Addressing Parameter Uncertainty in Regional Forecast Basins Using Similarity Indices

Terri S. Hogue
Department of Civil and Environmental, UCLA

Kristie Franz
Department of Geological and Atmospheric Sciences, ISU

Sonya Lopez and Janet Barco
UCLA

HEPEX 3rd Annual Workshop, Stressa, Italy
June 26-28th, 2007

Introduction

Motivation

- Numerous ungauged basins worldwide which require parameter estimates
- Operational forecasting system (NWSRFS) being implemented in developing countries
- No standard procedure for regionalization (traditional = linear regression)
- Model parameters may be badly defined (equifinality)
  i.e. regression methods have limited application (over-parameterized models!)
- Typically *limited* number of catchments available (most published work uses very large data sets)

Objectives

- Find parsimonious method to regionalize parameters for NWSRFS operational rainfall-runoff model (over-parameterized?)
- Start with homogenous region to evaluate method (NWS operational basins)
- Implement more rigorous testing of model (parameter) performance
  → linear regression and advanced regionalization methods use historical simulations and related statistics (DRMS or NSE)

General Approach

1) Selection of regional forecast basins (~small sample size!)

“Ecoregion” classification – Omernik, 1987

2) Development of similarity indices

- Precipitation (P), PE, Temp, PE/P ratio, Q/P, Area, etc.

3) Calibration of all regional basins with sufficient data (~13)

- MACS method (Hogue et al., 2000) and SCE (Duan et al., 1992)

4) Selection of “transferable” parameter sets

- Climate Neighbor (similarity indices)
- Nearest Neighbor
- Median value of regional parameter sets
- Mean value of regional parameter sets

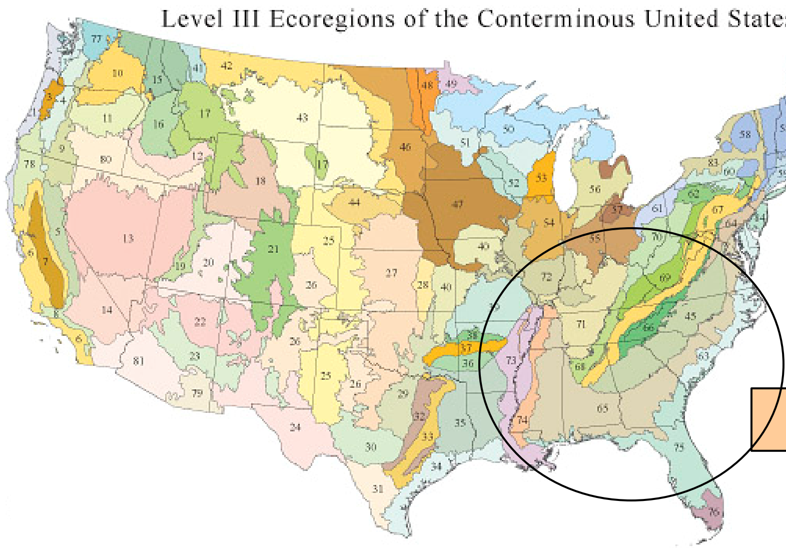
Comparison to:

- Calibration (MACS)
- MOPEX parameter set (~RFC)

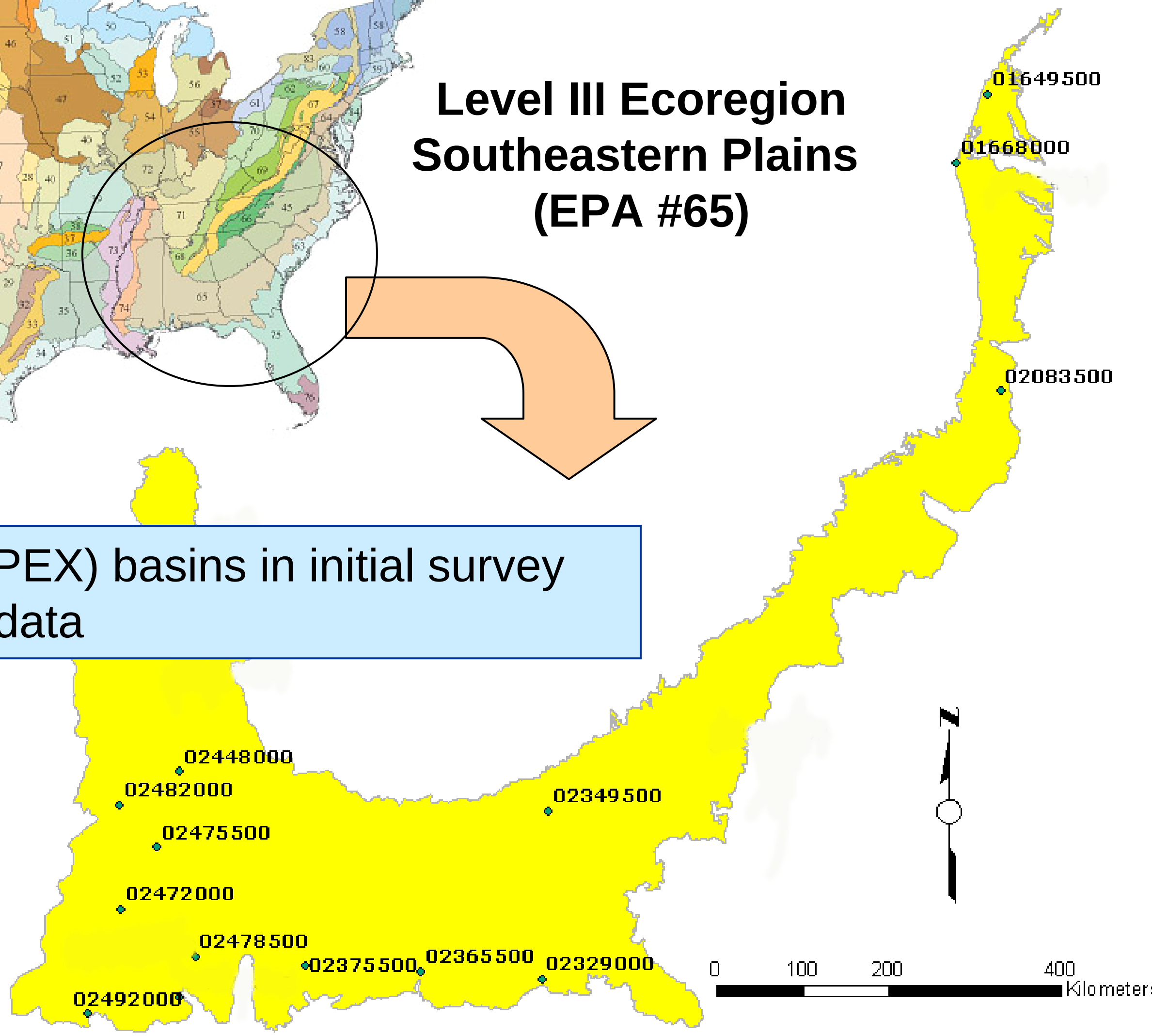
5) Evaluation

Traditional statistics and hydrographs (visual)
Hindcasting for “forecast performance”– RPSS (daily and weekly) (Franz et al., 2003)

Southeast Region Study Basins (SERFC and LMRFC)



