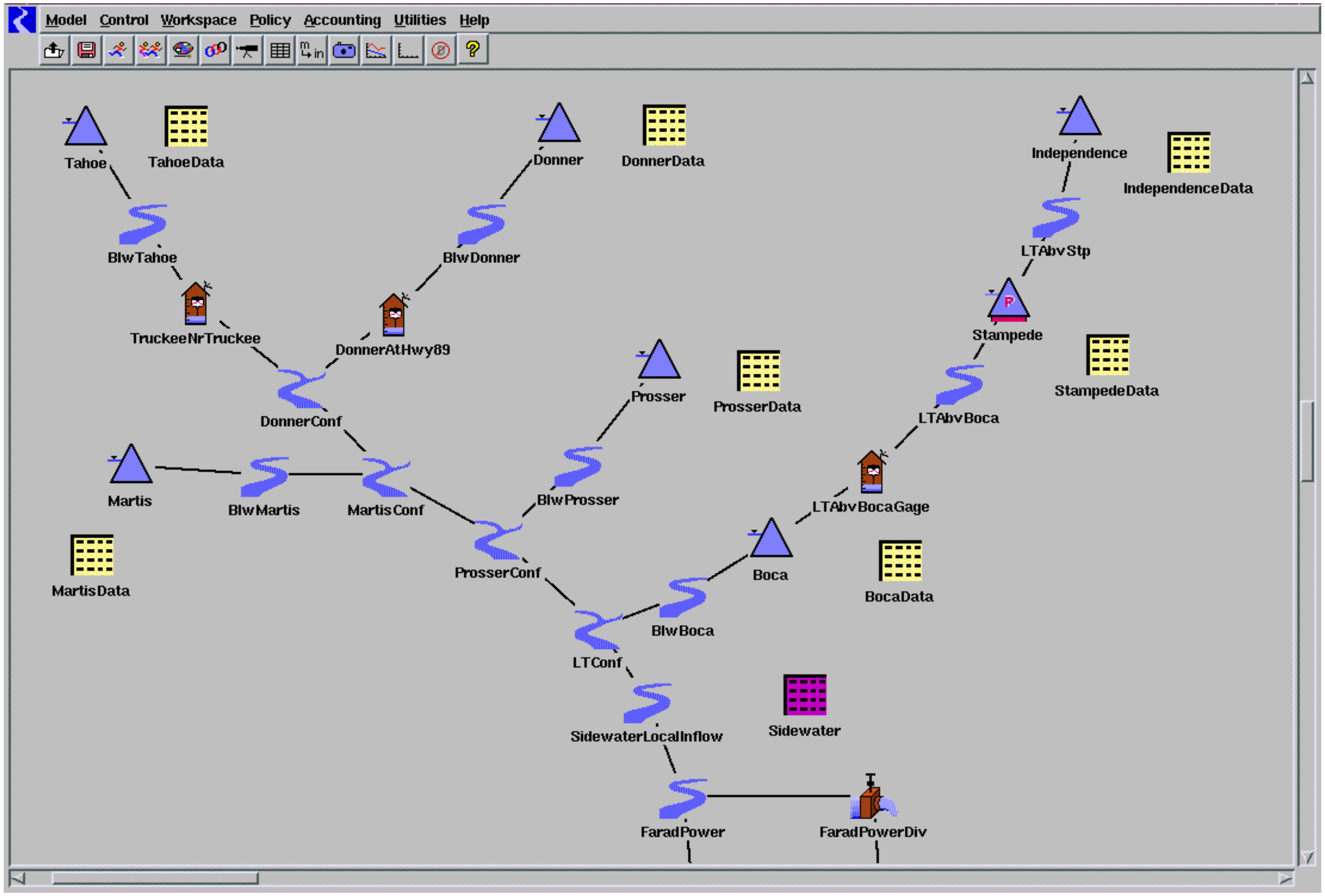
- April 1st forecast can be further refined
- Improved objective criterion
- Role of vegetation feedback, cloud cover, melt patterns, topography and shading factors

Regonda, S.K., B. Rajagopalan, M. Clark, and E. Zagona, A Multi-model Ensemble Forecast Framework: Application to Spring Seasonal Flows in the Gunnison River Basin , *Water Resources Research*, 42,W09404,2006.

