

USE OF HPC QPF CONFIDENCE INTERVAL FORECASTS TO PRODUCE A HYDROLOGIC ENSEMBLE OF RIVER FORECASTS

John Halquist *

National Weather Service, Chanhassen, Minnesota

ABSTRACT

The NCRFC has been using the HPC deterministic QPF in its hydrologic modeling since 1994 with reasonable success; however we have received a number of requests from our users over the years for alternative QPF inputs (i.e., less or more) or multiple QPF solutions. Although we were able to meet some of these requests by manually manipulating the data, it was cumbersome, introduced significant overhead to the forecast process, and resulted in questionable, subjective results that lacked consistency.

The NCEP HPC began producing 95% Confidence Interval (CI) Quantitative Precipitation Forecasts (QPF) in early 2004. These forecasts continue to be under review and refinement based on verification results and user feedback. In 2005, the NCRFC began a project to utilize the Maximum 95% CI QPF and Minimum 95% CI QPF as routine additional inputs to the National Weather Service River Forecast System (NWSRFS). With these additional inputs we now produce a three member ensemble of river forecasts for selected locations. The results of the analysis of these ensembles will be presented.

1. INTRODUCTION

In 1994 the North Central River Forecast Center (NCRFC) began using Quantitative Precipitation Forecasts (QPF) as input to the National Weather Service River Forecast System (NWSRFS) hydrologic model in the

production of river forecasts. Accurate QPFs enhance river forecasts by indicating a rise in the river which may otherwise not be seen using observed precipitation alone. Therefore, users of National Weather Service (NWS) river forecasts will have more time to take action if necessary. The QPF, currently in use at the NCRFC are created by the NWS National Centers for Environmental Prediction (NCEP) Hydrometeorological Prediction Center (HPC), and are for 6 hour intervals out through 72 hours in the future. The NCRFC routinely uses 24 hours of QPF, however on occasion, longer periods may be used; particularly with large scale organized weather systems when confidence in the QPF is high. Our primary objective in including QPF in the short term deterministic hydrologic forecasts is to provide users of NWS river forecasts with the maximum lead time in advance of rising river levels. The river forecasts utilize ongoing precipitation as well as precipitation likely to occur; based on the best information and meteorological models available. The QPF are merged with observed precipitation data and input to the hydrologic model; and forecasts are produced for selected locations for seven days in the future.

In 1996 the NCRFC began implementation of the Ensemble Streamflow Prediction (ESP) system within our forecast domain (Braatz et al., 1997). With ESP, longer term (i.e., >14day) products are created based on the current hydrologic conditions, and the historical precipitation and temperature time-series (climate). For each year of available climate data, a streamflow hydrograph is produced, resulting in an ensemble of hydrographs conditioned to current model states. Through statistical analysis of these ensembles, a variety of probabilistic forecasts are produced for 360 locations.

* *Corresponding author address:* John Halquist, North Central RFC, NOAA/NWS, 1733 Lake Drive West, Chanhassen, MN 55317-8580; email: John.Halquist@noaa.gov

Although the deterministic and probabilistic forecasts are each useful and meet differing customer requirements, we have received increasingly frequent requests for a variety of short term products which provide more information than is available with our routine deterministic forecasts. The requests vary from some wanting less QPF while others would prefer more QPF, and often the requests conflict over a given area. While the NWS Hydrology Laboratory has been working towards delivering a forecast system to prepare short term probabilistic forecasts, this is not yet available for general River Forecast Center (RFC) use (Krzysztofowicz, et al. 1992, 1993, 1997) (Herr et al., 2002). We therefore began investigating methods for providing the type of information requested, while maintaining a link to realistic future hydrometeorological conditions. To this end, we chose to rely on

the expertise at the HPC and to use their best estimates of the most likely QPF fields.

2. DATA INPUTS

The NCRFC routine hydrologic forecasts are based on observed precipitation, temperatures and 24 hours of QPF provided by the HPC in 6-hour time steps. These 6-hour deterministic QPF are produced four times per day (0000, 0006, 0012, and 1800 UTC) and delivered to the RFCs. In 2003 the HPC began producing an experimental product based on Confidence Intervals (CI) using the Short Range Ensemble Forecasts (SREF) (Im, et al., 2006). By analyzing the HPC QPF, RFC observational data, and the SREF; the absolute error (AE) confidence intervals (CI) were derived. HPC provides predicted AE and CI twice a day (0000 and 1200 UTC) in 6-hour intervals (figure 1).