Level III Ecoregion
Southeastern Plains
(EPA #65)



19 RFC (and MOPEX) basins in initial survey
13 with sufficient data

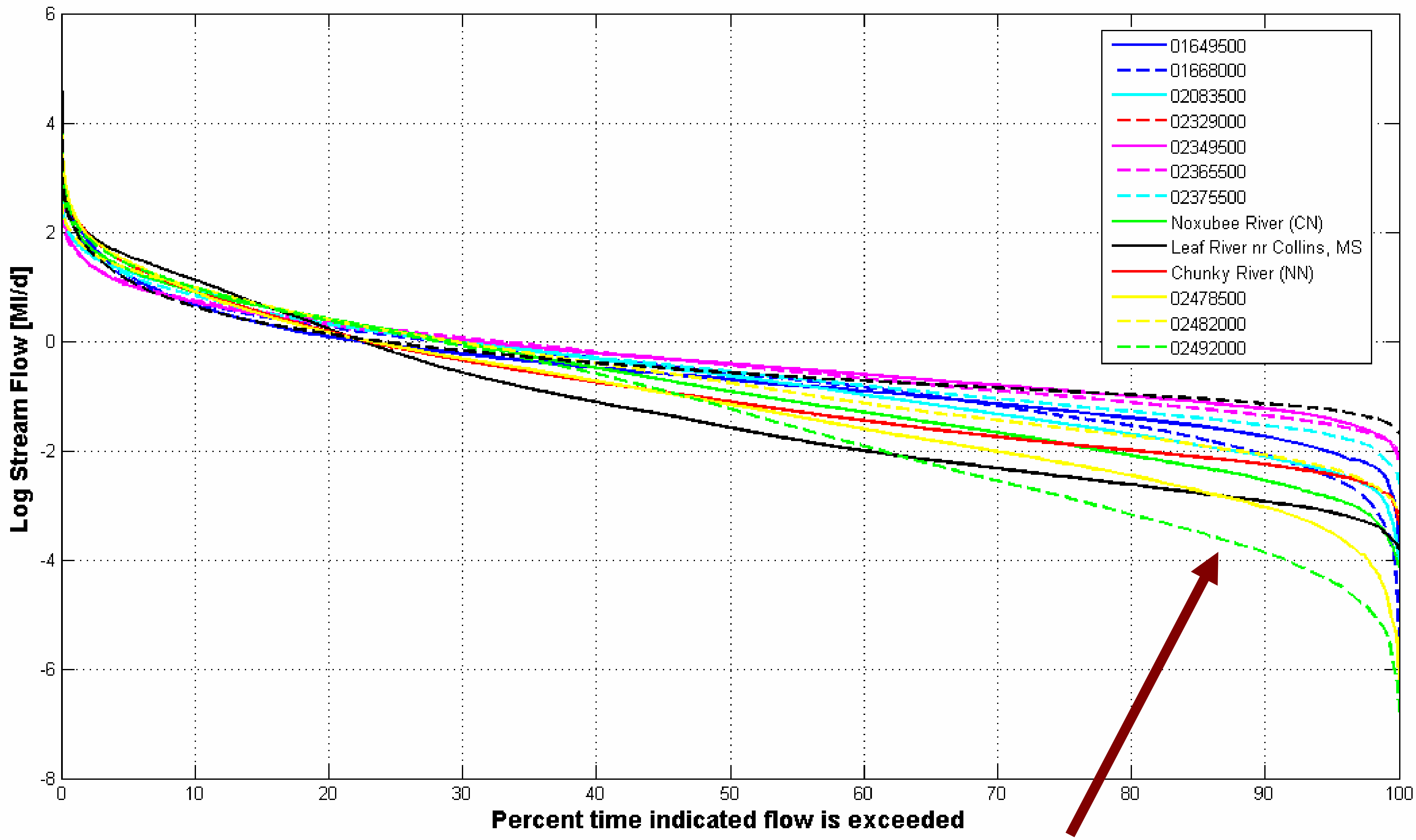
Ecoregions Level III (USEPA, 2007, Omernik, 1987)

Regional Basins – General Climatology

Site ID	Site Name	Size [mi^2]	Annual Ppt (mm)	Annual Q (mm)	Annual PE (mm)	RO Ratio (Q/P)	Evap. Ratio (PE/P)
01649500	**North East Branch Anacostia River at Riverdale, MD**	73	1058	414	965	.384	.937
01668000	**Rappahannock River Near Fredericksburg, VA**	1596	1034	357	921	.336	.914
02083500	**Tar River at Tarboro, NC**	2183	1145	321	993	.284	.886
02329000	**Ochlockonee River Nr Havana, FL**	1140	1318	323	1099	.233	.861
02349500	**Flint River at Montezuma, GA**	2900	1227	408	1039	.327	.868
02365500	**Choctawhatchee River at Caryville, FL**	3499	1386	543	1059	.384	.784
02375500	**Escambia River Near Century, FL**	3817	1448	547	1059	.369	.752
02448000	**Noxubee River at Macon, MS**	768	1381	489	1054	.337	.793
02472000	**Leaf River Nr Collins, MS**	**743**	**1444**	**508**	**1058**	**.338**	**.759**
02475500	**Chunky River Nr Chunky, MS**	369	1409	461	1052	.315	.771
02478500	**Chickasawhay River at Leakesville, MS**	2690	1441	487	1054	.326	.752
02482000	**Pearl River at Edinburg, MS**	904	1386	458	1052	.315	.787
02492000	**Bogue Chitto River near Bush, LA**	1213	1551	562	1069	.354	.713
	Mean	**1684**	**1325**	**453**	**1036**	**.331**	**.814**
	Std Dev	**1217**	**160**	**81**	**.071**	**.041**	**.071**

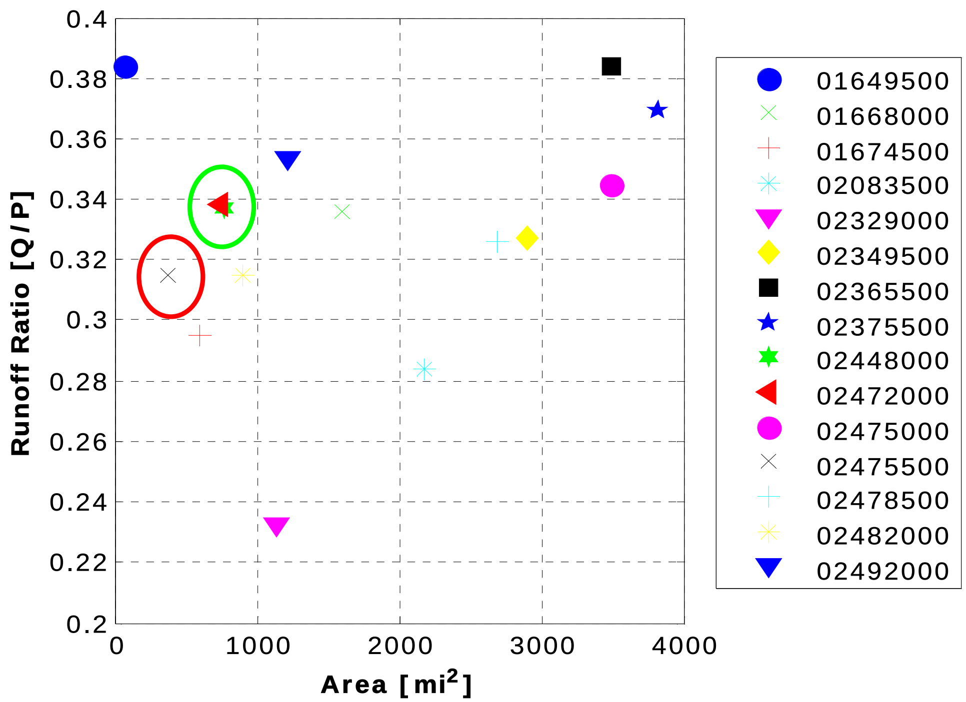
- **Homogeneous Climate – Ppt, PE, RO and PE ratios**
- **Largest Variation – Size, Q**