For decision making run forecast through the decision model

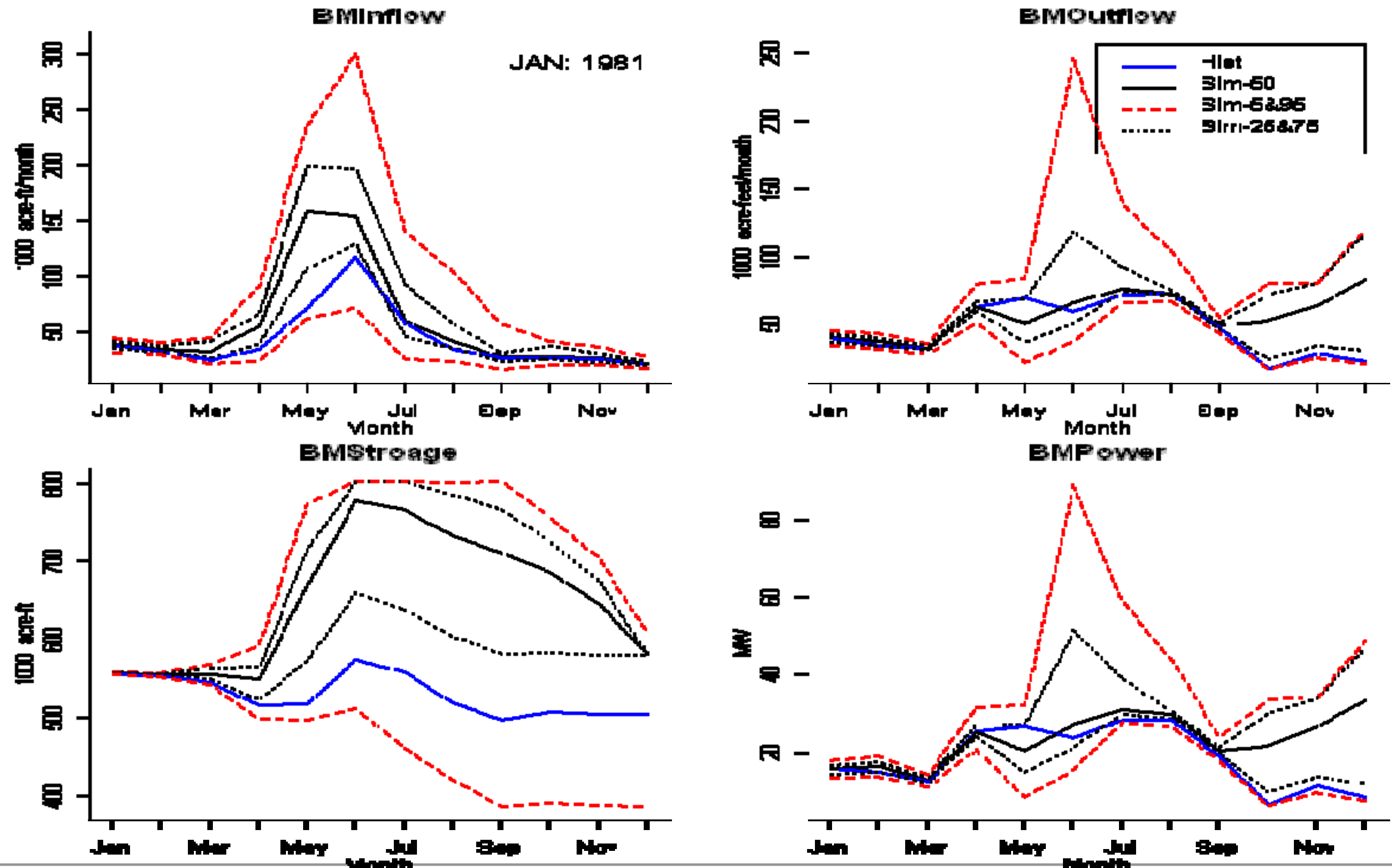


RiverWare model as an example for decision support system

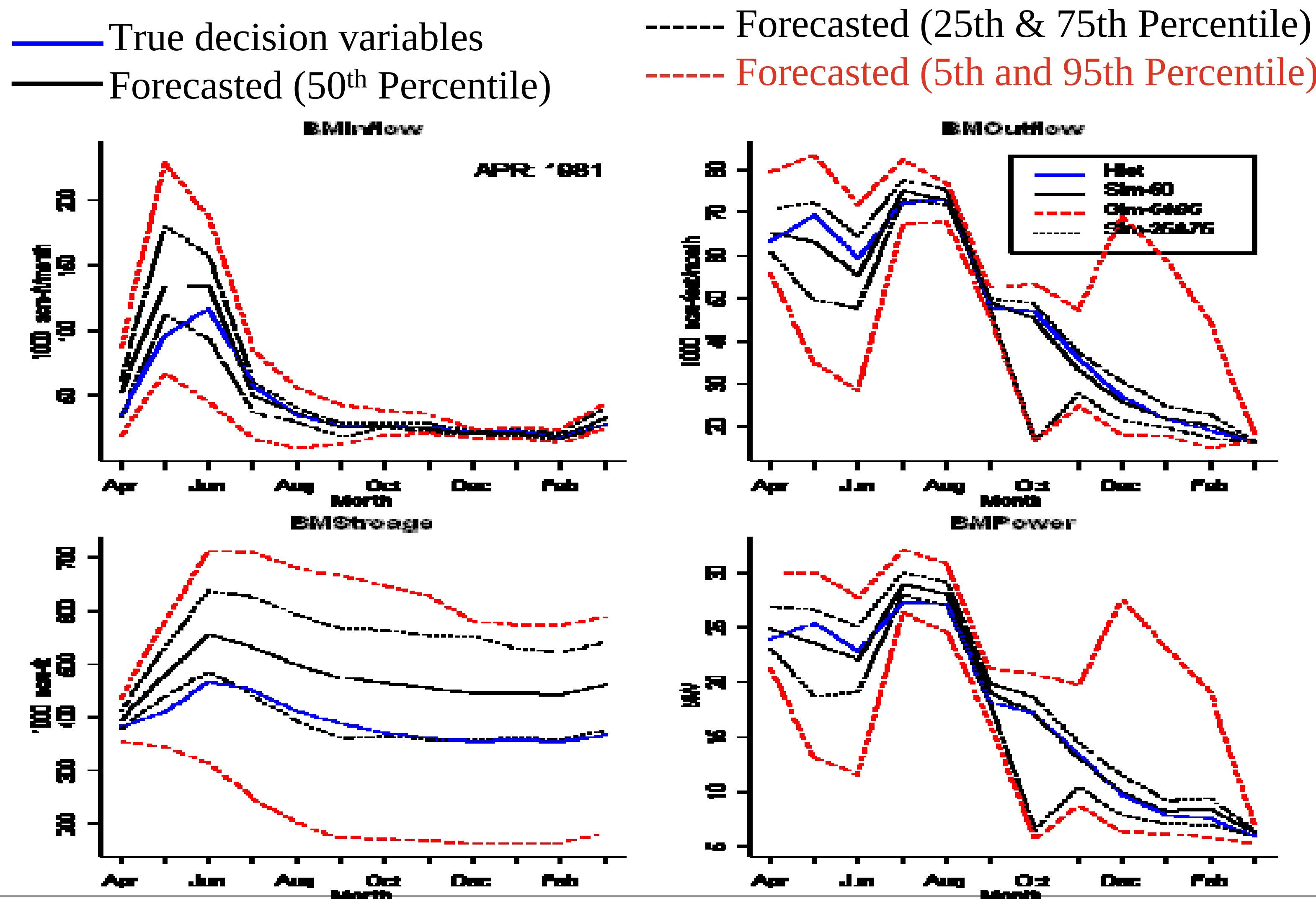


Decision variables corresponding to the January 1981 forecast (dry year)

True decision variables
Forecasted (50th Percentile)
------ Forecasted (25th & 75th Percentile)
------ Forecasted (5th and 95th Percentile)



