HPC QPF
Confidence Interval Forecasts

- Background
- Introduction
- Methodology
- Results
- Verification
- Future Plan

NWS/NCEP/Hydrometeorological Prediction Center
I.M. System Group
Jung-Sun Im
November 29, 2006

Background

- Funding Provided by the AHPS (Advanced Hydrologic Prediction System) program of NOAA NWS.

- Objectives: The HPC at the NCEP has produced a suite of deterministic QPFs for over 40 years. While the operational forecasts have proven to be useful in their present form, they offer no information concerning the uncertainties of individual forecasts.

  The purpose of this study is

  to develop a methodology to quantify
  the uncertainty in manually produced 6-h HPC QPF
  using NCEP short-range ensemble forecasts (SREFs).

- Benefits: 1) This research produces probabilistic error forecasts as well as deterministic error forecasts. 2) This study is also the first attempt to relate model produced ensemble forecasts with manually derived QPFs, and its operational application eventually could aid in increasing the forecast lead-time and accuracy of RFC streamflow model forecasts. 3) The methodology developed for QPF could have broader applicability and perhaps could be applied to other hydrometeorological parameters.

Timeline

- Work begun August 2003
- Prototype CI Forecasts Available – December 2003
- Real-time CI Forecasts using 15 QPF Cat. Reg. Available – March 2004
- AGU Presentation – May 2004
- Real-time CI Forecasts for 12h QPF Available – July 2004
- Upgrade (Two method combination) – September 2004
- Start providing point or grid data products to several RFCs from diverse geographic and hydrologic areas of the country

  1) ABRFC – starting July 2004

  2) NCRFC – starting December 2004

  ⇒ 2006 AMS Presentation by John Halquist:

  "Use of HPC QPF CI forecasts to produce a hydrologic ensemble of river forecasts"

  3) LMRFC – starting May 2005

  4) MBRFC – starting October 2006
- Publication in *Wea. Forecasting* – February 2006
- Future Plan: New scheme implementation - December 2006

Introduction

The quality of manual QPF forecasts, especially beyond 12 hrs, vary considerably from forecast to forecast.

: This can be related directly to the inherent uncertainty in model predictions.

⇑

Basis for all HPC Forecasts

*** One of the possible ways to quantify the uncertainty is the use of ensemble forecasts (SREF).**

Hypothesis
used in this approach

The larger spread in SREF, the greater uncertainty in HPC QPF

: To test this hypothesis,
Investigate the relationship
between
HPC QPF Absolute Error (AE)
& Several parameters available from the SREF

Spread / forecast error relationships

- **Investigate relationship between HPC QPF AE and SREF spread for 500mb heights: Result: Low correlation (-0.2 < cc < 0.2)**
- **Investigate relationship between HPC QPF AE and SREF spread for 850mb RH: Result: Low correlation (-0.2 < cc < 0.2)**
- **Investigate relationship between HPC QPF AE and SREF QPF spread: Result: Higher correlation**

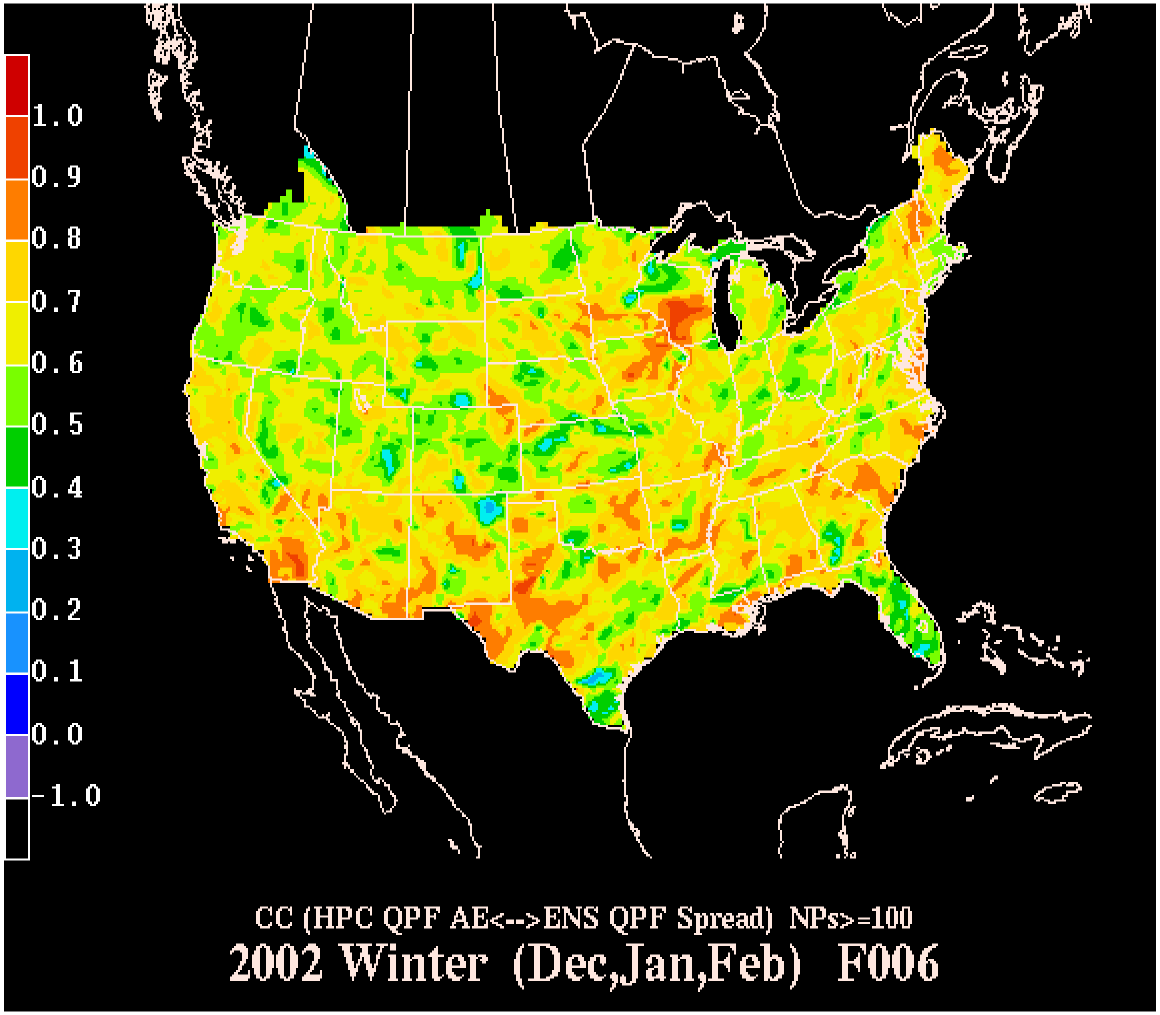
-compute cc for various regression types (e.g., simple linear, logarithmic, power, exponential, polynomial, and multiple regressions…)

-best fit was found in the simple linear regression

- test the null hypothesis (Ho: slope=0)

: rejected at the 0.001 error rate (99.9% confidence level)

Correlation Coefficient between HPC QPF AE and ENS QPF Spread



cc ≥ 0.5

at most US grid points (90.5%)