Regional Basins - Flow Duration Curves

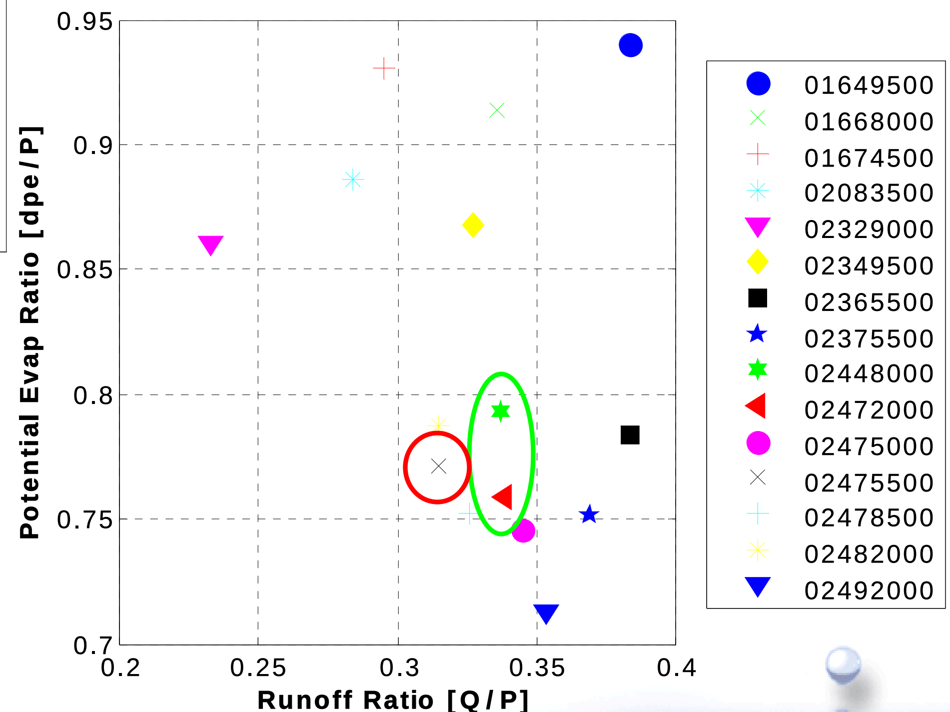


Significant variation in response during low-flow periods

Similarity Indices



- 7 indices evaluated:
  - f (Hydroclimatology, Area)
  - all from MOPEX database
- "Climate Neighbor" selected Noxubee River (24755)

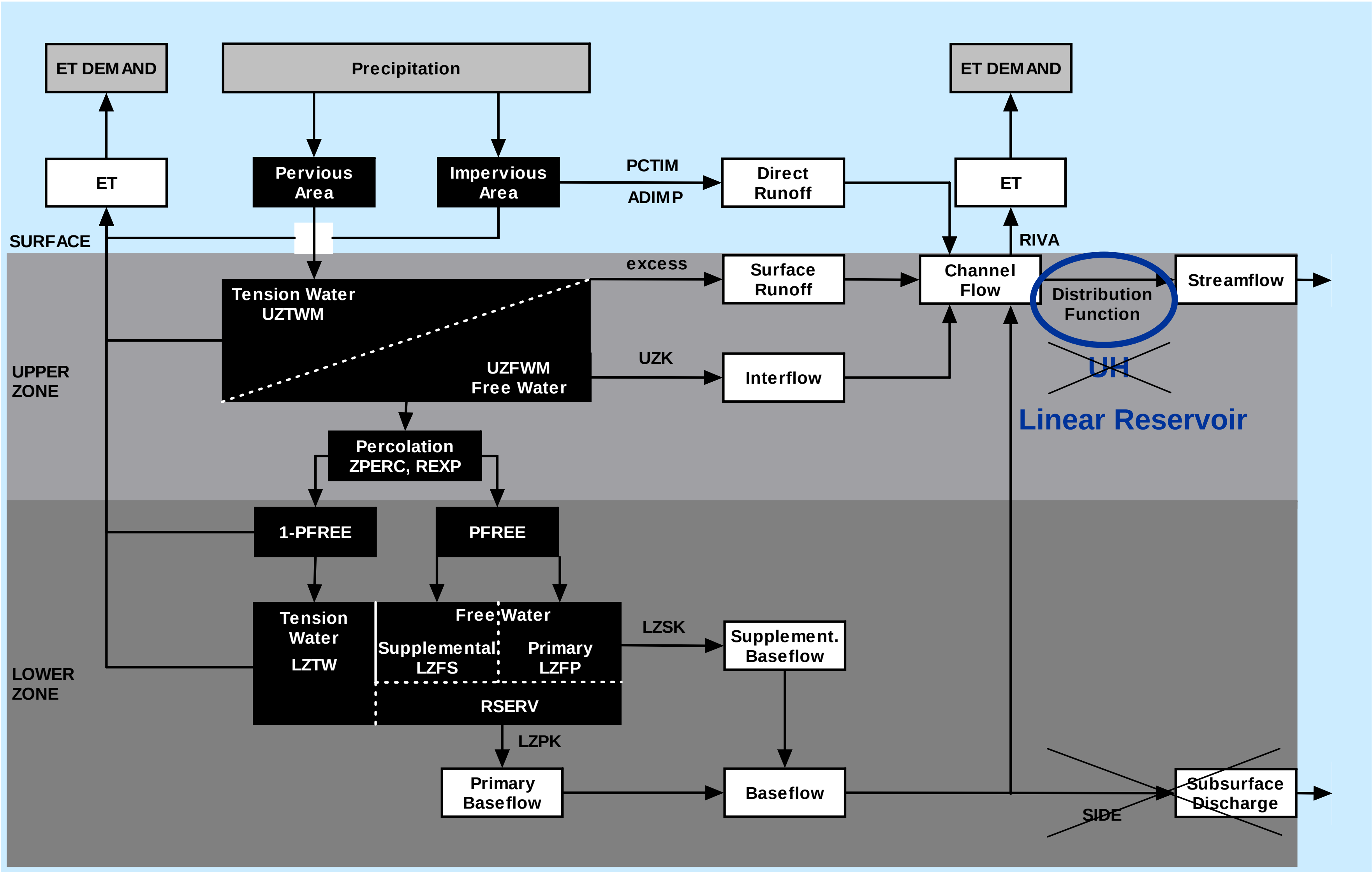


◀ Leaf River (24720)

★ Climate Neighbor (24480)

X Nearest Neighbor (24755)

Sacramento Soil Moisture Accounting Model (SAC-SMA)

