OHD Prototype
Ensemble Pre-Processor (EPP)

Limin Wu[1,2] and Julie Demargne[1,3]

Acknowledgments
John Schaake, James Brown, DJ Seo

1. NOAA/NWS Office of Hydrologic Development
2. TCoombs & Associates, LLC
3. University Corporation for Atmospheric Research

Outline

- Capabilities
- Methodology
- Results
- Planned work

Capabilities

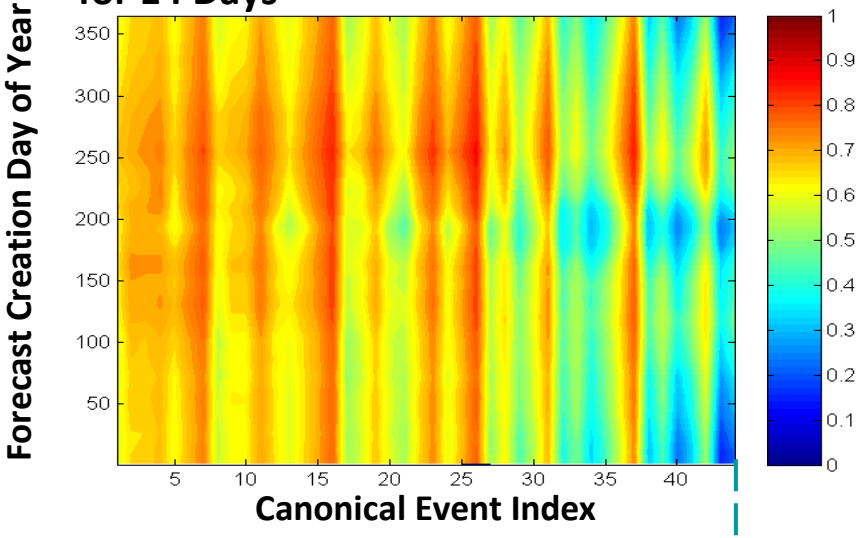
The EPP can generate short- to long-term precipitation and temperature ensembles at RFC basin scale from the following sources:

- HPC/RFC single-valued forecasts (up to 5 days) to capture skill added by RFC forecaster
- Frozen version (~circa 1998) of the GFS (up to 14 days)
- CFS forecasts (up to 8 months)

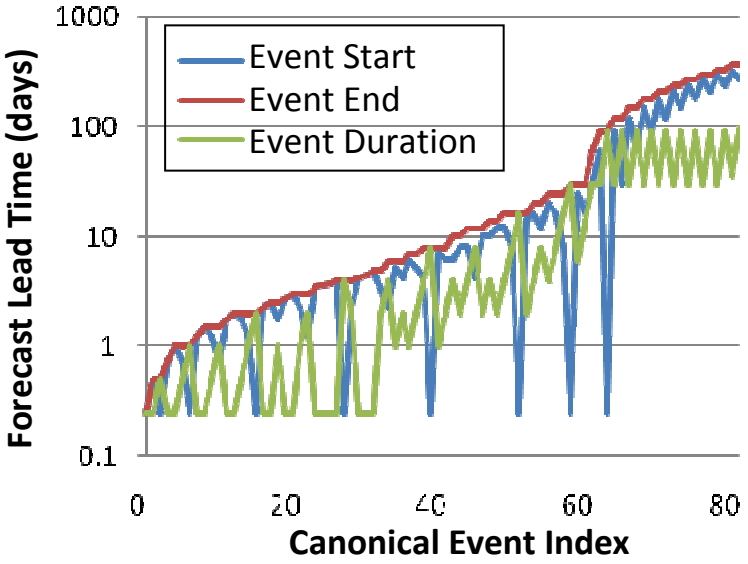
The method could be applied to any single-valued forecast for which a multi-year archive is available

Capabilities – Skill at multiple temporal scales

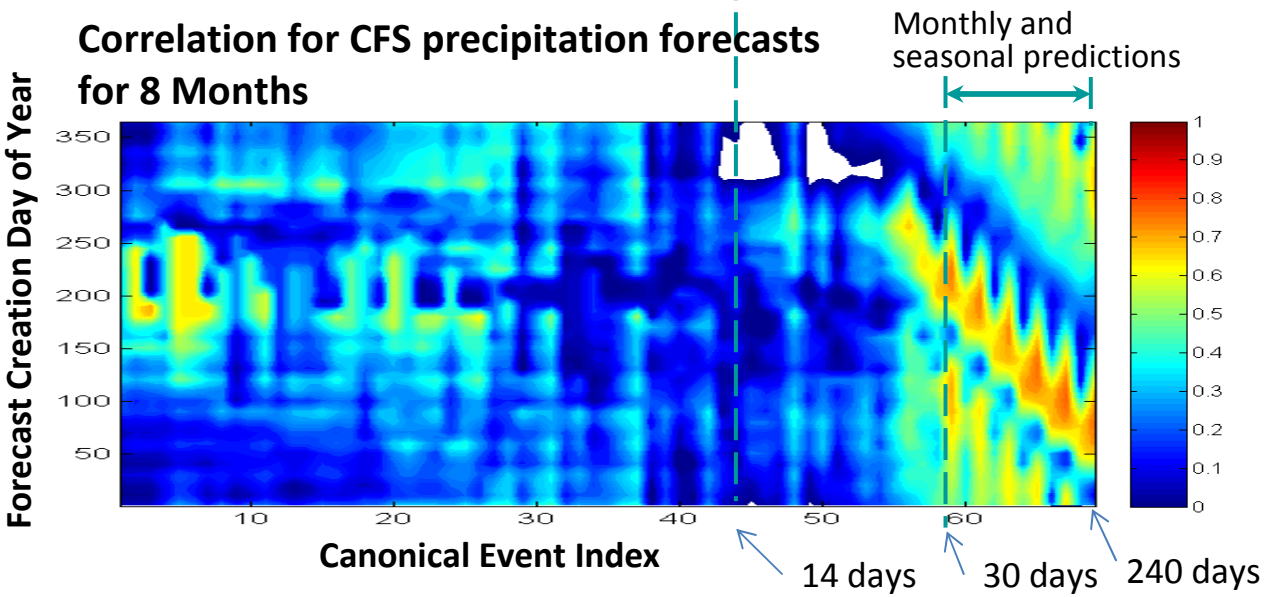
Correlation for GFS precipitation forecasts for 14 Days