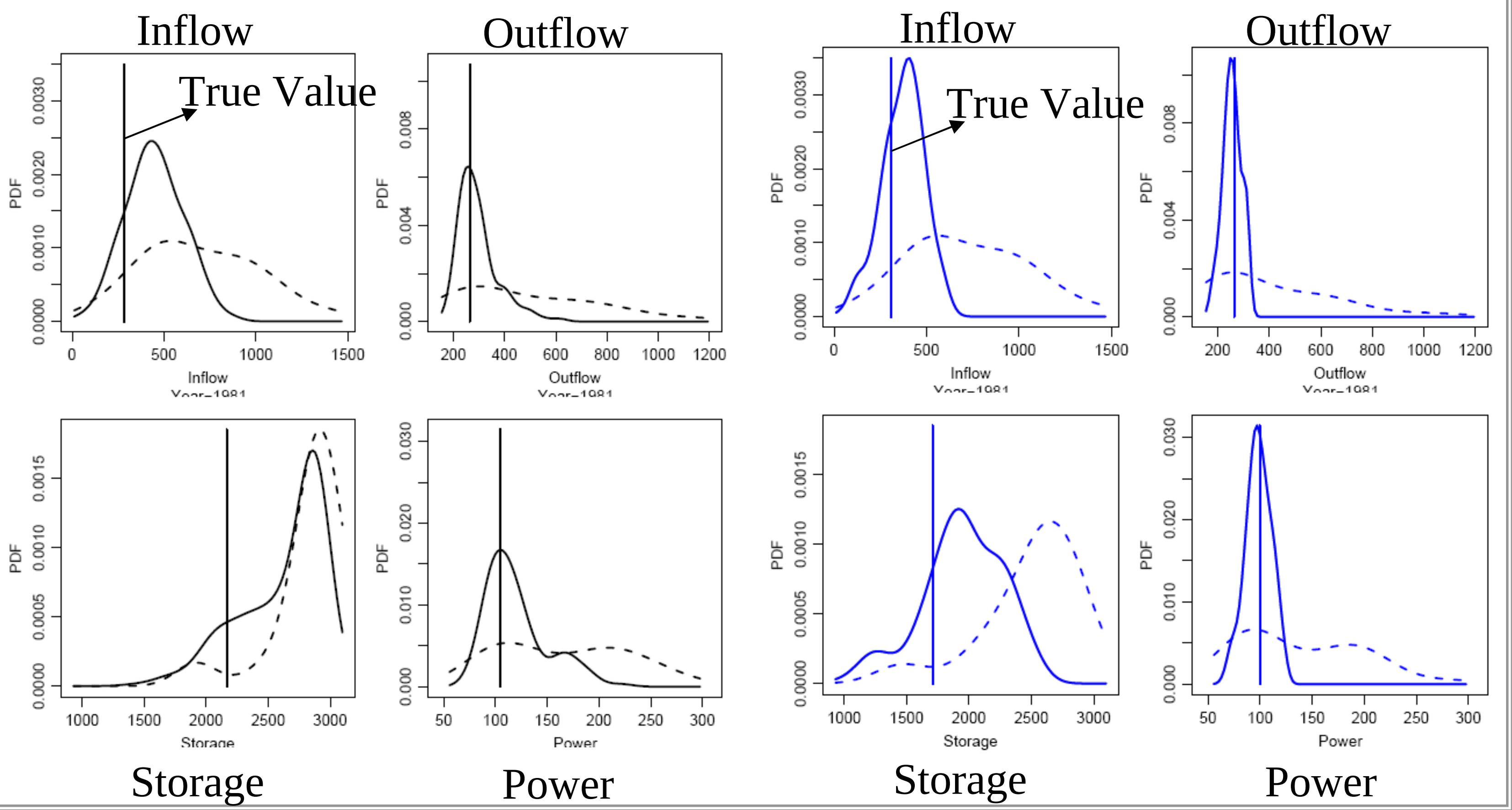
Decision variables corresponding to the April 1981 forecast (dry year)



PDF plots of spring season decision variables at Blue Mesa, 1981 (dry year)