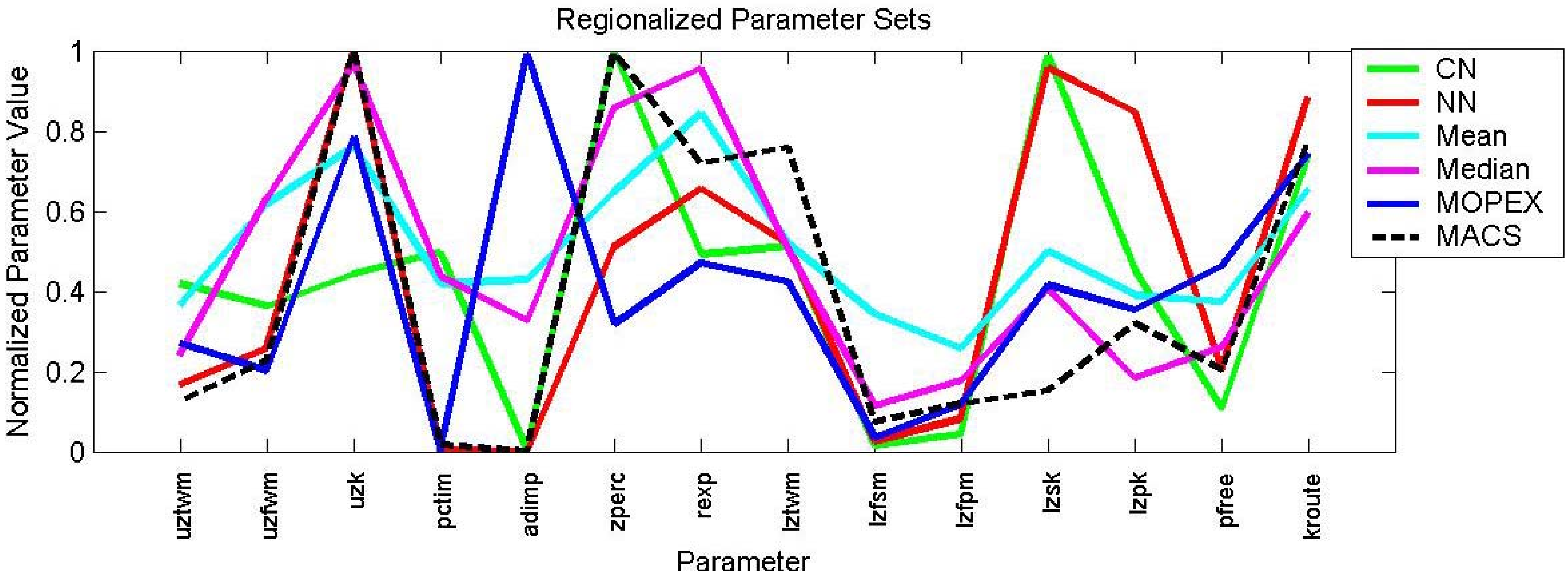


➢16 Model parameters ~ 13 typically calibrated

Calibrated and "Regional" Parameter Sets

13 basins calibrated with automatic algorithm (MACS and SCE)

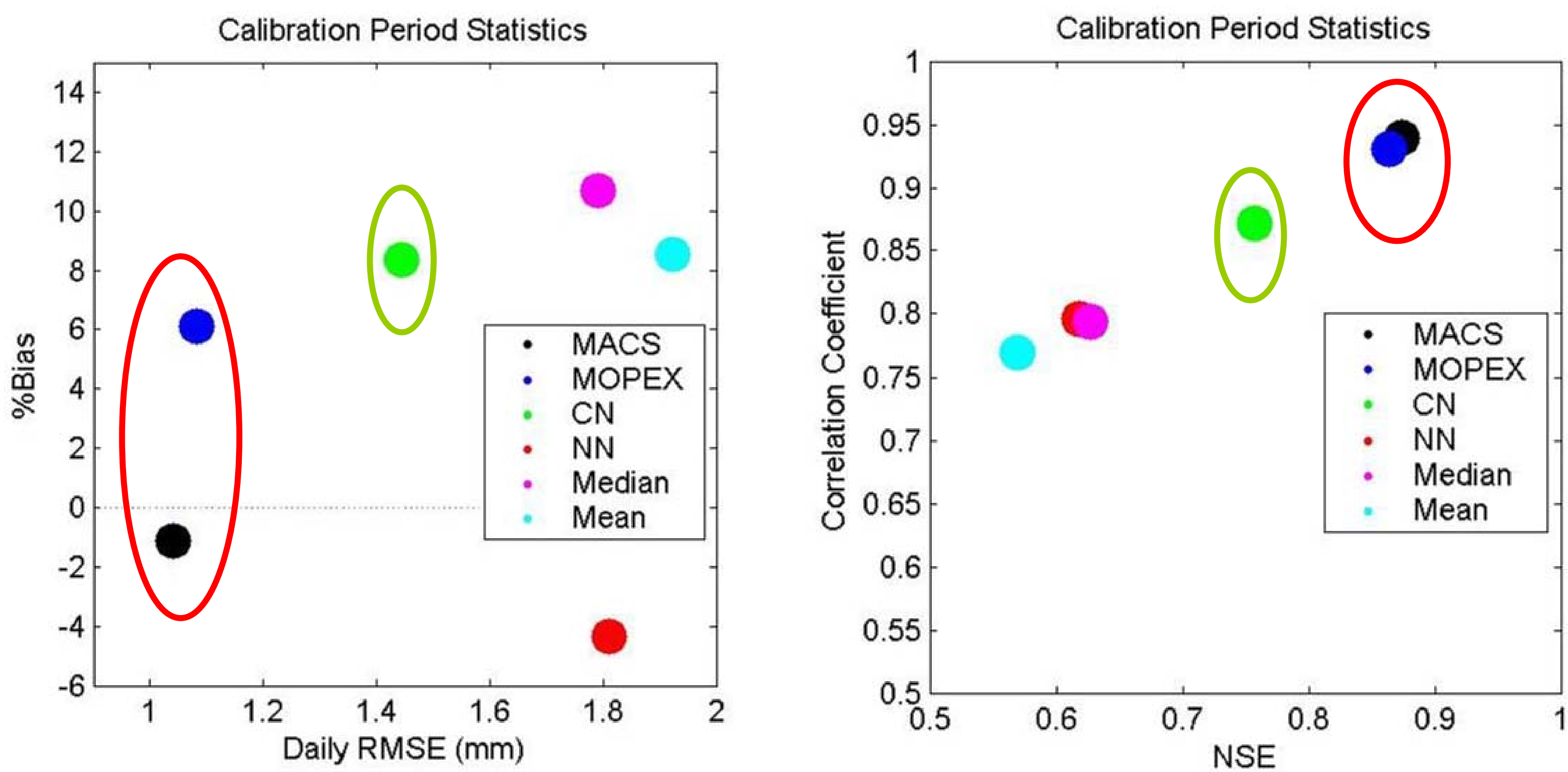
- Six parameter sets selected for evaluation on "test basin" → Leaf River
- Climate Neighbor (CN), Nearest Neighbor (NN), Mean and Median of 12 regional basins, MOPEX on Leaf, MACS on Leaf