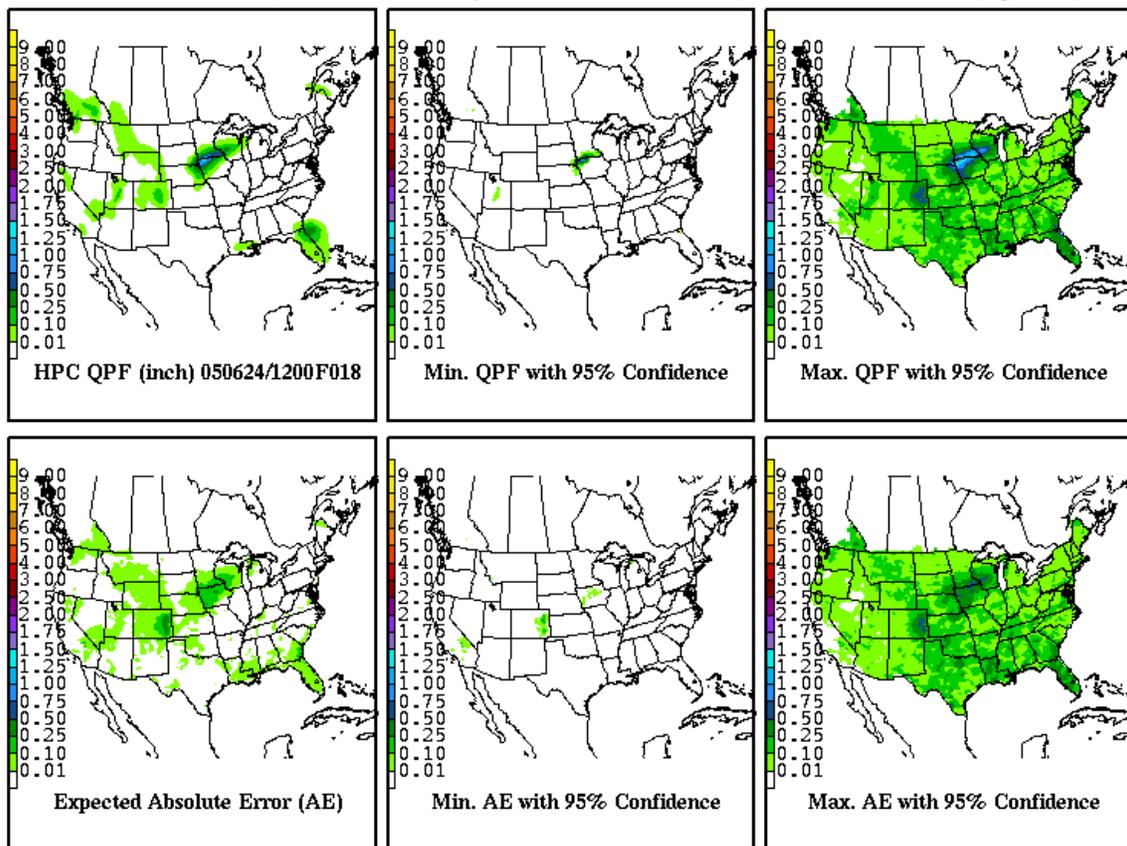


Figure 1: HPC QPF Confidence Interval Web graphics.

For our purposes, the 95% Confidence Interval for both maximum and minimum QPF are used. These represent the

maximum QPF (QPF_{CX}) and minimum QPF (QPF_{CN}) that can be expected with 95% confidence for a given 6-hour period. These

QPF CI forecasts are available on the Internet, and also as data grids.

3. ENSEMBLE PROCESS

During routine morning operations at NCRFC the HPC deterministic QPF (HQPF) are used in the updating of the hydrologic model and in the production of hydrologic forecasts. Typically, these model updates are completed by or before 1500 UTC. The QPF CI grids for the 1200 UTC forecast

cycle are made available at 1430 UTC, these grids are retrieved from the HPC and processed at the NCRFC to create the future mean areal precipitation (FMAP) inputs as well as graphical displays for local use. Copies of the operational NWSRFS datasets are made so that independent model simulations can be produced using the same initial conditions, but with varying QPF (Figure2).