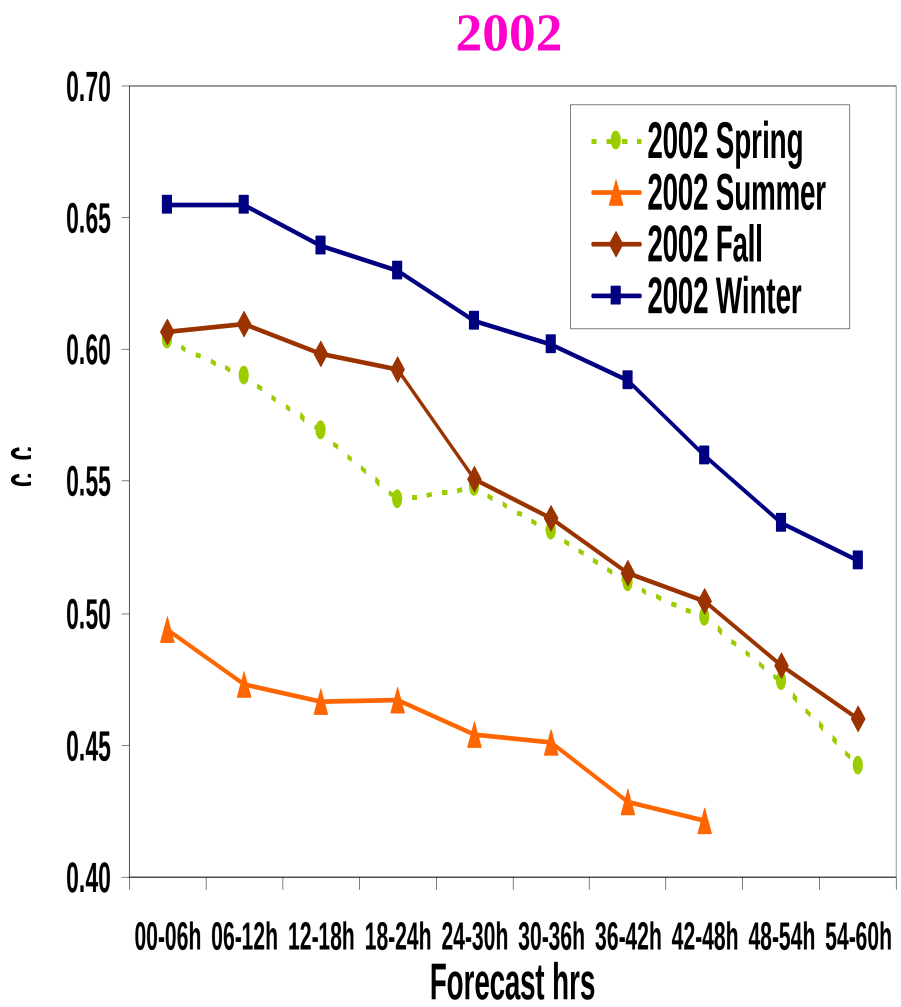
cc ≥ 0.8

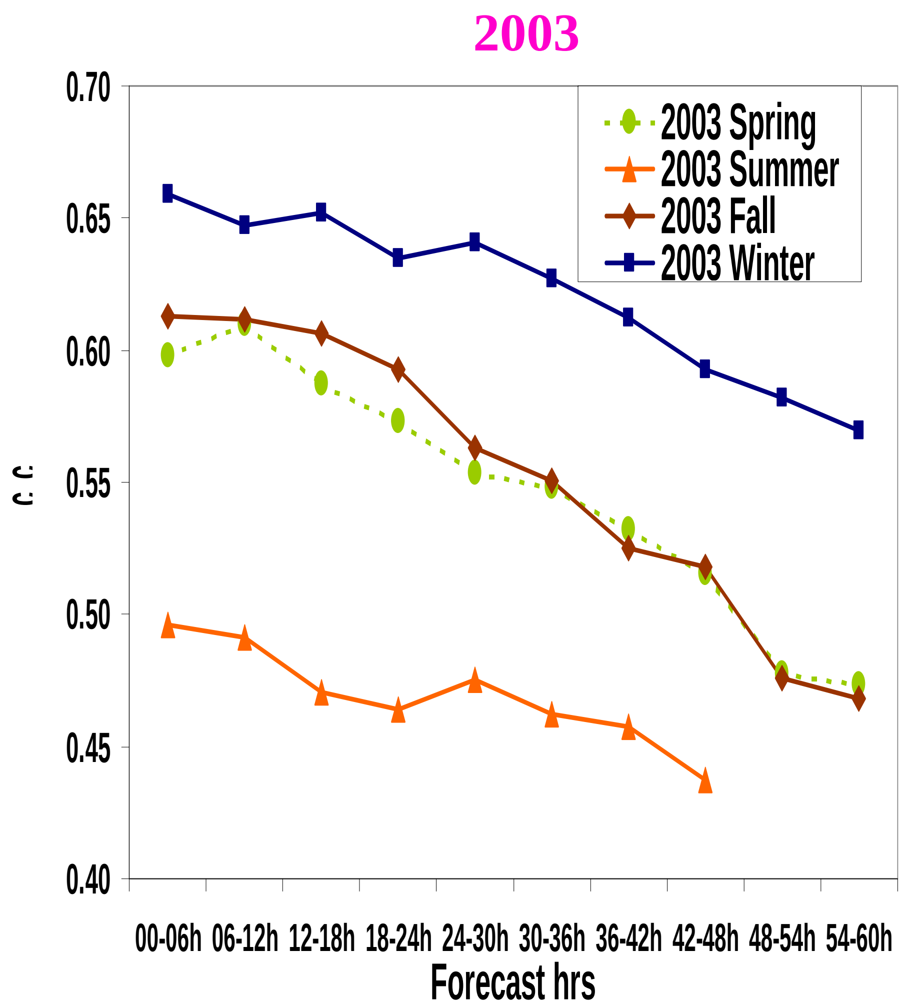
at many grid points (10.5%)

⇓

This indicates the SREF can be used to predict the uncertainties of the HPC QPFs. The cc is greater than those previously reported in works (e.g., Grimit and Mass 2002; Stensrud and Yussouf 2003).