- ➢ **Better consistency in UZ and LZ storage volumes**
- ➢ **Recession parameters more variable (esp. lower zone).**
- ➢ **Percolation and impervious parameters highly variable**
- ➢ **Routing parameter (kroute) is consistent between sets**

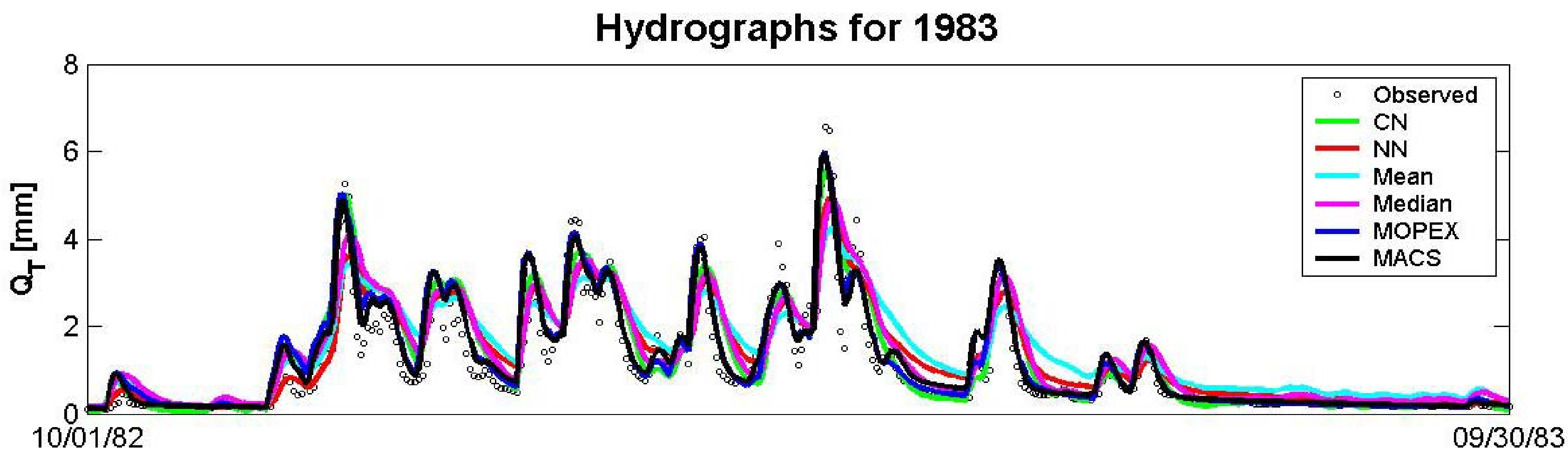
Testing of Parameter Sets - Calibration Period

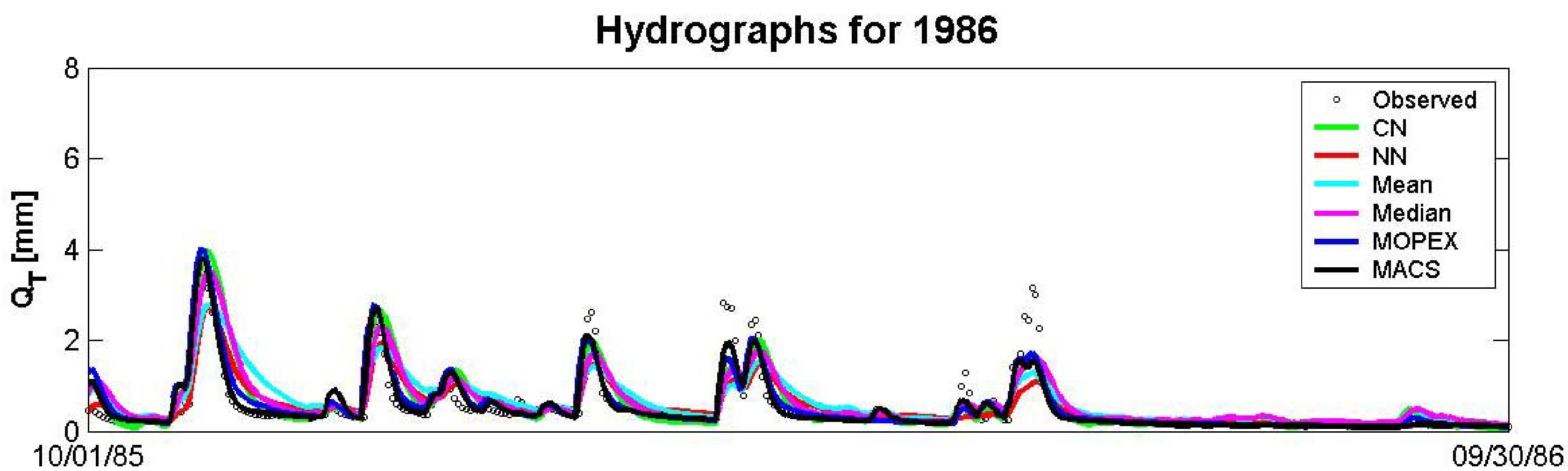
Leaf River (Collins, MS) is initial "ungauged test basin"



- **MACS and MOPEX best performance overall (explicit calibrations)**
- **Climate neighbor (CN) next best performance**
- **Mean / Median show similar performance**