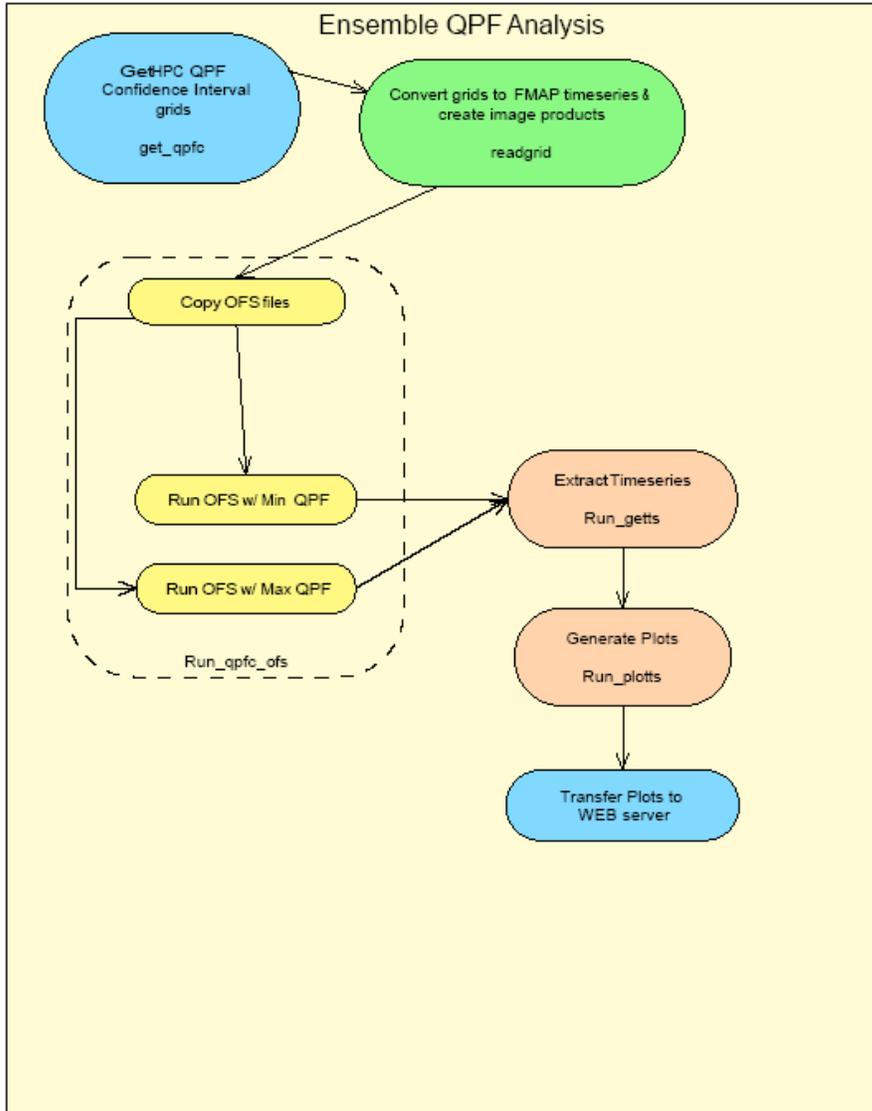


Figure 2: Diagram of hydrologic ensemble process

This results in three independent solutions based on HQPF, QPFCX, and QPFCN. River stage time-series are extracted from

the NWSRFS for selected forecast locations and are used in the generation of graphical displays (Figure 3).

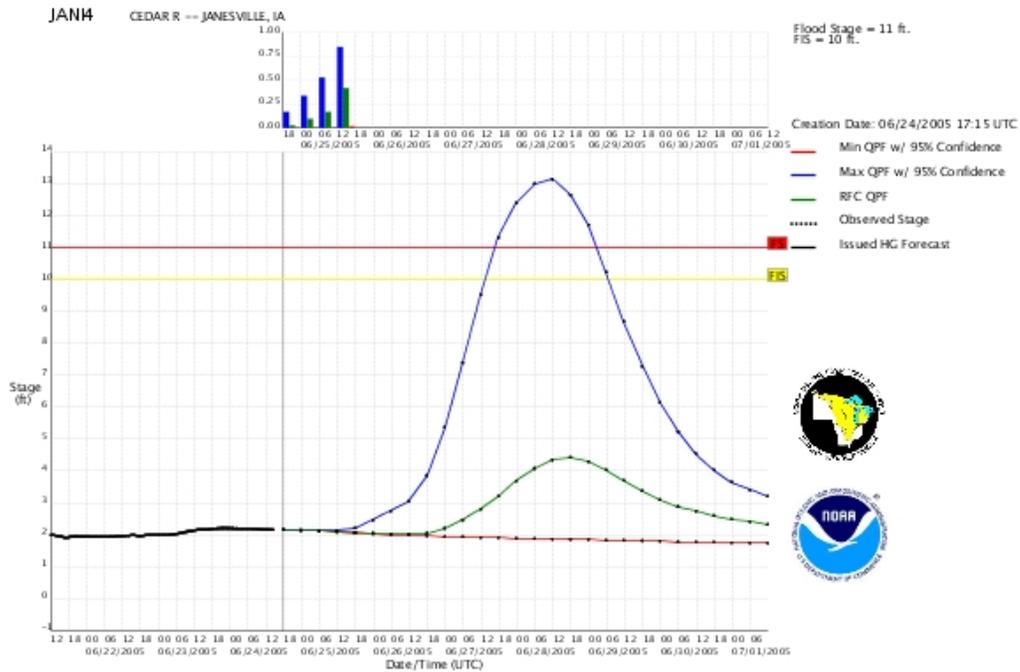


Figure 3: Sample ensemble hydrograph plot

Graphics are produced in several image formats, including Scalable Vector Graphics (SVG) which allow for interactive query of stage and times from the hydrographs.

An example of how these graphics would be used is illustrated by Figure 4 which shows the ensemble created on the 25th of June. The HQPF and QPFCN hydrographs were clearly under forecast, based on the observed data (green line) while the QPFCX hydrograph indicated a significant rise and provided a reasonable upper limit for potential river stages.

4. ANALYSIS

The ensemble forecasts for the period April through September 2005 were analyzed using traditional statistical measures commonly used in the NWS (NWSI 10-1601, 2004) and as recommended in Wilks (1995). For the purposes of this paper, the analysis will be presented for locations on Cedar River and tributaries in Minnesota and Iowa

(Figure 5). The focus is on the analysis of the hydrologic model ensemble members rather than the QPF inputs, and data are analyzed both by location and as a combined dataset of all locations. The equations used are contained in Appendix 1.

In examining the data for all forecast locations combined with sample sizes of 11,593 forecast/observed pairs, and 49,211 modeled/observed pairs (QPF, QPFCN, QPFCX), we see favorable results (Figures 6 – 9). Forecast data are actual river stage forecasts issued by the NCRFC and are typically issued based on river stages exceeding a given level; modeled data are available regardless of river stage. In general, when combining the full suite of measures, although not for each individual measure, the statistics for the forecast are slightly better than any of the model output statistics (Table 1). For the model ensembles, the HPC QPF produces better results than either of the QPF CI QPF members.