Canonical Precipitation Events



Correlation for CFS precipitation forecasts for 8 Months



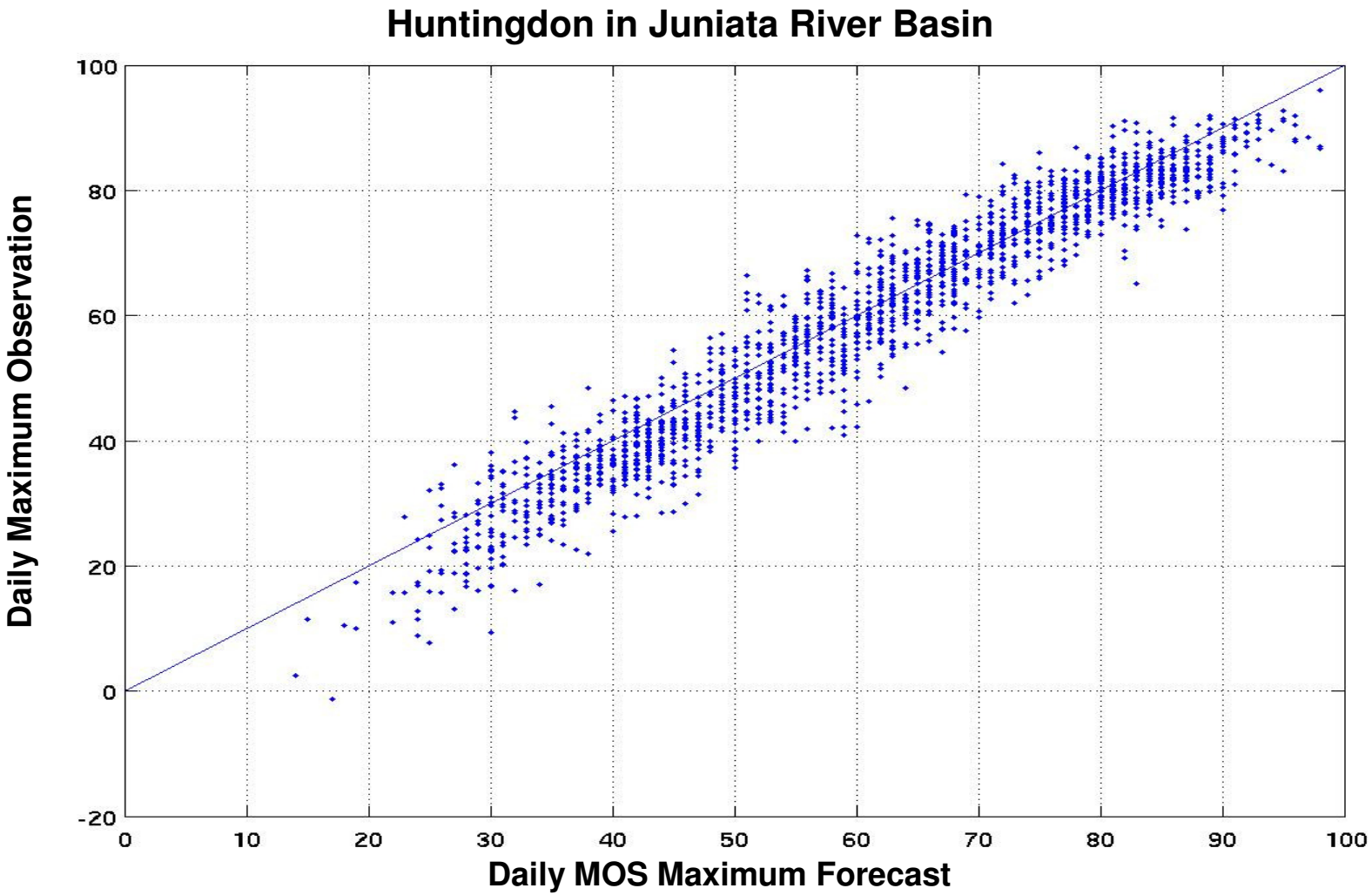
Weather and climate forecasts are projected to multiple temporal scales called Canonical Events and then correlated with corresponding observations at each temporal scale

Methodology

- Model the joint probability distribution between the single-value forecast and the verifying observation (calibration)
  - Quantify the uncertainty in the single-value forecasts
  - Capture skill in the single-value forecasts
- Sample the conditional probability distribution conditional given the single-value forecast (ensemble generation)
- Apply Schaake Shuffle to the sample from the conditional distribution to obtain the forecast ensemble that behaves similarly to the historical ensemble in space-time variability
- Use multiple aggregate "canonical" events to capture and preserve the forecast skill at longer temporal scales (e.g. several days/weeks)