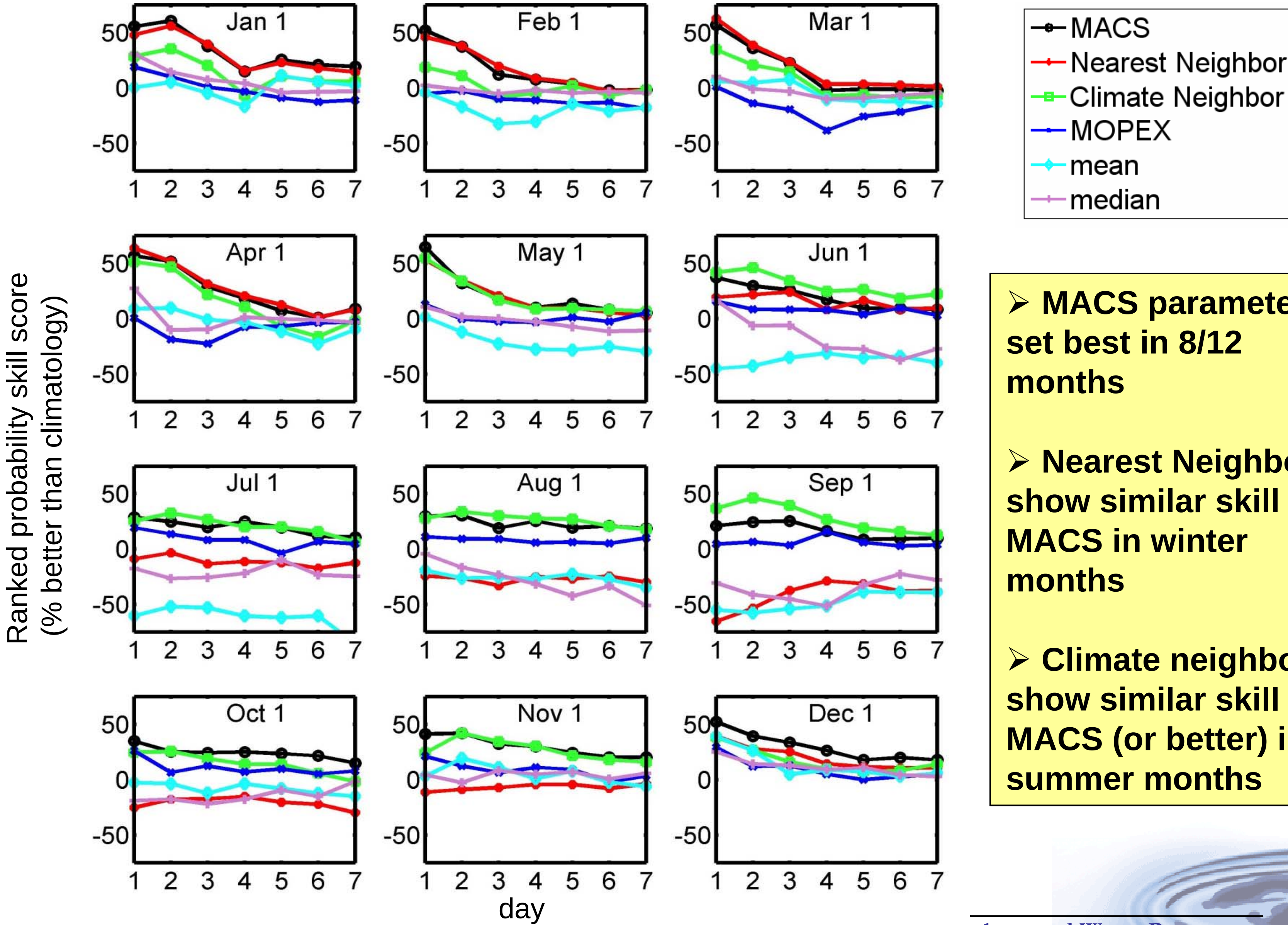
Ensembles – Calibration Period





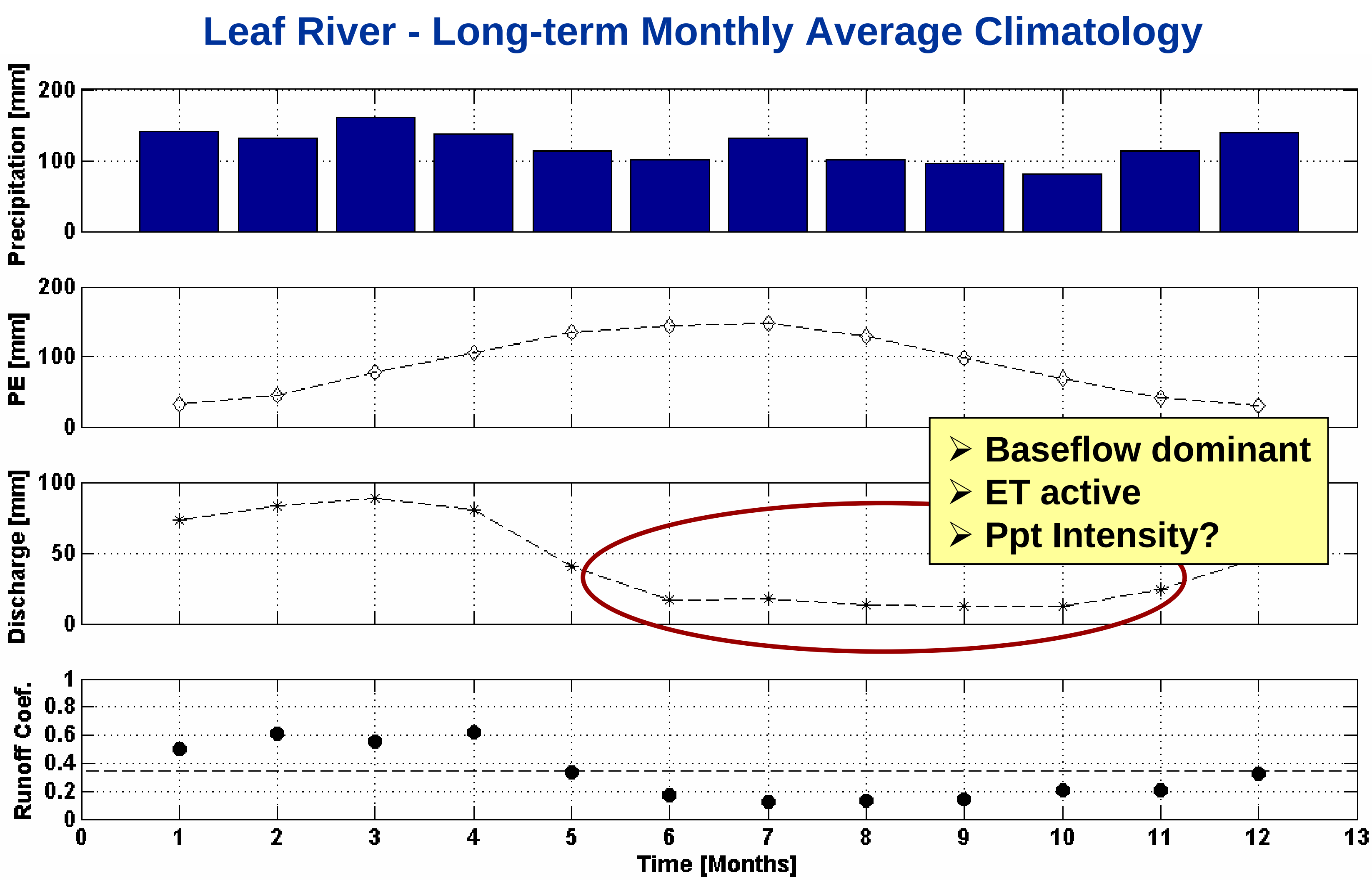
- **MACS, MOPEX, CN simulate most high flows**
- **Mean does poorly on recessions**