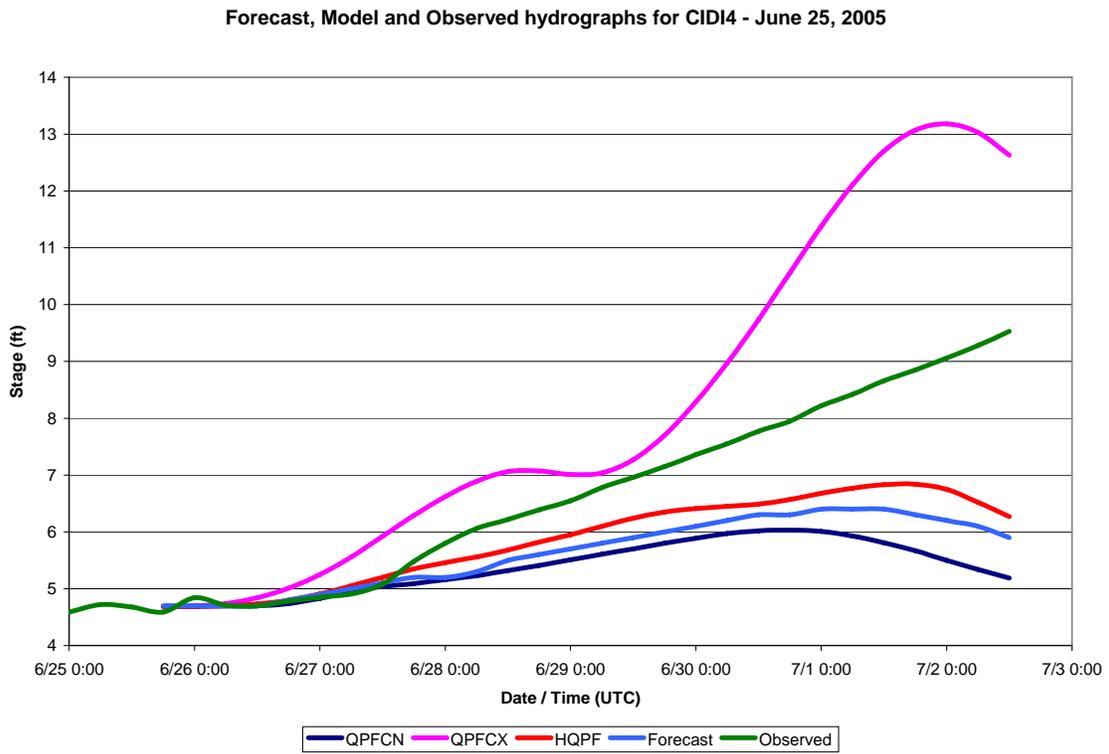


Figure 4: Example plot showing ensembles generated on June 25, 2005 with actual observed