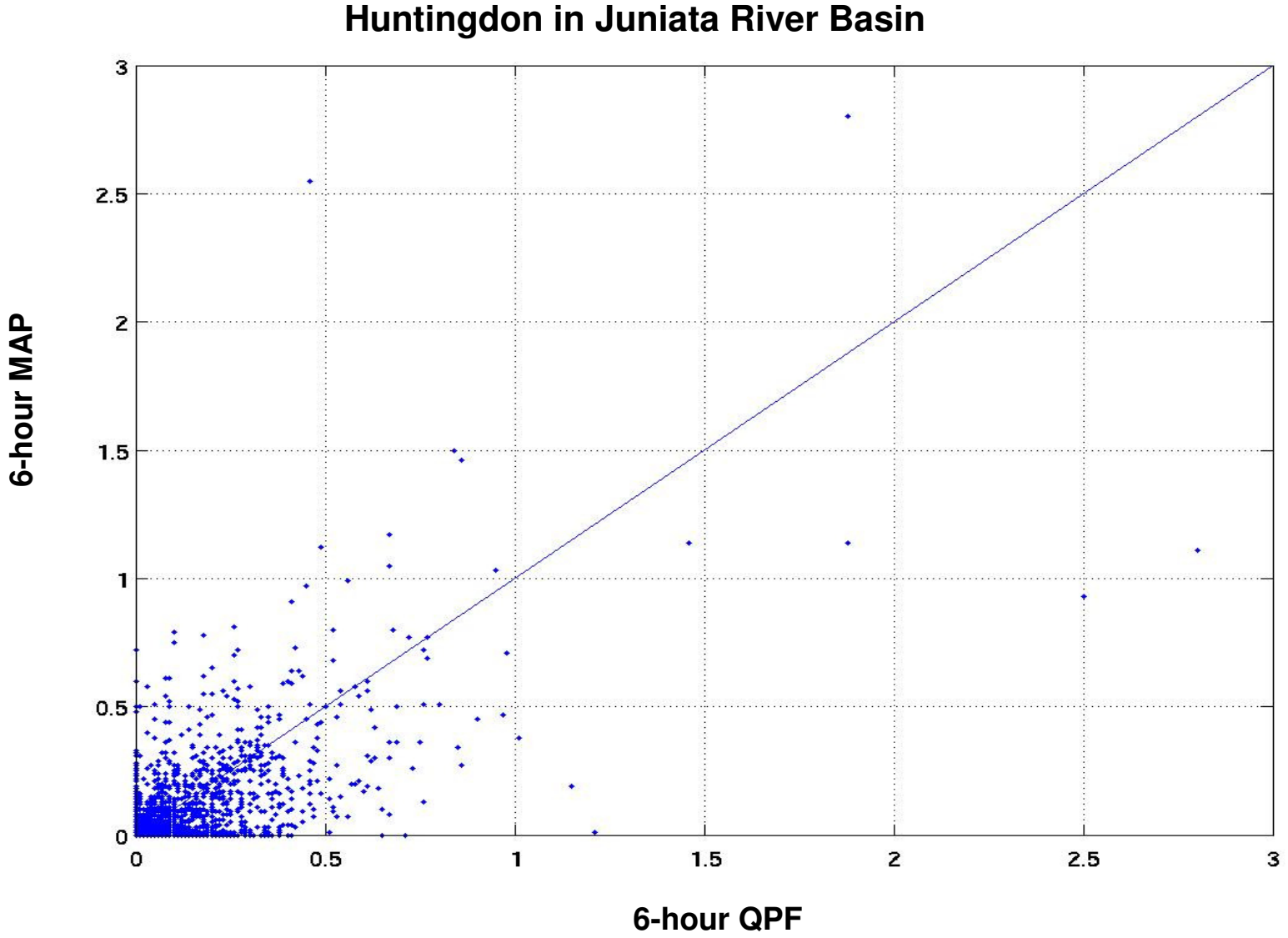
Methodology

Joint relationship between fcst and obs (T_{max})