Skill Score for Mean Daily Discharge Hindcasts

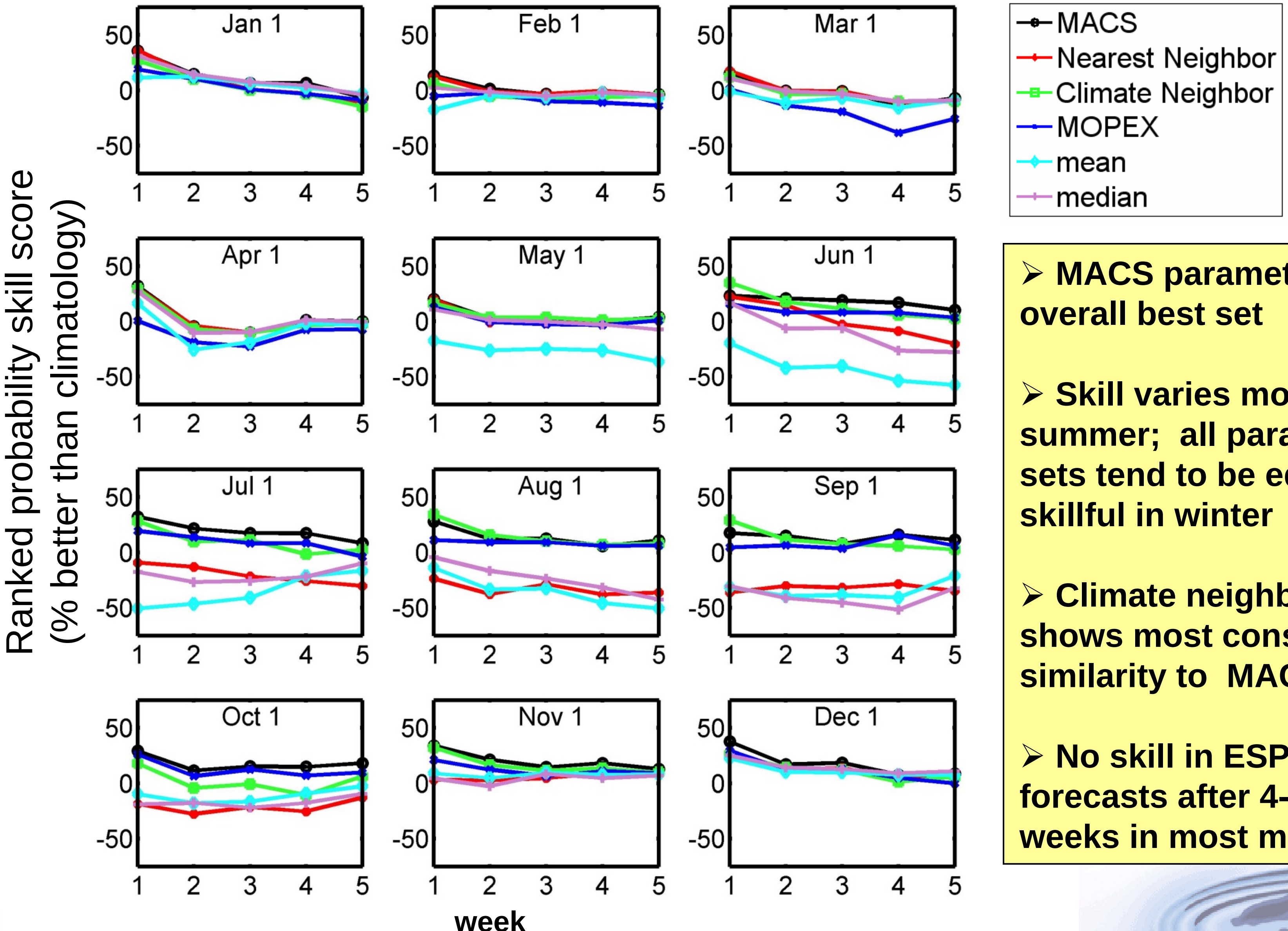


- ➢ **MACS parameter set best in 8/12 months**
- ➢ **Nearest Neighbor show similar skill to MACS in winter months**
- ➢ **Climate neighbor show similar skill to MACS (or better) in summer months**

Why the Poorer Performance in Summer?



Skill Score for Mean Weekly Discharge Hindcasts



- **MACS parameters overall best set**
- **Skill varies most in summer; all parameter sets tend to be equally skillful in winter**
- **Climate neighbor shows most consistent similarity to MACS**
- **No skill in ESP forecasts after 4-5 weeks in most months**

Summary and Future Plans

Climate Neighbor "best transferable parameter set"

Nearest neighbor, Median, Mean – Less satisfactory overall

Hindcasting provides valuable information on "operational" performance

Poorer model performance during summer period

- Low flow regimes, ET interactions, Ppt Intensity?

Daily forecasts show better overall performance

No skill in forecasts after ~ 5 weeks for all sets

Uncertainty in forcing – but some parameter sets handle this better

Further develop similarity indices

- Incorporate physiography – slope, watershed length, channel density, sinuosity, etc.
- GIS platform with MOPEX/EDNA (USGS) database

Further refine evaluation methods

Quantifying uncertainty in regionalized sets

Use traditional statistics and hydrographs (visual)

Continue hindcasting evaluation - RPSS (daily and weekly)

- Add reliability, discrimination statistics

Selection of "snow-dominated" regional forecast basins

- Western U.S. contains multi-tiered forecast (semi-distributed) systems

Coupling of MOPEX and HEPEX

- Involvement with HEPEX Hydrological Uncertainty Test-bed Project

Source: NOAA

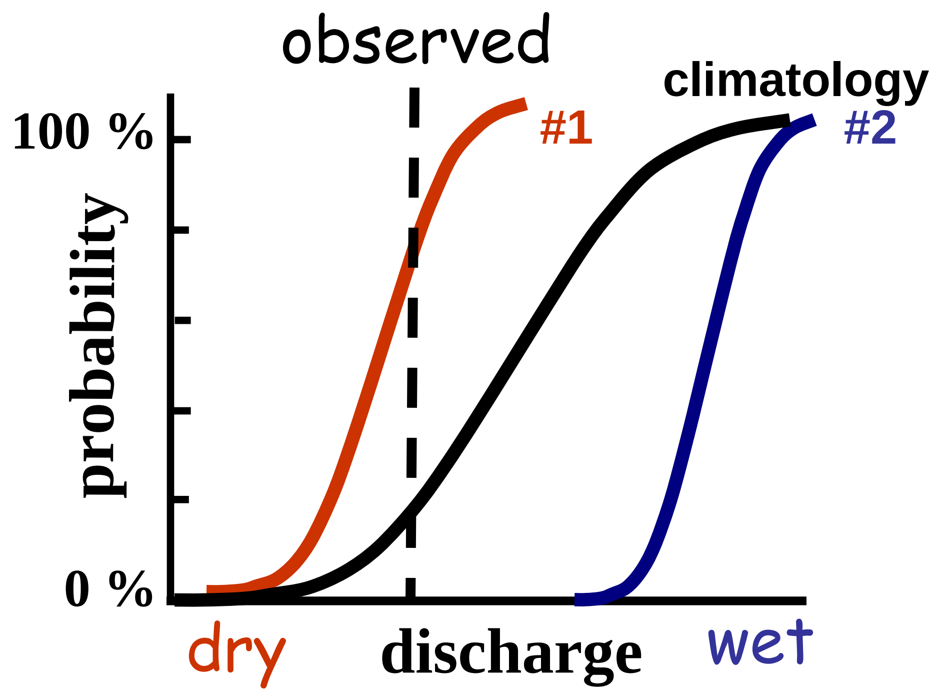
Ranked Probability Skill Score

- overall accuracy measure
- increasingly penalizes forecast (F) for being far from observed (O)