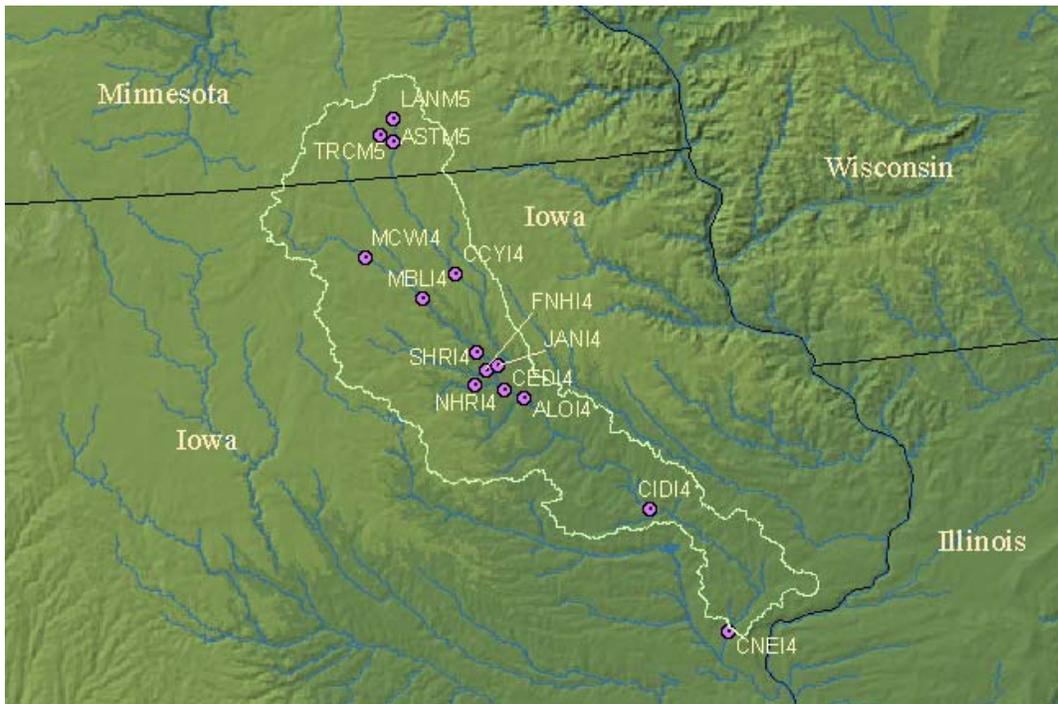


Figure 5: Map showing Cedar River Basin

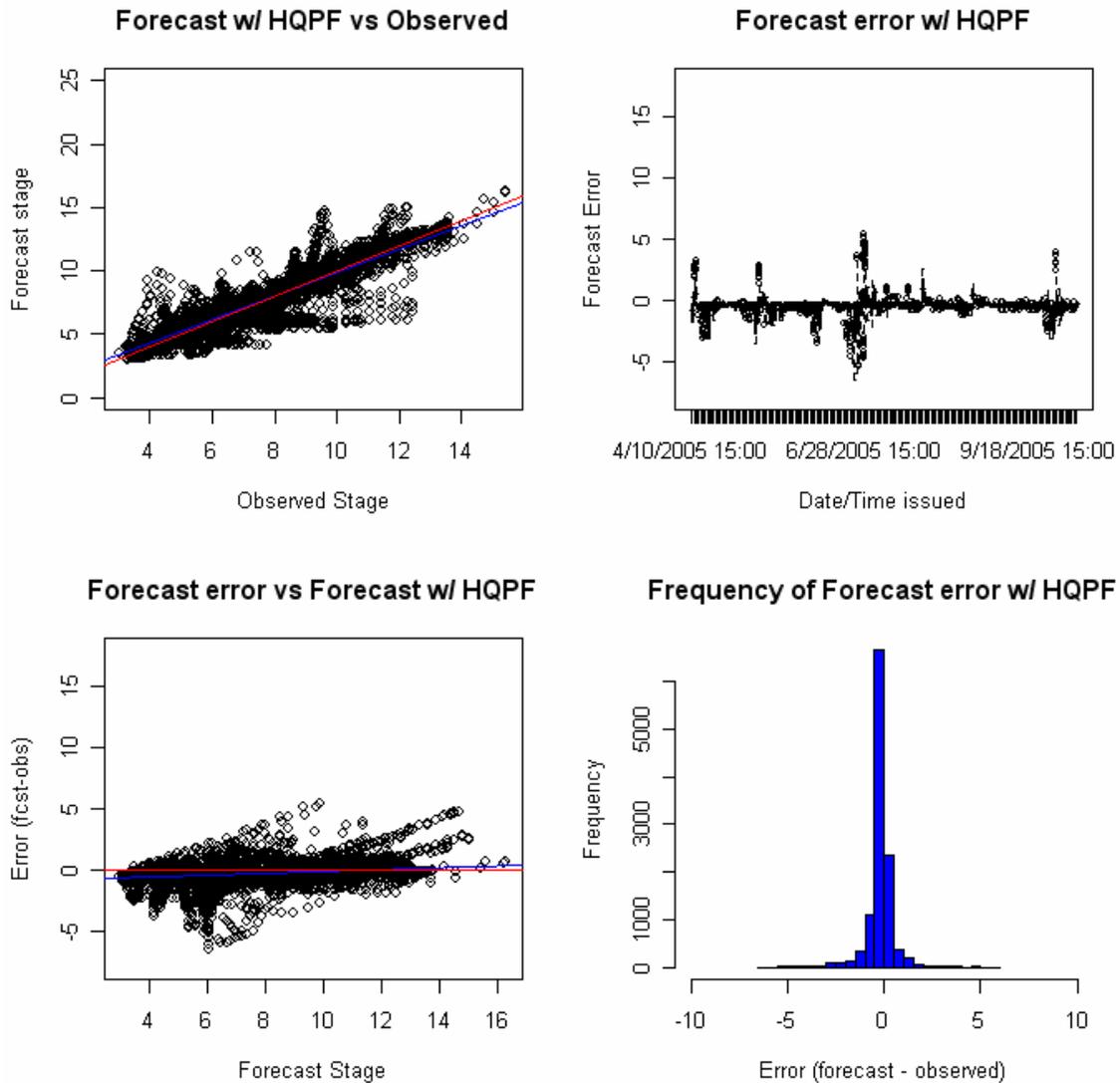


Figure 6: Forecasts w/ QPF vs Observed [Blue line is fitted to data, Red line is perfect]

lm(formula = fe ~ ff) [Applies to fitted line for upper left graphic]

Residuals:

Min	1Q	Median	3Q	Max
-6.05709	-0.16417	0.09379	0.25194	5.56672

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.886276	0.020163	-43.95	<2e-16 ***
ff	0.072682	0.003059	23.76	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.695 on 11590 degrees of freedom
 Multiple R-Squared: 0.04646, Adjusted R-squared: 0.04638
 F-statistic: 564.7 on 1 and 11590 DF, p-value: < 2.2e-16