CC Variations with Forecast Hours (CC averaged for US)





Linear Regression of HPC QPF AE on ENS QPF Spread

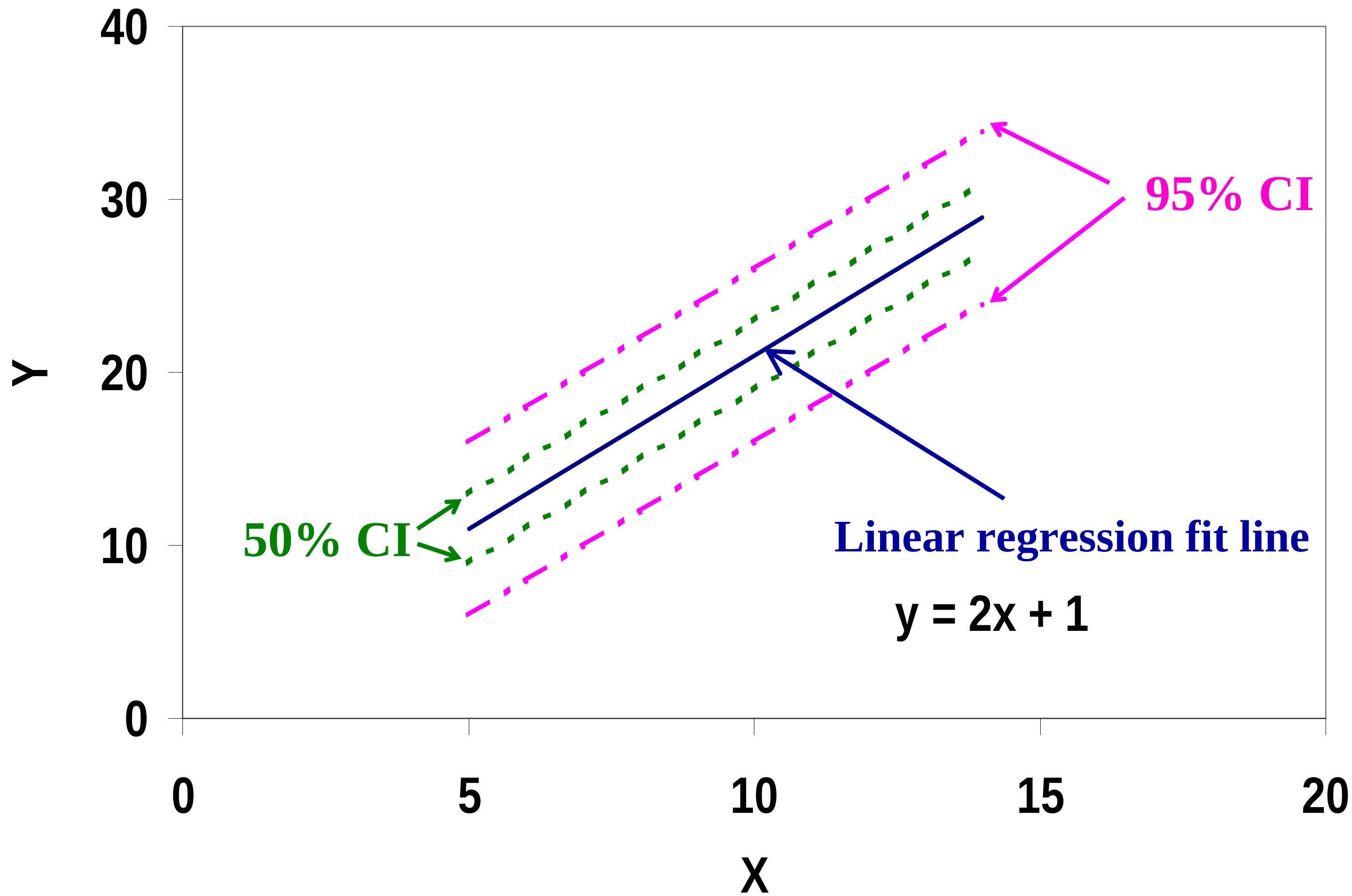
$$|RFC\_QPE_{ij} - HPC\_QPF_{ij}| = b0_{ij} + b1_{ij}\, SP_{ij}$$

$$\uparrow$$

"AE_{ij}"

- **RFC_QPE: observed precipitation (i.e., ground truth)**
- **HPC_QPF:HPC precipitation forecast**
- **b0:intercept**
- **b1:slope**
- **SP:ensemble QPF spread**
- **i, j : horizontal position of each grid point**

Schematic illustration of Confidence Interval (CI) for df=∞



Confidence Interval (CI) for the Predicted AE at SP

$$CI(pred.AE_{ij} at SPo_{ij}) = b0_{ij} + b1_{ij} SPo_{ij} \pm t_{ij} \sqrt{MSE_{ij}\left[1 + \frac{1}{n_{ij}} + \frac{(SPo_{ij} - \overline{SP}_{ij})^2}{SS(SP)_{ij}}\right]}$$

(1) (2) (3) (4) (5) (6) (7)

(1) b0: Intercept
(2) b1: Slope
(3) SPo: SP value for the individual point we are trying to predict
(4) t: appropriate percentile of the t distribution
(5) MSE: Mean Squared Error (estimate of the true variance of residuals)
(6) n: Number of data points
(7) SS(SP): Sum of Squares for SP

⇒ When we compute the minimum (i.e., b0 + b1*SPo – t $\sqrt{\ }$) and maximum (i.e., b0 + b1*SPo + t $\sqrt{\ }$), we can say we are 95% confident that HPC QPF with SP=SPo will have AE between the minimum value and the maximum value.

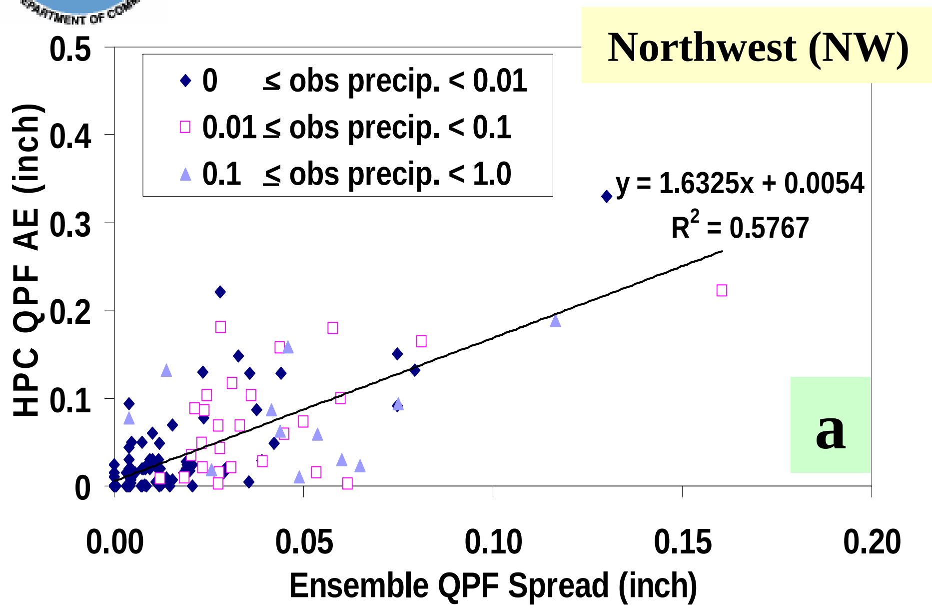
MSE
(Estimate of the true variance of residuals)