—— Forecasted ------ Climatology

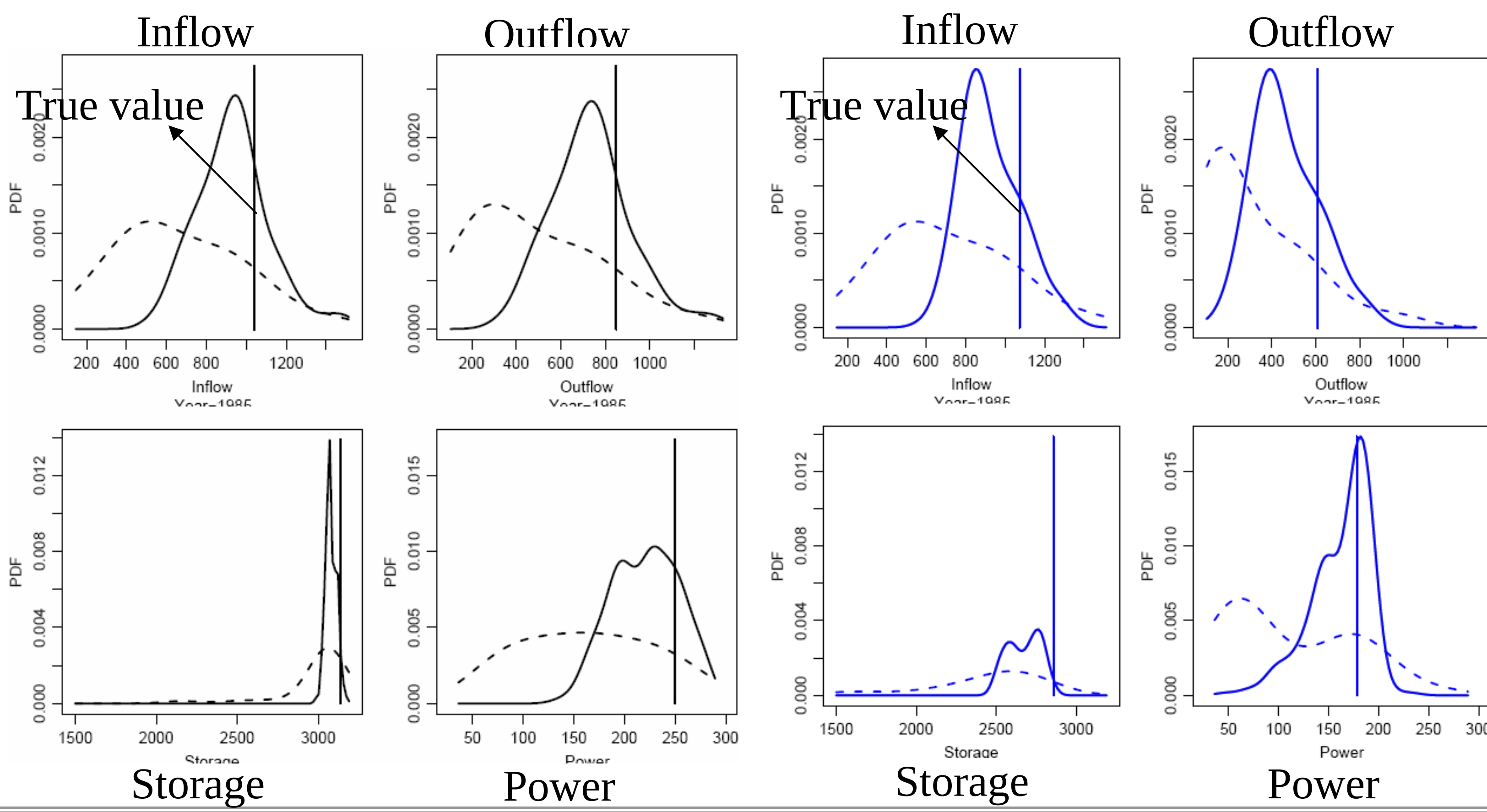
January 1st forecast April 1st forecast



PDF plots of spring season decision variables at Blue Mesa, 1985 (wet year)