$$RPS = \sum_{m=1}^{J} (F_m - O_m)^2$$

$$O_m = \sum_{j=1}^{m} o_j, m = 1, \ldots J$$

$$F_m = \sum_{j=1}^{m} f_j, m = 1, \ldots J,$$

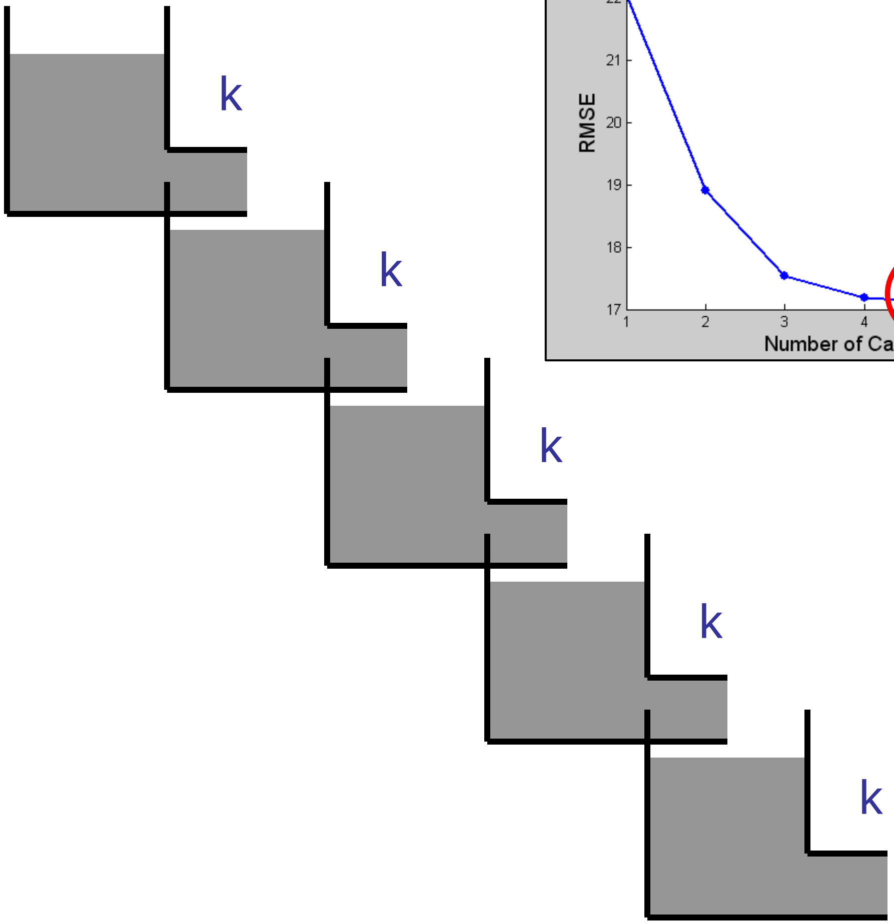


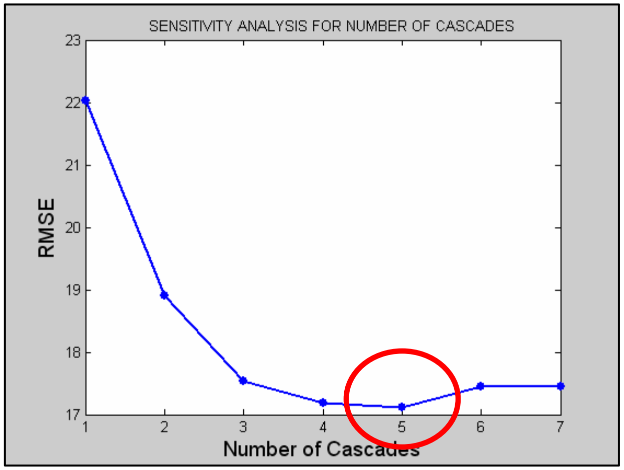
$$RPSS = 1 - \overline{RPS}_{forecast} \Big/ \overline{RPS}_{reference} \; x100\%$$

International Version – Cascade of Linear Reservoirs (5)

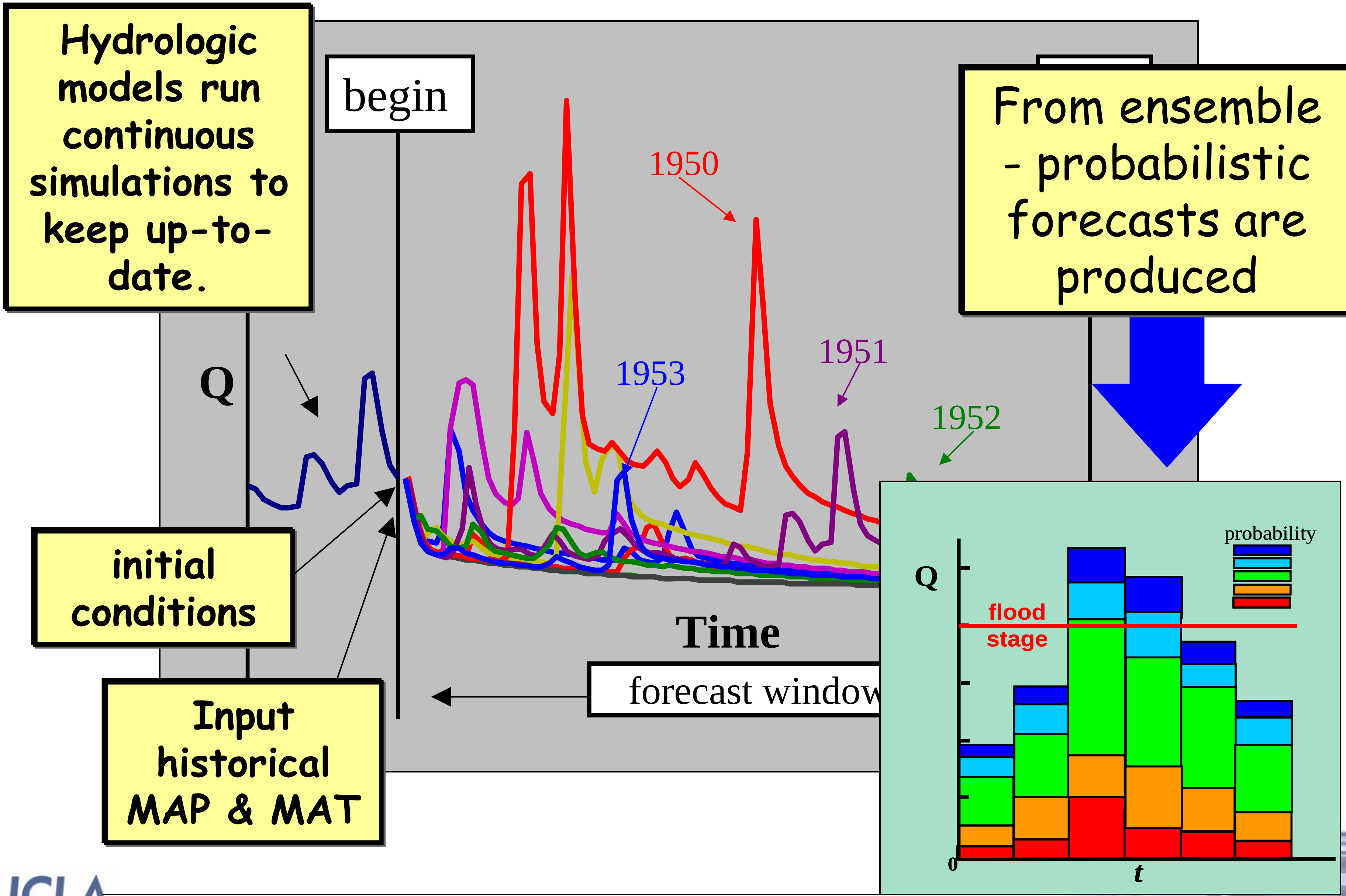
What distribution function to use?

- UH
- Synthetic UH
- Linear Reservoir
- **Nash Cascade**
- Kinematic Wave





Verification using ESP (Ensemble Streamflow Prediction)



Noxubee Climatology (climate neighbor)