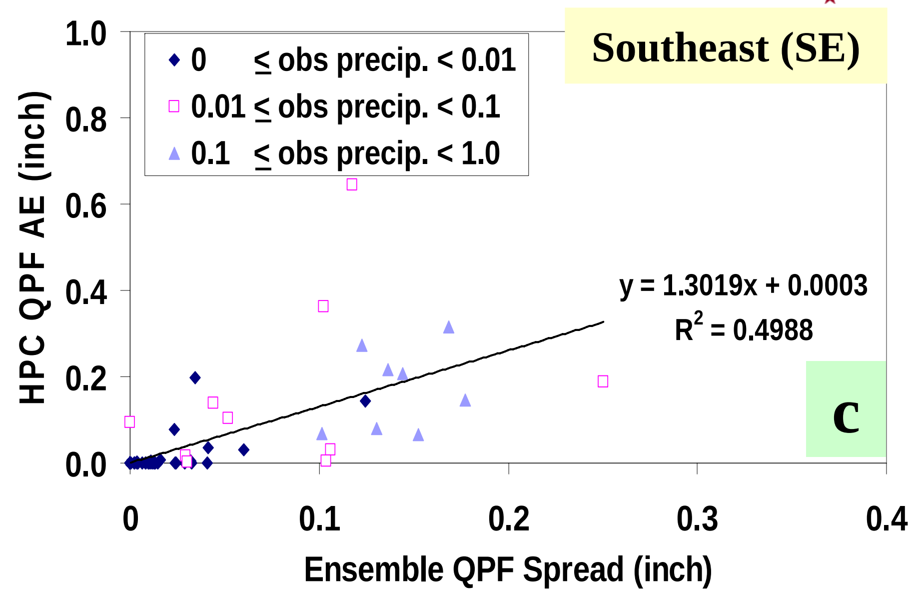
In order to construct the CI, we usually assume the deviations of the observed AE's from the true fit line (i.e., true errors, ε) satisfy the following assumptions

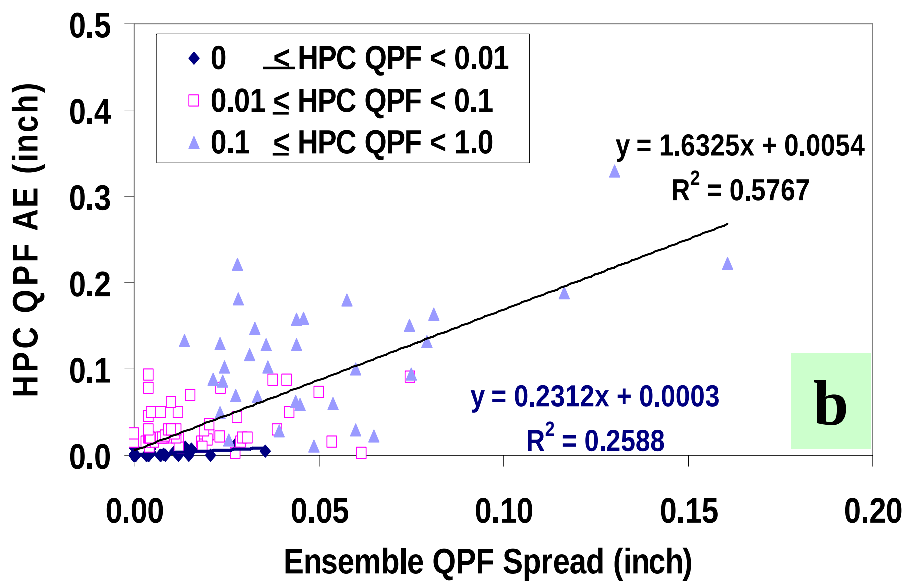
1. They all have mean 0.
2. They all have the same variance σ^2.
3. They are uncorrelated.
4. They are normally distributed.

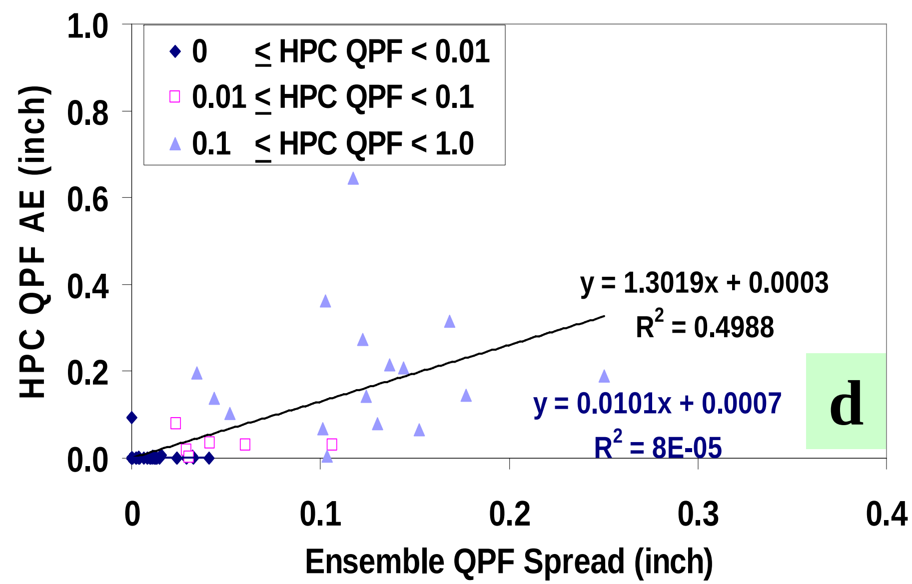
To test our data sets satisfy the above assumptions, the estimates of ε were computed and examined. First, it was demonstrated that the methodology used for the first step approach satisfy the assumptions 1 and 4. Second, all the MSEs (estimates of the ε's variance) computed for HPC_QPF categorized subsets were not the same. Especially, the MSEs computed for the wet QPF categorized subsets were much greater than the MSE for dry QPF subset.

Scatterplots of HPC QPF AE vs ENS QPF SP for observed precip. subsets & for HPC QPF subsets









How to Handle Different MSEs for Dry and Wet Precipitation…

Stratification Methodology

(1) Method 0 (Dry/All Regressions)

- For Dry HPC QPF (QPF=0) CI forecast: Apply regression model equation parameters derived using previous year data of $0 \le QPF < 0.01$ inch
- For Wet HPC QPF (QPF>0) CI forecast: Apply regression model equation parameters derived using previous year data of $QPF \ge 0$

(2) Method 1 (Dry/Wet Regressions)

- Dry QPF (QPF=0): Apply Reg. Parm. derived from $0 \le QPF < 0.01$
- Wet QPF (QPF>0): Apply Reg. Parm. derived from QPF>0

(3) Method 2 (Dry/Light/ModerateHeavy Regressions)

- Dry QPF (QPF=0): Apply Reg. Parm. derived from $0 \le QPF < 0.01$
- Light QPF (0<QPF<0.1): Apply Reg. Parm. derived from 0<QPF<0.1
- Moderate/Heavy QPF ($QPF \ge 0.1$): Apply Reg. Parm. derived from $QPF \ge 0.1$

(4) Method 3 (Dry/Log-Log Regressions)

- Dry QPF (QPF=0): Apply Reg. Parm. derived from $0 \le QPF < 0.01$
- Wet QPF (QPF>0): Apply Reg. Parm. obtained from the logarithmically transformed data (both AE and SP) of QPF > 0

Stratification Methodology (Cont.)