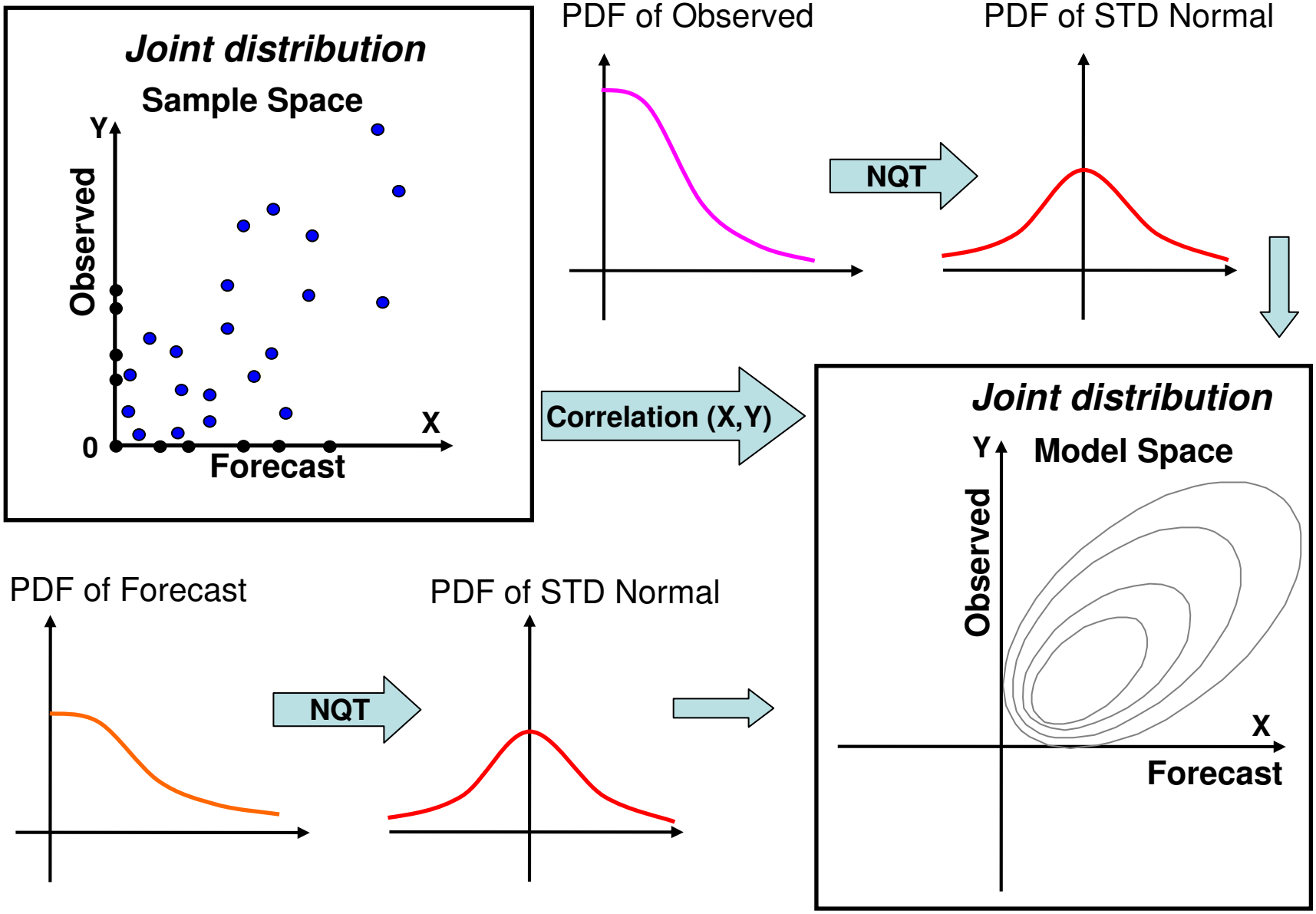


Methodology

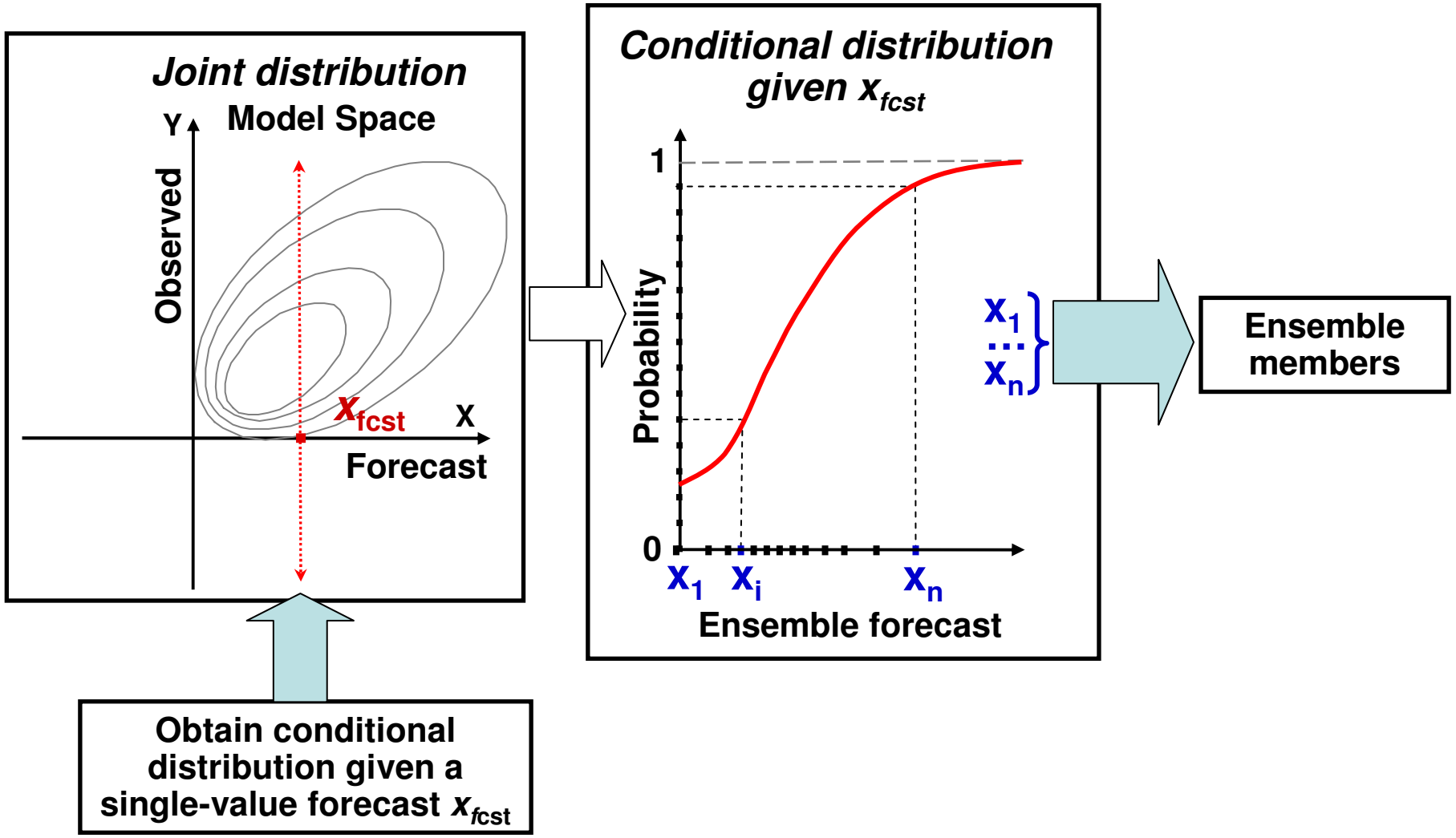
Joint relationship between fcst and obs (precipitation)