——— Forecasted   ------ Climatology

January 1st forecast   April 1st forecast

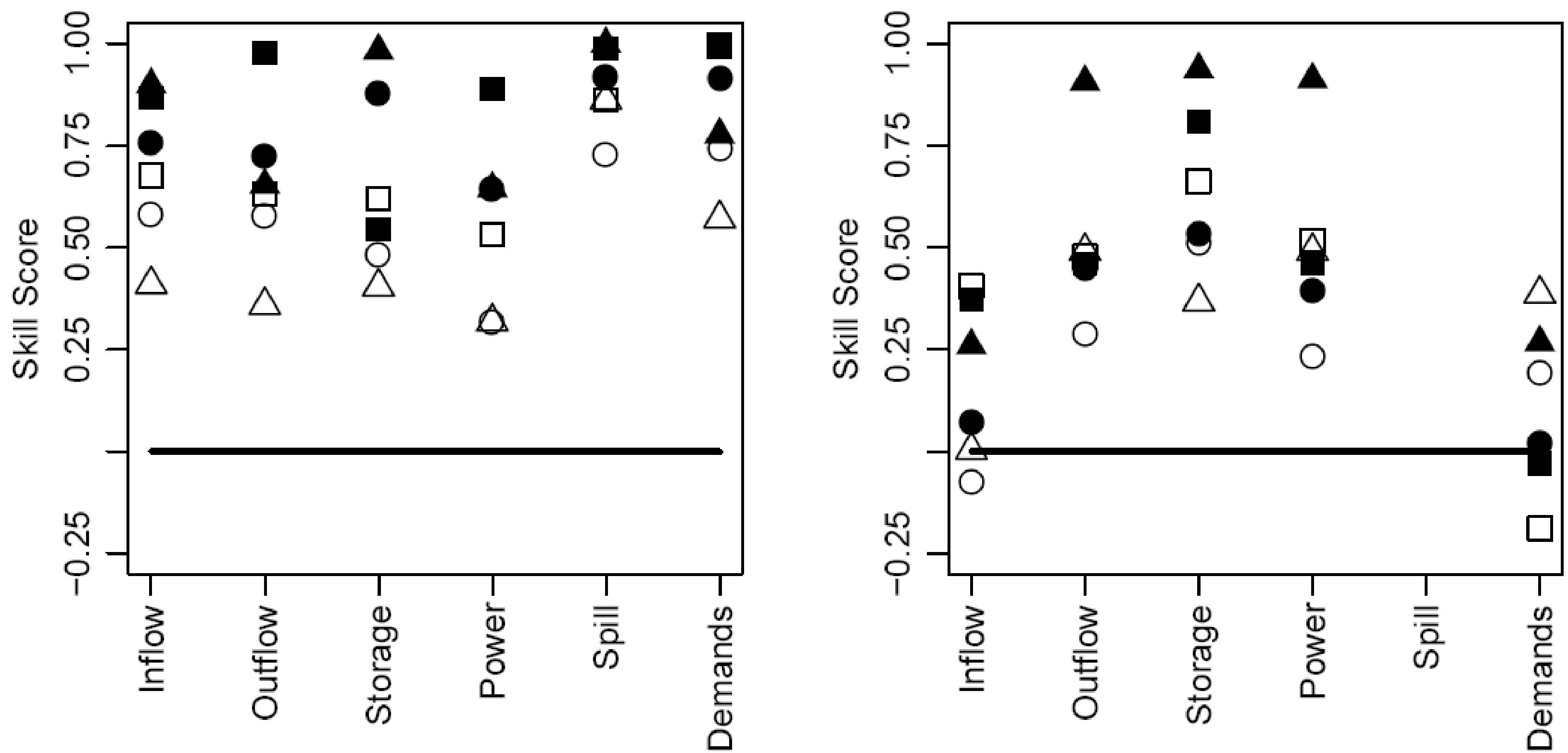


RPSS of decision variables – Blue Mesa

January 1st Forecast ○ All Years □ Wet Years △ Dry Years

April 1st Forecast ● All Years ■ Wet Years ▲ Dry Years

Spring (April – July) Winter (September – December)



Summary (decision support model)

- Skill in streamflow ensembles translated to ensembles of decision variables
- Substantial long-lead skills in the decision variables obtained.
- Nonlinearity in the skills of inflows and decision variables observed.

Regonda, S.K. (2006), Intra-annual to Inter-decadal Variability in the Upper Colorado Hydroclimatology: Diagnosis, Forecasting, and Implications for Water Resources Management, Ph.D. thesis, University of Colorado at Boulder, Colorado.