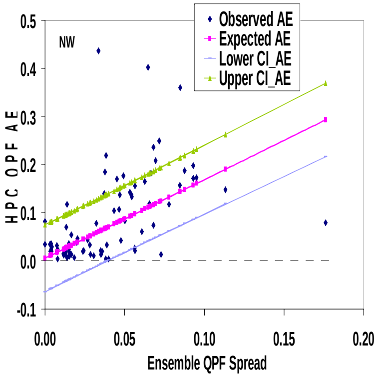
(5) Method 4 (15 QPF Categorized Regressions)

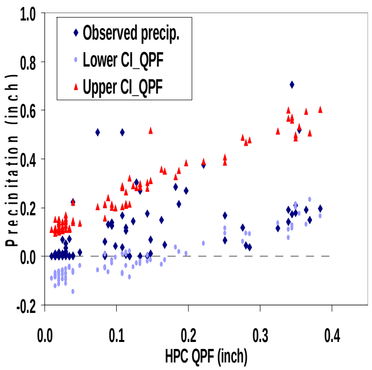
- **Dry QPF (QPF=0): Apply regression model equation parameters derived using previous year data of 0≤QPF<0.01 inch**
- **Wet QPF 1 (0<QPF<0.10): Apply regression model equation parameters derived using previous year data of 0<QPF<0.10 inch**
- **Wet QPF 2 (0.10≤QPF<0.25): Apply Reg. Parm. derived from 0<QPF< 0.25**
- **Wet QPF 3 (0.25≤QPF<0.50): Apply Reg. Parm. derived from 0<QPF<0.50**
- **Wet QPF 4 (0.50≤QPF<0.75): Apply Reg. Parm. derived from 0<QPF<0.75**
- **Wet QPF 5 (0.75≤QPF<1.00): Apply Reg. Parm. derived from 0<QPF<1.00**
- **Wet QPF 6 (1.00≤QPF<1.25): Apply Reg. Parm. derived from 0<QPF<1.25**
- **Wet QPF 7 (1.25≤QPF<1.50): Apply Reg. Parm. derived from 0<QPF<1.50**
- **Wet QPF 8 (1.50≤QPF<1.75): Apply Reg. Parm. derived from 0<QPF<1.75**
- **Wet QPF 9 (1.75≤QPF<2.00): Apply Reg. Parm. derived from 0<QPF<2.00**
- **Wet QPF10 (2.00≤QPF<2.50): Apply Reg. Parm. derived from 0<QPF<2.50**
- **Wet QPF11 (2.50≤QPF<3.00): Apply Reg. Parm. derived from 0<QPF<3.00**
- **Wet QPF12 (3.00≤QPF<4.00): Apply Reg. Parm. derived from 0<QPF<4.00**
- **Wet QPF13 (4.00≤QPF<5.00): Apply Reg. Parm. derived from 0<QPF<5.00**
- **Wet QPF14 (5.00≤QPF< infinite): Apply Reg. Parm. derived from 0<QPF<infinite**

Method Comparisons:

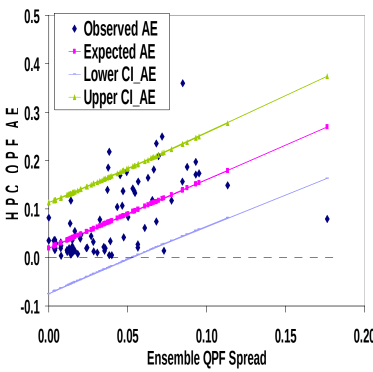
Evaluation of CI forecasts for AE and HPC QPF for 5 methods

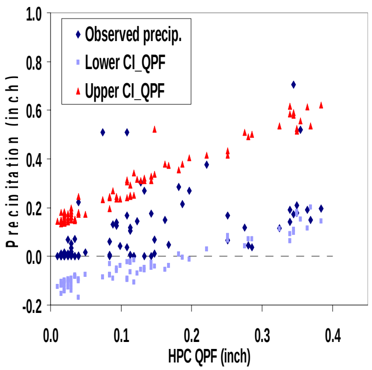
Dry/All (M0)



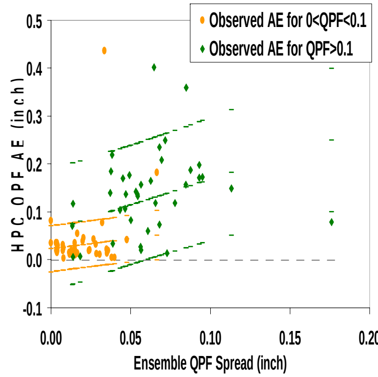


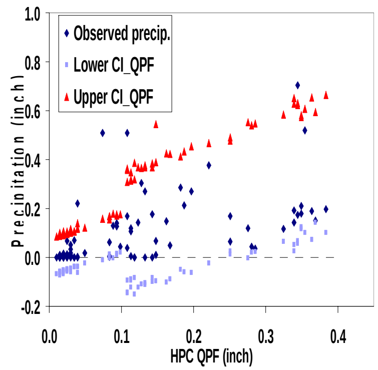
Dry/Wet (M1)



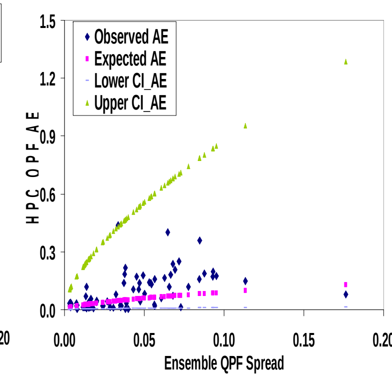


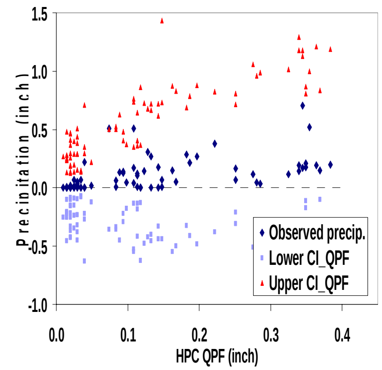
3 QPF (M2)



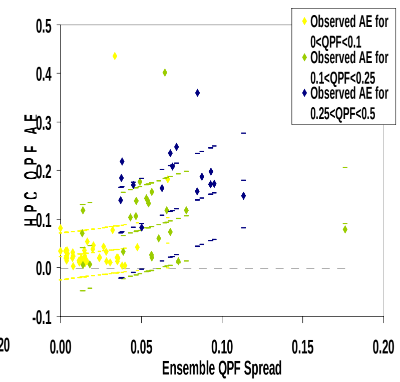


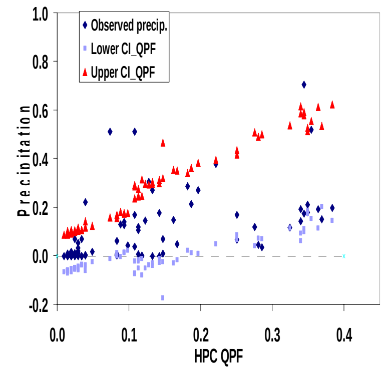
LN trans. (M3)





15 QPF (M4)





Note: Reg. parms were obtained using previous year data (i.e., 2001 winter data) and then CIs were computed (predicted) for 2002 winter

Method Comparison Results

The best CI forecasts
satisfying
Narrow CI and High Hit Rate **(Lower_CI≤OBS≤Upper_CI),**
are found in
Method 4 (15 QPF categorized regression approach).

Final competitors are

Method 0 and Method 4
in Probabilistic CI forecasts for HPC QPF and HPC QPF AE

Method 2 and Method 4
in Deterministic HPC QPF AE forecasts.