Methodology – Calibration

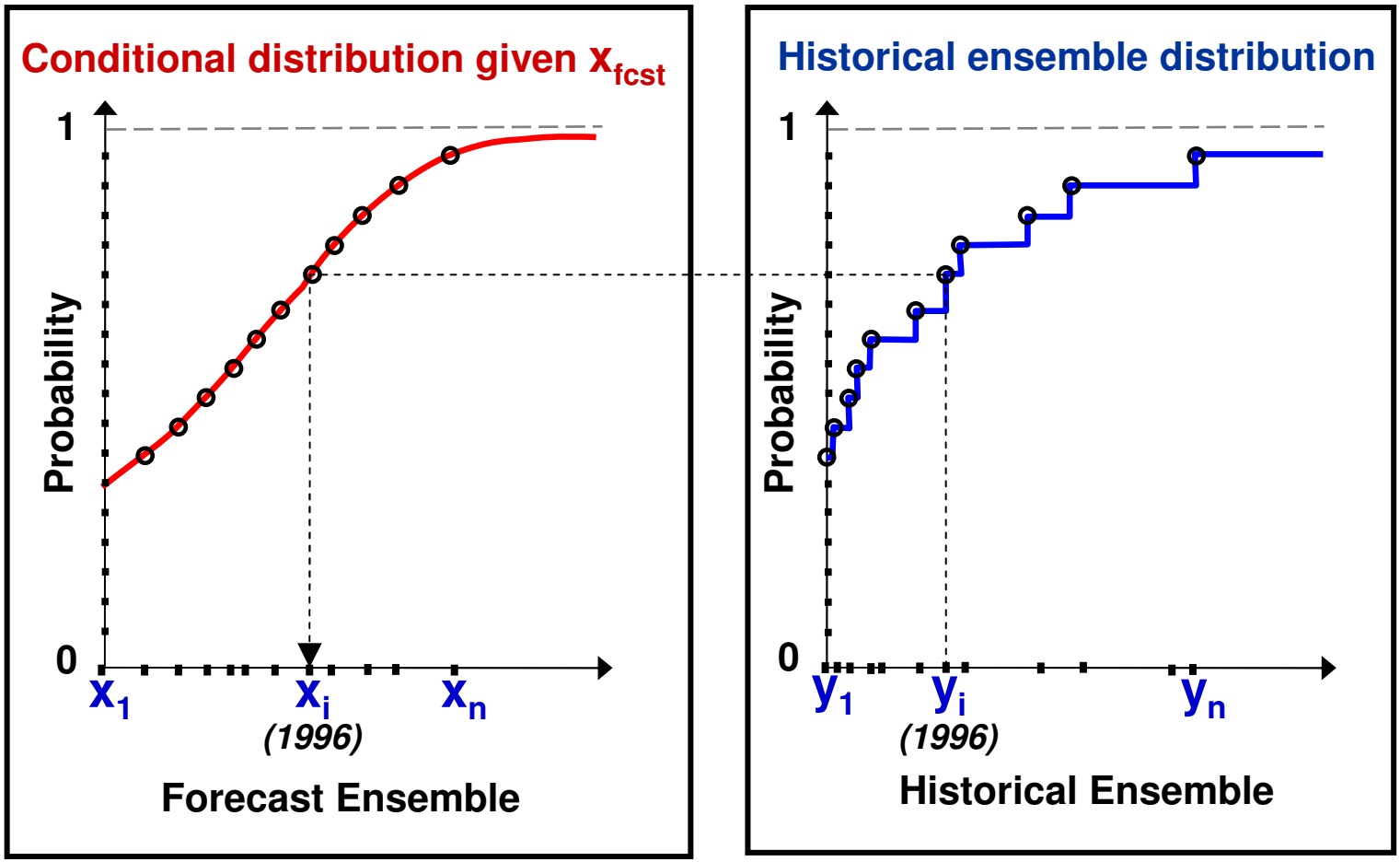


Methodology – Ensemble Generation



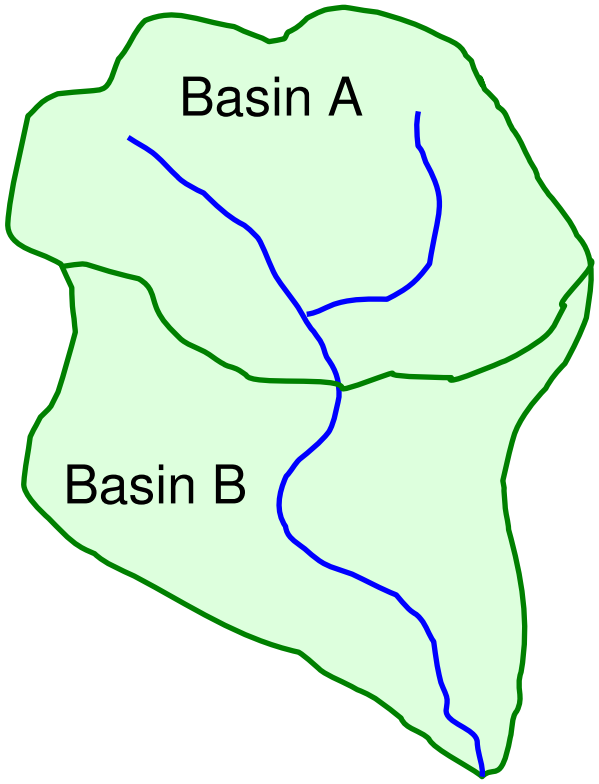
Methodology – Schaake Shuffle

For each segment, at each time step, associate forecast ensemble members (left panel) with historical ensemble members (right panel) by year



Methodology – Schaake Shuffle

The spatial variability between two neighboring MAP basins is preserved (in terms of rank correlation) in the forecast ensembles



For a given time step

Basin A

Year	Historical ensemble	Forecast ensemble
1958	*	*
…		
1996	y_{Ai}	x_{Ai}
…		
2007	*	*

Basin B

Year	Historical ensemble	Forecast ensemble
1958	*	*
…		
1996	y_{Bi}	x_{Bi}
…		
2007	*	*

Correlated

Similarly Correlated

Similarly, temporal variability, as well as co-variability with temperature, is preserved (in terms of rank correlation) in the forecast ensembles

Evaluation and Implementation

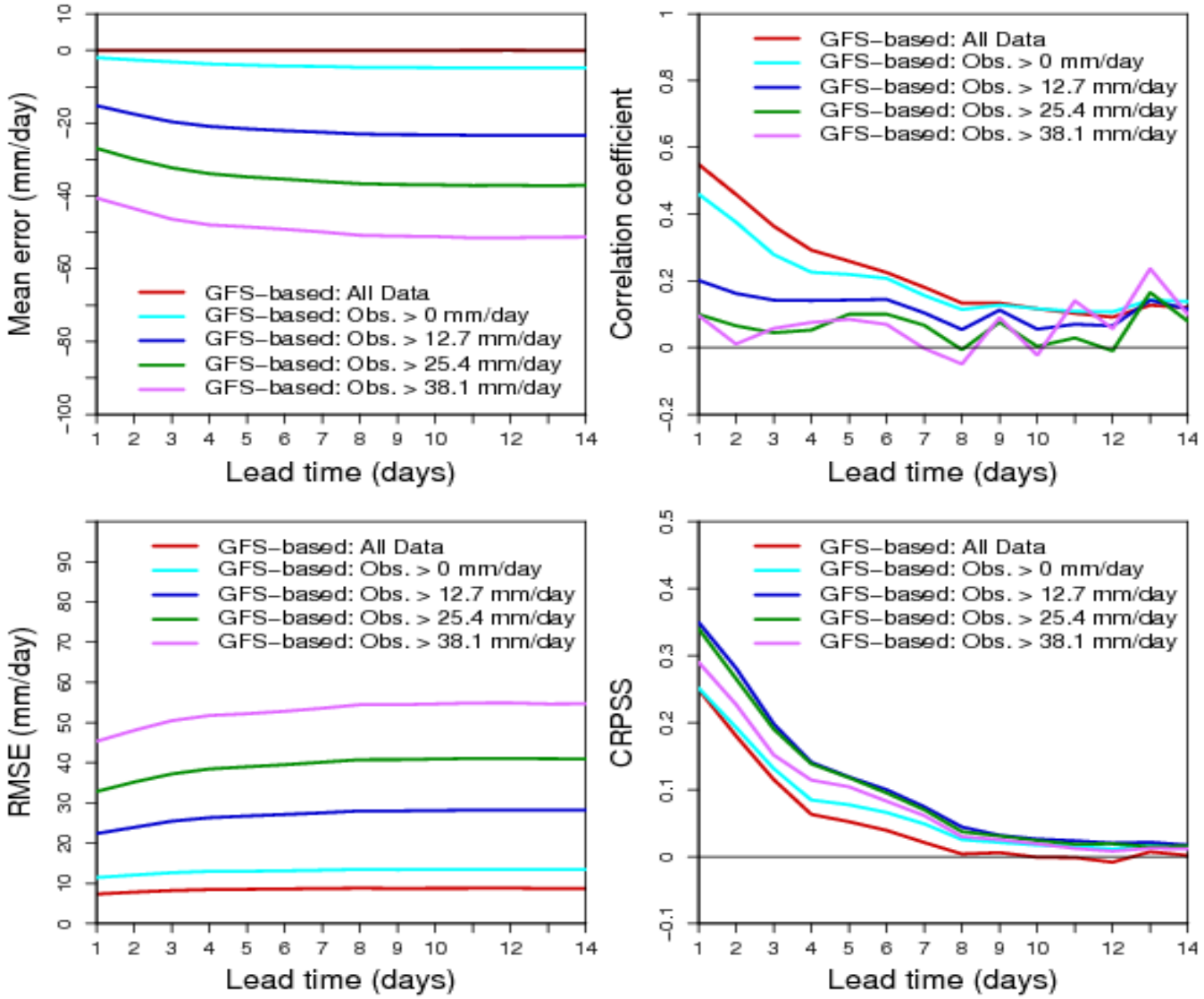
- Evaluation via multi-year hindcasting on several test basins at MA-, AB-, CN-, and NW-RFCs
- EPP interfaced with the Community Hydrologic Prediction System (CHPS) and released to selected RFCs for testing in mid-Nov. 2010
- Planned
  - Extensive evaluation on multiple temporal scales and for various basins to guide future enhancements
  - Use new forecast datasets (CFSv2, GEFS)
  - Compare with other techniques to guide improvements
  - EPP to become operational by Jan. 2014