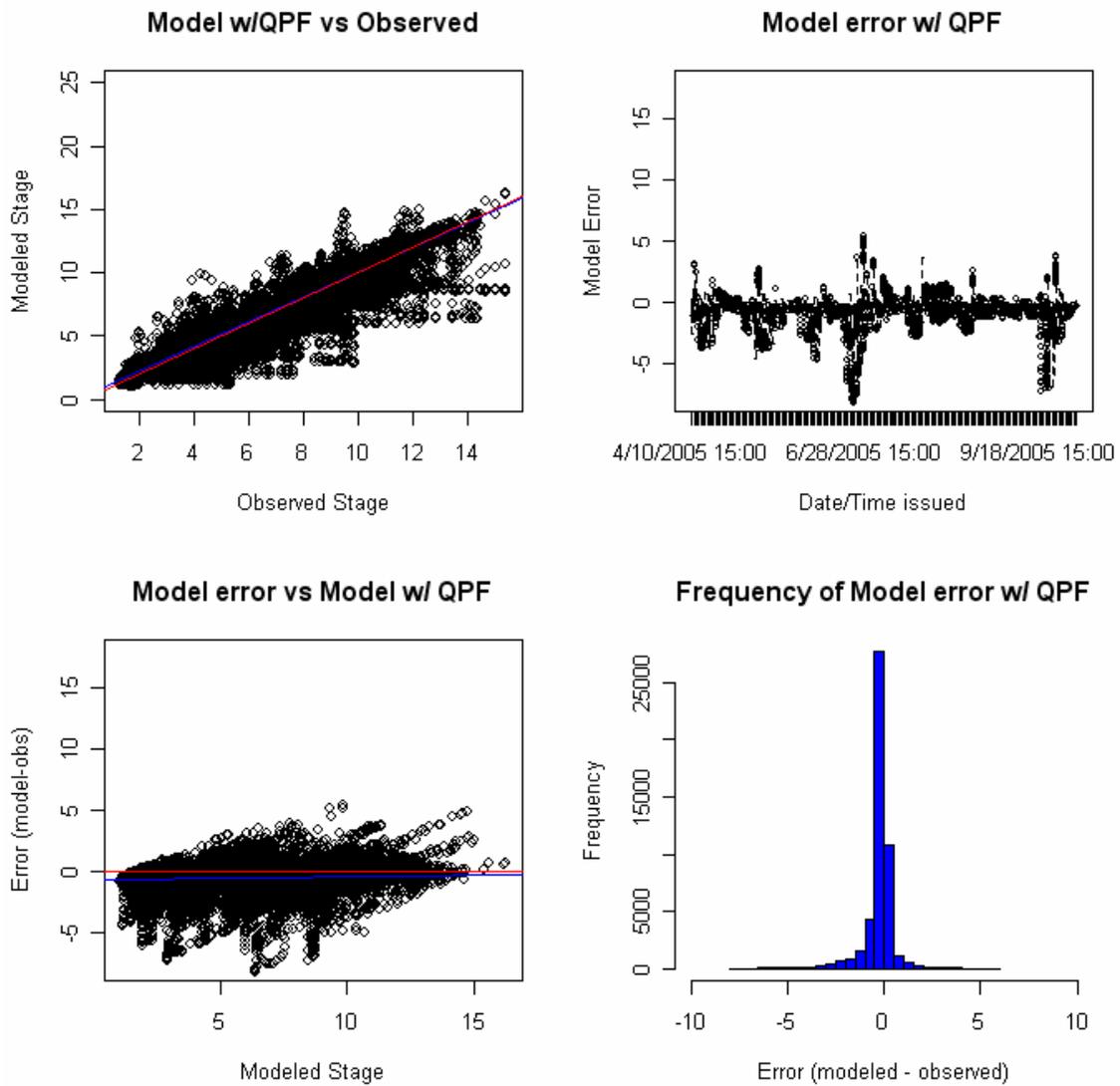


Figure 7: Modeled w/ QPF vs Observed [Blue line is fitted to data, Red line is perfect]

lm(formula = me ~ fa) [Applies to fitted line in upper left graphic]

Residuals:

Min	1Q	Median	3Q	Max
-7.80095	-0.07796	0.13868	0.29935	5.81114

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.660229	0.008201	-80.50	<2e-16 ***
fa	0.025263	0.001300	19.43	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7834 on 49208 degrees of freedom
 Multiple R-Squared: 0.007612, Adjusted R-squared: 0.007591
 F-statistic: 377.4 on 1 and 49208 DF, p-value: < 2.2e-16