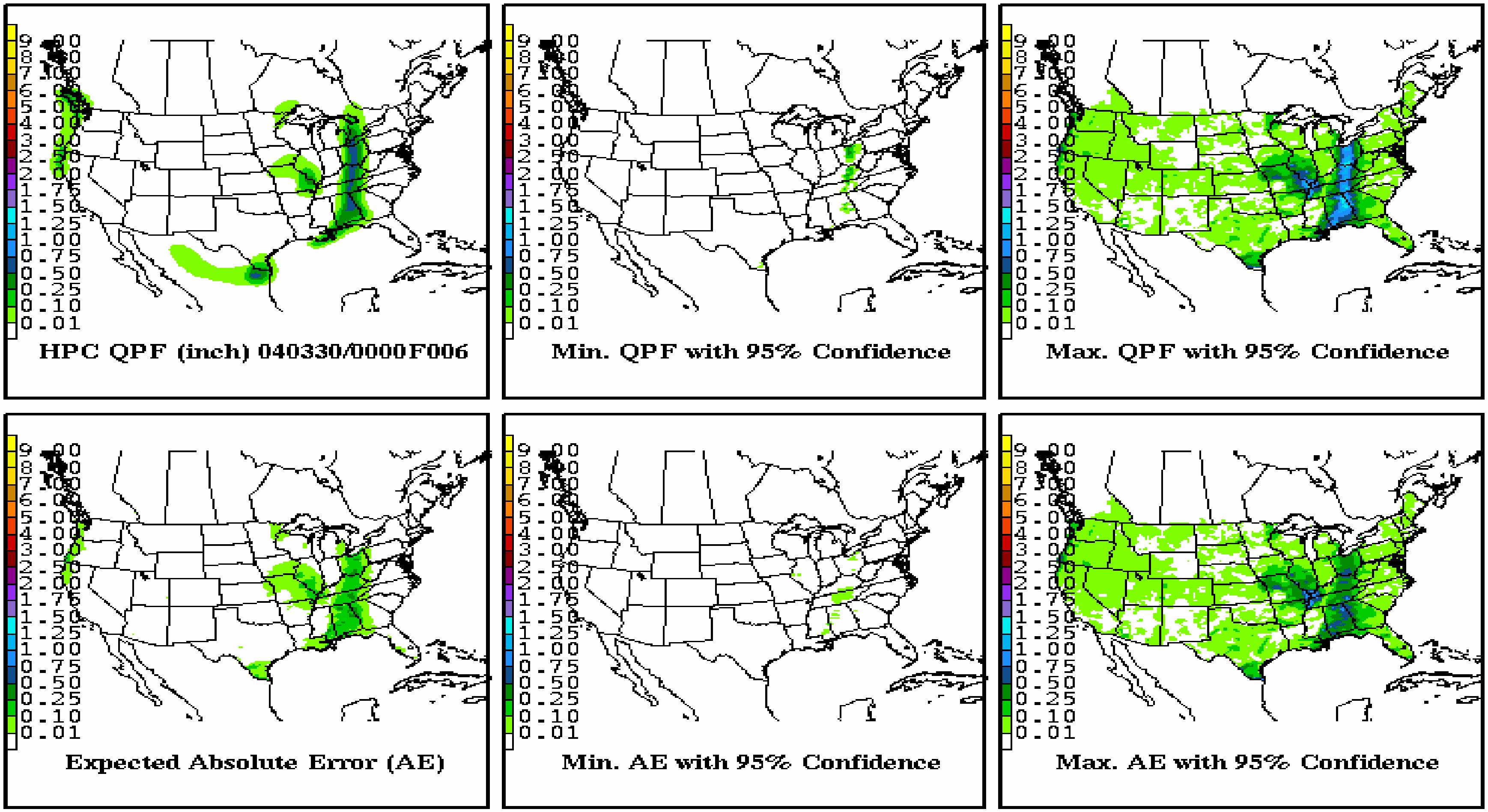
(Details are described in Im et al. (*Wea. Forecasting*, 2006)

Confidence Interval Forecasts

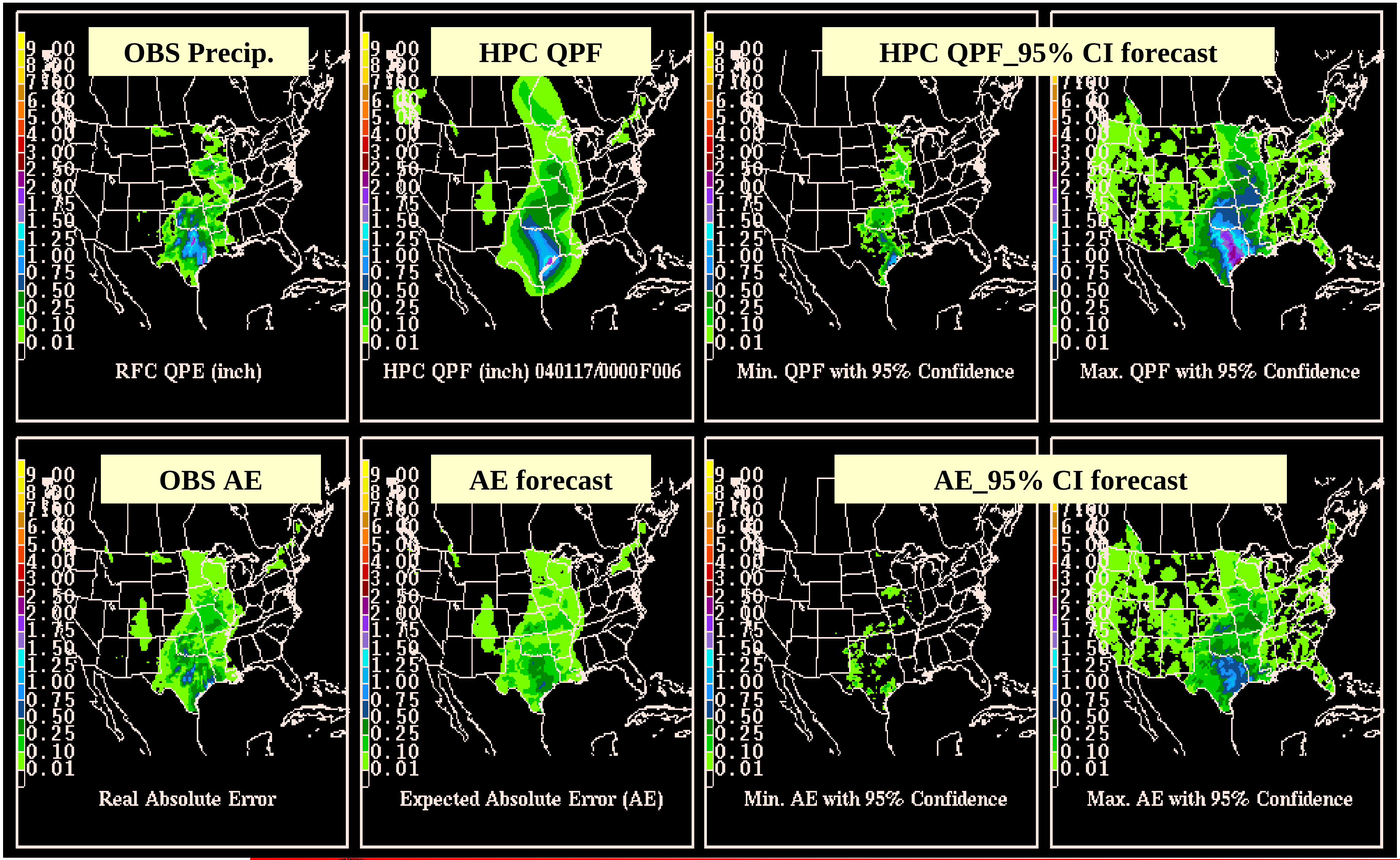
Using 15 QPF categorized regressions

(Real Time Data Processing)