Results: ABRFC Case Study

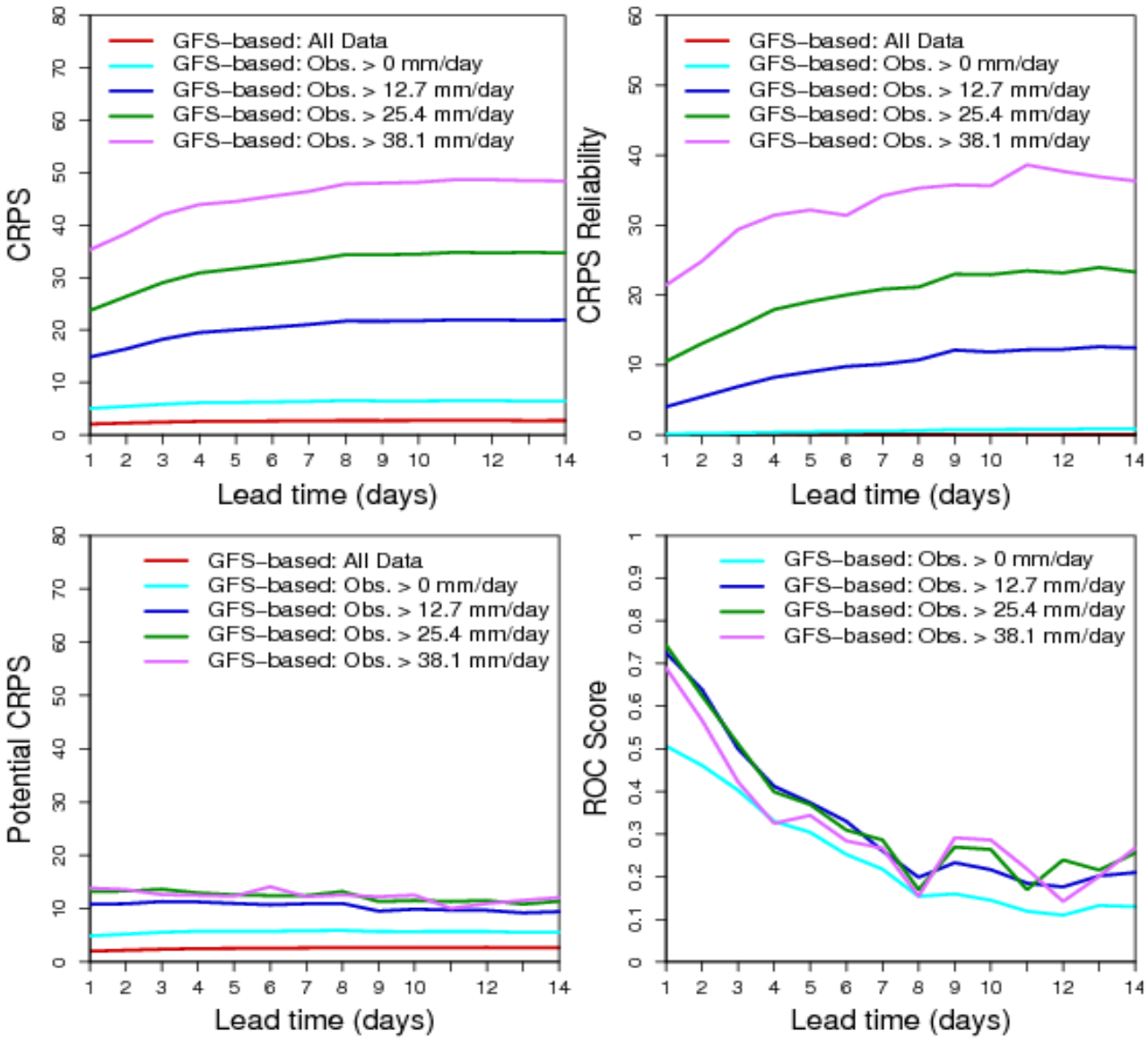
Dependent Verification Results
ABRFC basin BLUO2

- Single-valued forecasts: GFS ensemble mean values
- Verification window: Jan – Dec
- Temporal aggregation period in verification: 24 hours
- Verification period: 01/01/1979-12/31/1999