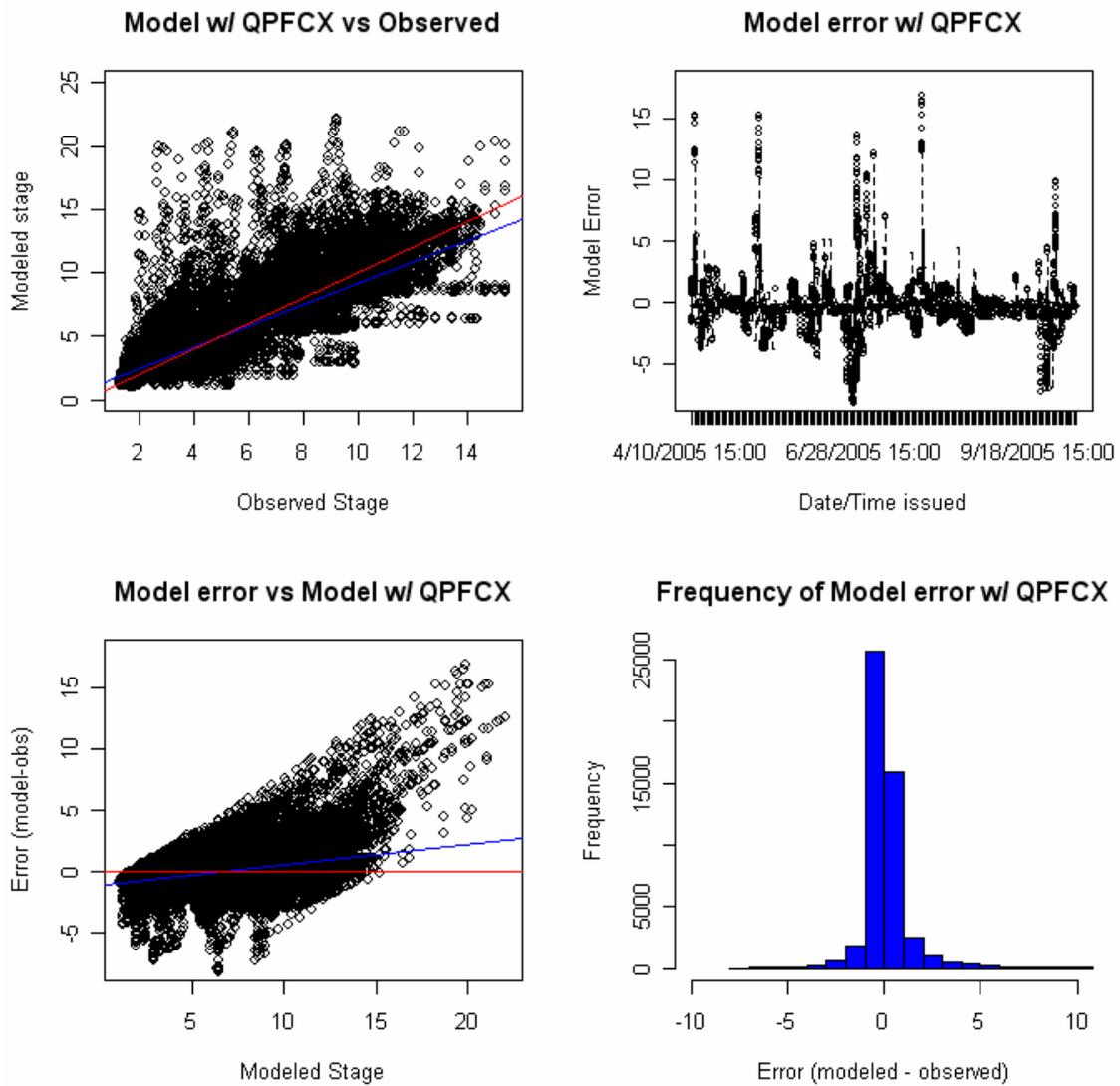


Figure 8: Modeled w/ 95% CI Maximum QPF vs Observed [Blue line is fitted to data, Red line is perfect]

lm(formula = xe ~ cx) [Applies to fitted line in upper left graphic]

Residuals:

Min	1Q	Median	3Q	Max
-8.21071	-0.50076	-0.04517	0.39570	14.86353

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-1.134978	0.012280	-92.43	<2e-16 ***
cx	0.163900	0.001817	90.20	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.199 on 49208 degrees of freedom