http://www.hpc.ncep.noaa.gov/qpfci/qpfci.shtml



2004 Jan 17 00z Case
(Predicted Precip. & Observed Precip.)



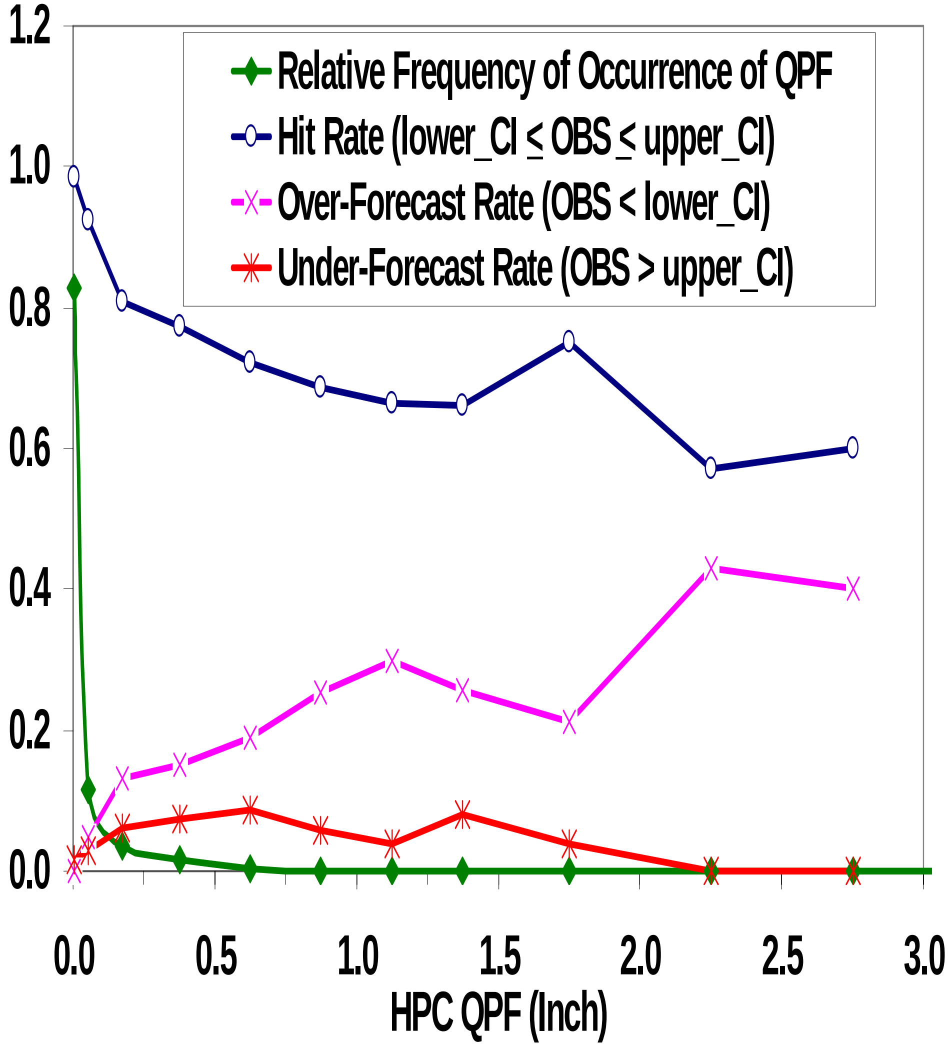
VERIFICATION

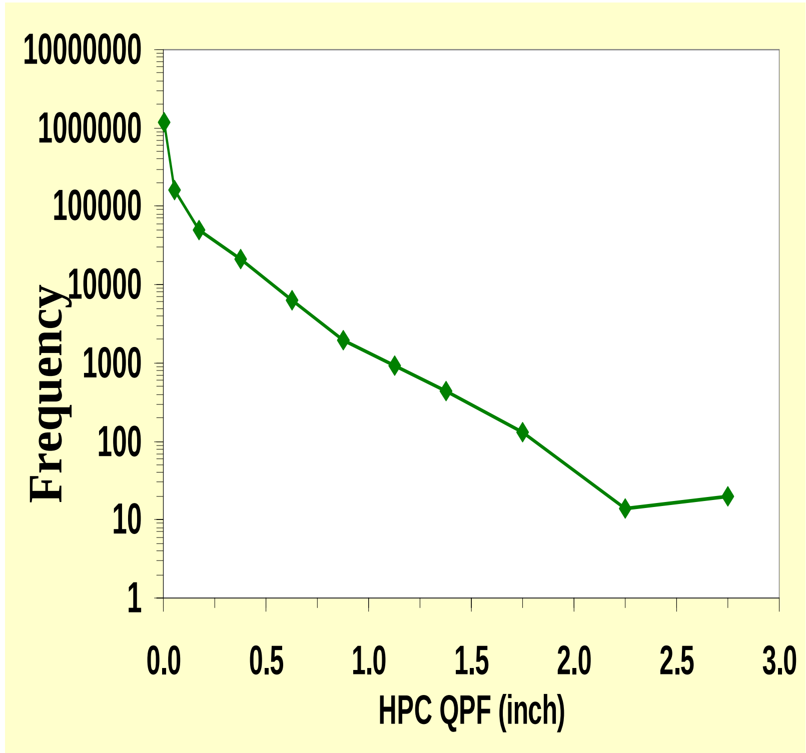
Detection % (Hit Rate x 100)

for "HPC QPF 95% CI" and "HPC QPF AE 95% CI"