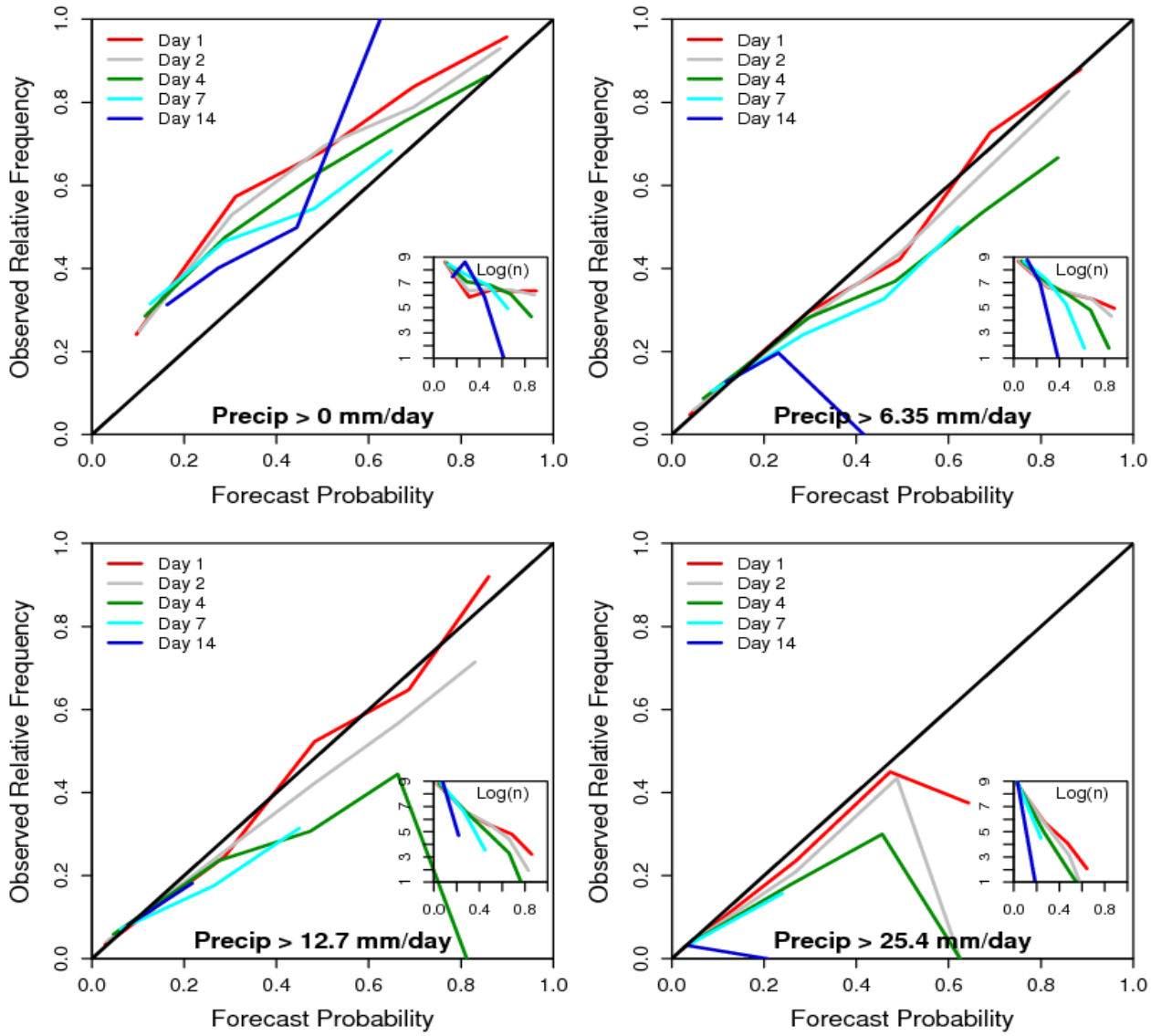
Results: ABRFC Case Study



Results: ABRFC Case Study



Results: ABRFC Case Study



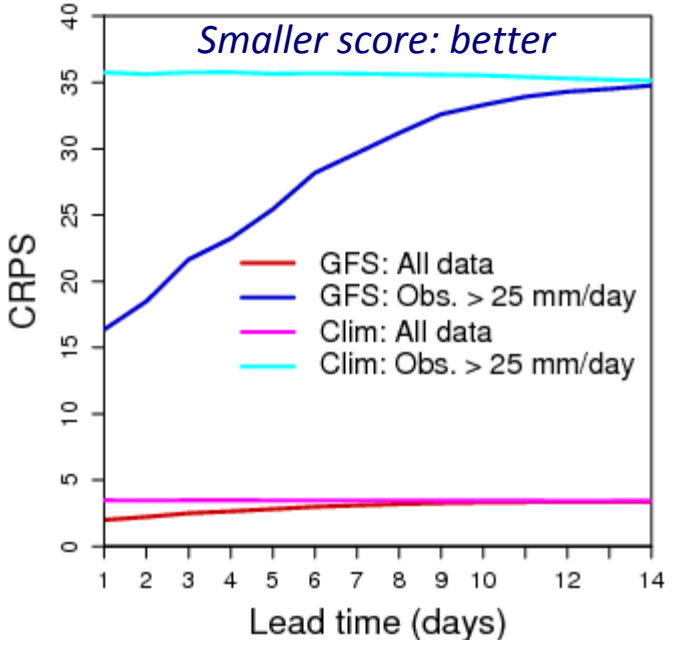
Results: CNRFC Case Study

EPP: Retain skill in single-valued input forecasts & generate unbiased ensembles

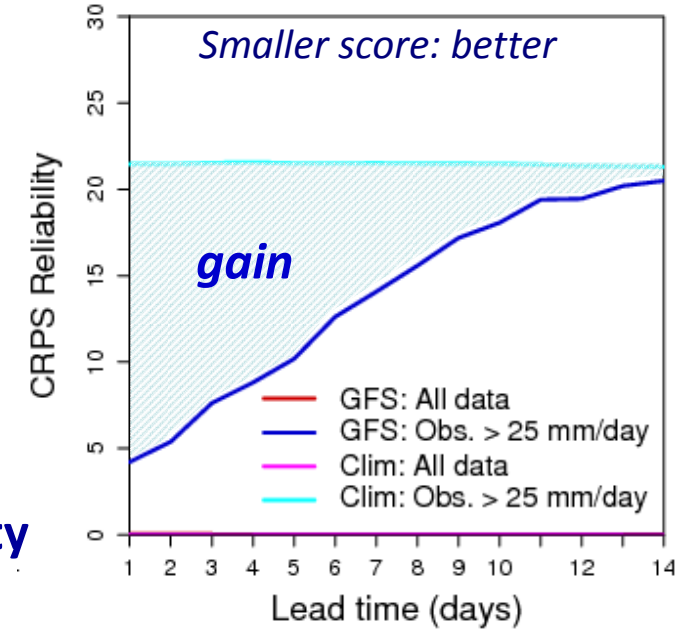
- North Fork of the American River (875 km^2) near Sacramento, California
- Daily products, 14 lead days, 45 members, 1979-2005
- GFS-based ensemble forecasts against climatology, evaluated via Ensemble Verification System (EVS)

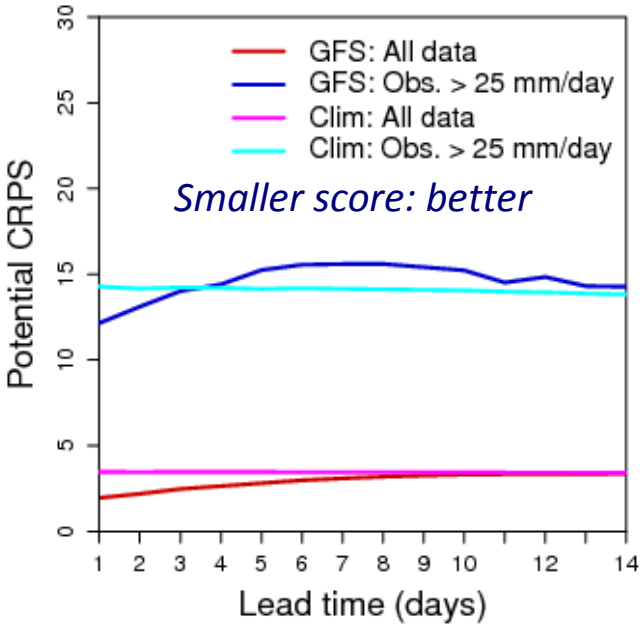
Results: CNRFC Case Study

- GFS-based 24-hr **precipitation ENS** from EPP vs. Climatology:
  - Mean Continuous Ranked Probability Score (CRPS)
  - Mean CRPS decomposition



Mean CRPS=
Reliability
+ Potential CRPS





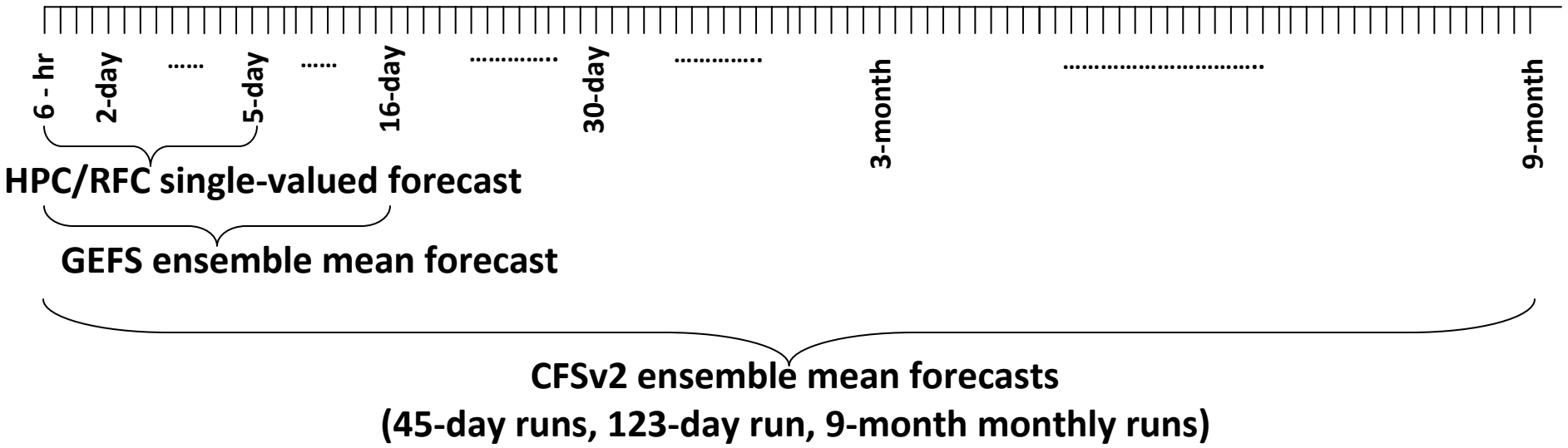
Gain is mostly in reliability

Planned Work

- Use new forecast datasets for which a multi-year archive of hindcasts is available:
  - CFSv2: processing of hindcasts has just started; should significantly improve forecast for the first 30 days
  - GEFS: hindcasts to be made available by June 2011
- Evaluate EPP extensively and compare with other techniques
  - Evaluate EPP performance on multiple time scales (from 6-hr to seasonal)
  - Evaluate/improve EPP grid processor
  - Establish EPP data requirements by analyzing sampling uncertainty
  - Develop guidance on how to best use current EPP (RFC, GFS and CFS components) and other forcing input ensembles (e.g., SREF ensembles)
  - Evaluate/improve Schaake Shuffle technique to generate more ensemble members w/ consistent space-time properties across multiple variables
- Coordinate w/ NCEP and ESRL on forcing input ensembles (e.g., techniques, reforecasts/forecast datasets)

Planned Work – Upgrading current EPP

- Combine new NWP forecast datasets to generate improved 6-hr to 1-yr forcing input ensembles



Example of aggregate events:

- 6-hr events for lead days 1-5
- 24-hr (1-day) events for lead days 6-8
- 2-day events for lead days 9-16
- 4-day events for lead days 17-24
- 6-day event for lead days 25-30
- 1-month events for lead months 2-9

Thank You !

limin.wu @noaa.gov
julie.demargne@noaa.gov