	2002 Winter (2002 Dec, 2003 Jan Feb)		2003 Spring (2003 Mar Apr May)		2003 Summer (2003 Jun Jul Aug)		2003 Fall (2003 Sep Oct Nov)		2003 Winter (2003 Dec, 2004 Jan Feb)	
fhrs	QPF Dec. %	AE Dec. %	QPF Dec. %	AE Dec. %	QPF Dec. %	AE Dec. %	QPF Dec. %	AE Dec. %	QPF Dec. %	AE Dec. %
00-06hrs	96.68 ± 2.09	95.78 ± 2.34	95.26 ±3.19	94.29 ± 3.55	95.57 ± 3.95	94.92 ± 4.35	97.08 ± 2.28	95.11± 2.69	96.41 ± 2.43	95.15 ± 2.82
06-12hrs	96.58 ± 2.19	95.75 ± 2.44	95.48 ±3.06	94.52 ± 3.41	95.81 ± 3.94	95.15 ± 4.61	97.07 ± 2.32	96.08 ± 2.80	96.27 ± 2.59	95.00 ± 2.96
12-18hrs	96.70 ± 1.97	95.77 ± 2.31	95.37 ± 3.18	94.38 ± 3.53	95.86 ± 3.48	95.28 ± 3.98	97.15 ± 2.15	96.19 ± 2.56	96.53 ± 2.39	95.48 ± 2.70
18-24hrs	96.49 ± 2.15	95.67 ± 2.40	95.41 ± 3.16	94.66 ± 3.43	95.62 ± 4.65	94.97 ± 5.37	96.94 ± 2.24	96.04 ± 2.61	96.32 ± 2.52	95.22 ± 2.73
24-30hrs	96.38 ± 2.03	95.46 ± 2.42	95.13 ± 3.19	94.20 ± 3.52	96.70 ± 3.95	94.97 ± 4.70	97.08 ± 2.12	96.31 ± 2.45	96.71 ± 2.39	95.74 ± 2.52
30-36hrs	96.29 ± 2.06	95.57 ± 2.33	95.16 ± 3.09	94.38 ± 3.38	95.66 ± 4.44	94.89 ± 5.36	96.82 ± 2.24	96.17 ± 2.54	96.42 ± 2.44	95.40 ± 2.65
36-42hrs	96.18 ± 2.14	95.47 ± 2.43	95.10 ± 3.18	94.31 ± 3.57	95.73 ± 3.56	95.05 ± 4.59	96.75 ± 2.38	96.12 ± 2.66	96.47 ± 2.55	95.57 ± 2.81
42-48hrs	95.92 ± 2.39	95.25 ± 2.61	95.19 ± 3.09	94.50 ± 3.48	95.61 ± 3.83	94.98 ± 4.99	96.69 ± 2.29	96.08 ± 2.51	96.22 ± 2.66	95.40 ± 2.86
48-54hrs	96.15 ± 2.19	95.46 ± 2.48	94.95 ± 3.31	94.19 ± 3.60	95.09 ± 4.34	94.33 ± 5.40	96.30 ± 2.55	95.65 ± 2.84	96.41 ± 2.58	95.63 ± 2.77
54-60hrs	95.90 ± 2.34	95.29 ± 2.52	95.16 ± 3.01	94.58 ± 3.29	94.69 ± 4.77	93.97 ± 6.04	96.22 ± 2.61	95.62 ± 2.87	96.16 ± 2.56	95.40 ± 2.77

Relative Frequency of Occurrence of QPF & CI forecast Hit / False rates versus QPF





Two Method Combination (M4+M0)

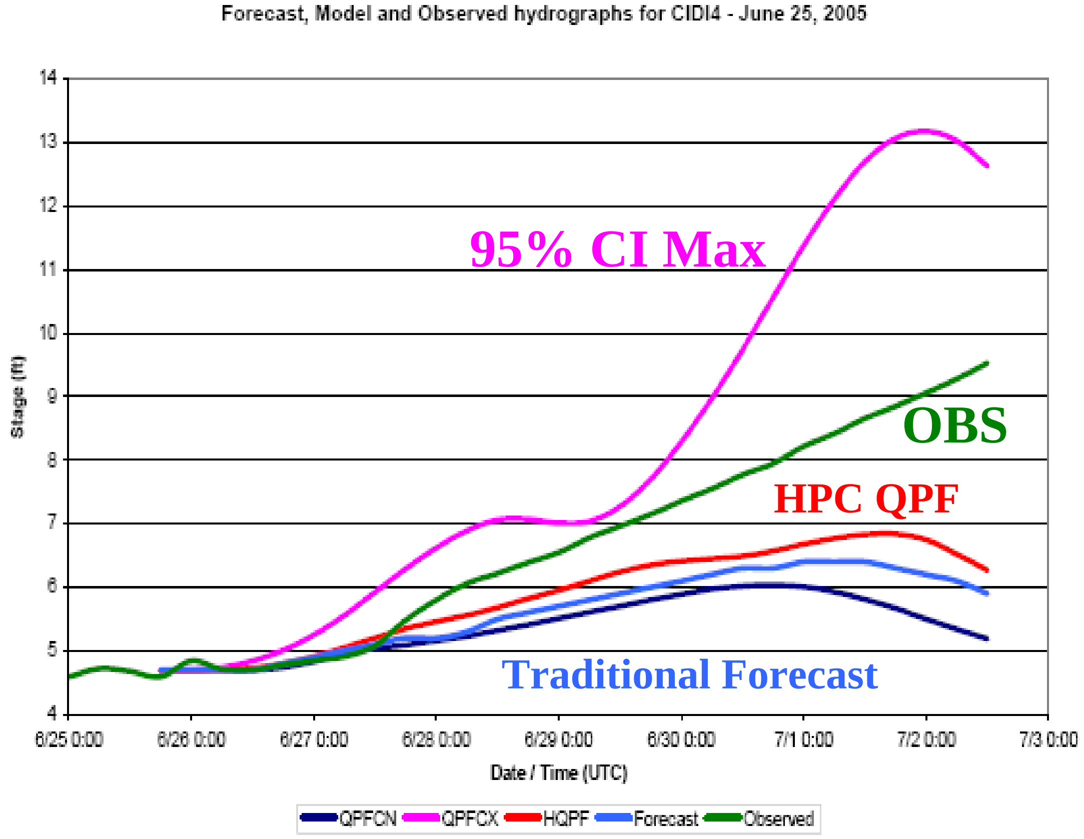
$$(MSE)_{M4} < \left[\frac{QPF_{\max}}{t\_value}\right]^2 \Rightarrow M4$$

$$(MSE)_{M4} \geq \left[\frac{QPF_{\max}}{t\_value}\right]^2 \Rightarrow M0$$

The verification statistics indicates improvements in the CI sizes (significantly reduced CI size in light rain ranges), while showing ignorable changes in hit rates.

Since September 2004, this method has been implemented to the real time confidence interval forecasts.

Use of HPC QPF CI forecasts to produce a hydrologic ensemble of river forecasts at NCRFC



While HPC QPF and NCRFC traditional forecast are under forecast, the forecast using upper bound value of HPC QPF 95% CI indicates a significant rise and provides a reasonable upper limit for potential river stages.

John Halquist, 2006: 20th Conf. on Hydrology

Coming soon in HPC….

Current Scheme

Compute Regression parameters (i.e., b1,b0, MSE, n, mean SP, SS for SP, etc) each season and apply to CI computations

Use 09z and 21z cycle SREF

New Scheme

Compute Regression parameters each day using most recent 3 month data and apply to CI computations

Use 03z and 15z cycle SREF

Benefits

- HPC 95% QPF CI Forecasts will be delivered approximately 6 hrs earlier than currently.
- New scheme will adapt more quickly to changes in the operational SREF.
- New scheme will use information more relevant to the current weather regime.

Acknowledgments

Edwin Danaher

Keith Brill

Chris Bailey

Tish Soulliard

Mark Klein

Mike Bodner

Peter Manousos

Joey Carr

Jun Du

Zoltan Toth

John Schaake