Multiple R-Squared: 0.1419, Adjusted R-squared: 0.1419

F-statistic: 8136 on 1 and 49208 DF, p-value: < 2.2e-16

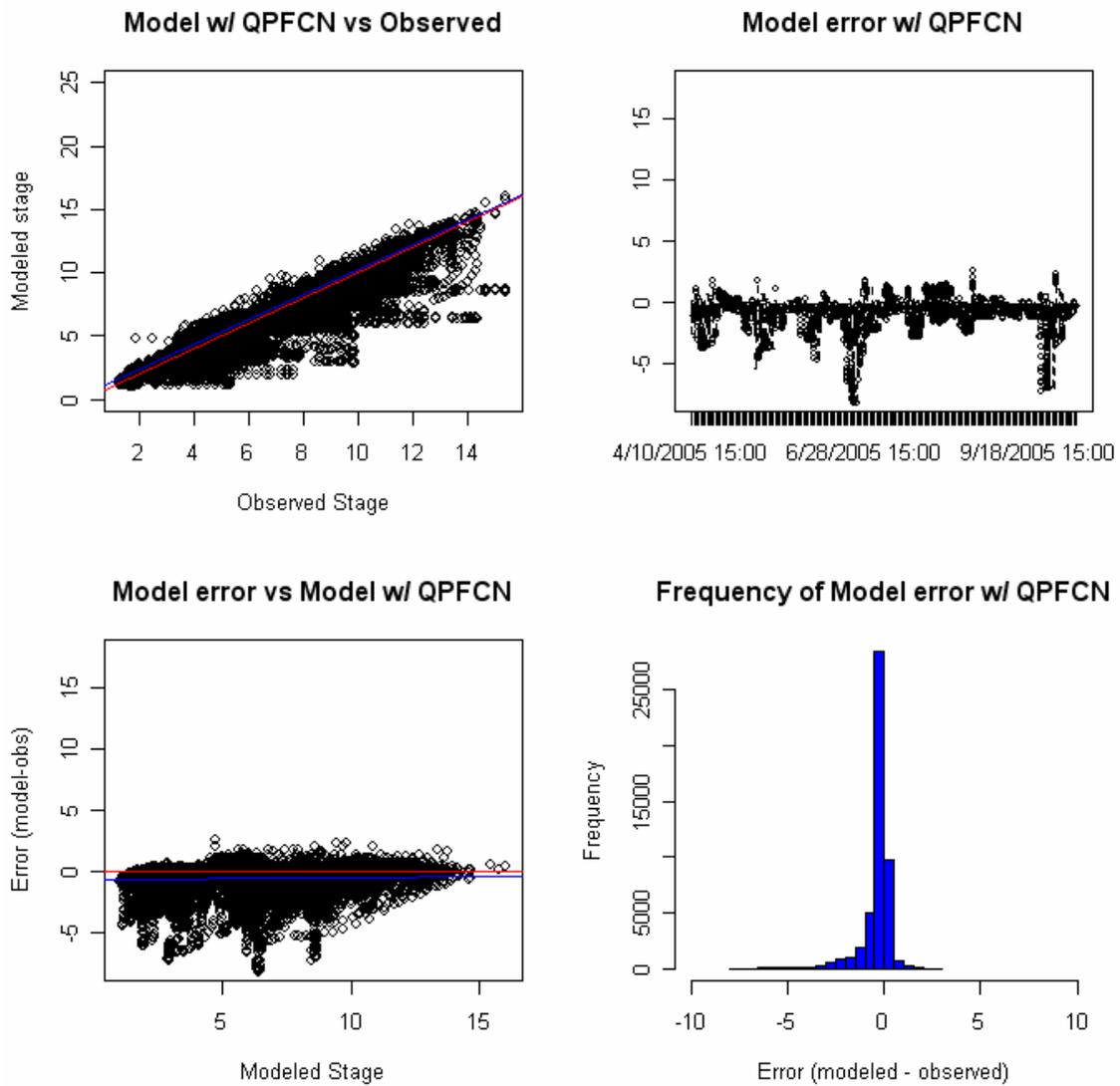


Figure 9: Modeled w/ 95% CI Minimum QPF vs Observed [Blue line is fitted to data, Red line is perfect]

lm(formula = ne ~ cn) [Applies to fitted line in upper left graphic]

Residuals:

Min	1Q	Median	3Q	Max
-7.70872	-0.05017	0.19555	0.34640	3.11521

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.682721	0.008173	-83.53	<2e-16 ***
cn	0.014243	0.001315	10.83	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7833 on 49208 degrees of freedom
 Multiple R-Squared: 0.002379, Adjusted R-squared: 0.002359
 F-statistic: 117.3 on 1 and 49208 DF, p-value: < 2.2e-16

Table 1: Combined dataset (10 forecast locations)

Combined dataset				
	Forecast 11,593 points	HQPF 49,211 pts.	QPFCN 49,211 pts.	QPFCX 49,211 pts.
BIAS	0.874	0.855	0.791	1.068
POD	0.773	0.777	0.747	0.819
FAR	0.115	0.091	0.056	0.233
CSI	0.702	0.721	0.715	0.656
PC	70.229	72.052	71.490	65.618
ME	-0.132	-0.201	-0.286	0.170
MAE	0.384	0.389	0.397	0.580
RMSE	0.724	0.777	0.797	1.297

Appendix 2 contains the individual basin scores, and shows that in most cases the HPC QPF performs best with an under forecasting bias, the QPF CI Minimum also under forecasting and the QPF CI Maximum over forecasting. While these results confirm our expectations, there are exceptions, such as the location FNHI4 with a much better BIAS score with the QPF CI Maximum; however the POD, FAR and CSI are minimally or negatively impacted.

One of our initial assumptions was that QPF would contribute to a larger error as basin size increased, due to the accumulation of any upstream errors. However, in comparing each score based on the total basin area for a given location, no clear signal is present to indicate that errors are directly related to basin size (Figure 10) (Table 2). Figures 11 and 12 show the modeled and observed data for CIDI4 and FNHI4.

Due to precipitation variability and given that only one day of future precipitation is input to the hydrologic model, one can assume that with model simulations seven days in the future there would be more uncertainty or error than for the first day. Figures 13 and 14 illustrate the error spread as related to time. The data show that for larger

downstream locations (CIDI4), QPF data outliers contribute to a +/- 2 foot maximum error spread out to five days, while for a smaller headwater locations (FNHI4) this same error spread is exceeded by the third day. However the upper and lower quartiles (as illustrated by the boxes) show minimal error throughout the full seven day period. It is also evident that the 95% Maximum QPF results in a larger spread, but when outliers are ignored, contributes a realistic upper limit.

5. CONCLUSIONS

The QPF Confidence Interval forecasts produced by the HPC provide a means for computing the expected upper and lower limits for short term deterministic river forecasts. Ongoing analysis over a longer period and with additional locations will be essential to demonstrate the full value of the methods presented here. Future improvements and enhancements to the HPC QPF process will contribute directly to improvements in these ensembles.

Further analyses are planned, particularly with respect to the impacts QPF timing has on river forecasting, and varying the number of hours of QPF used for additional model

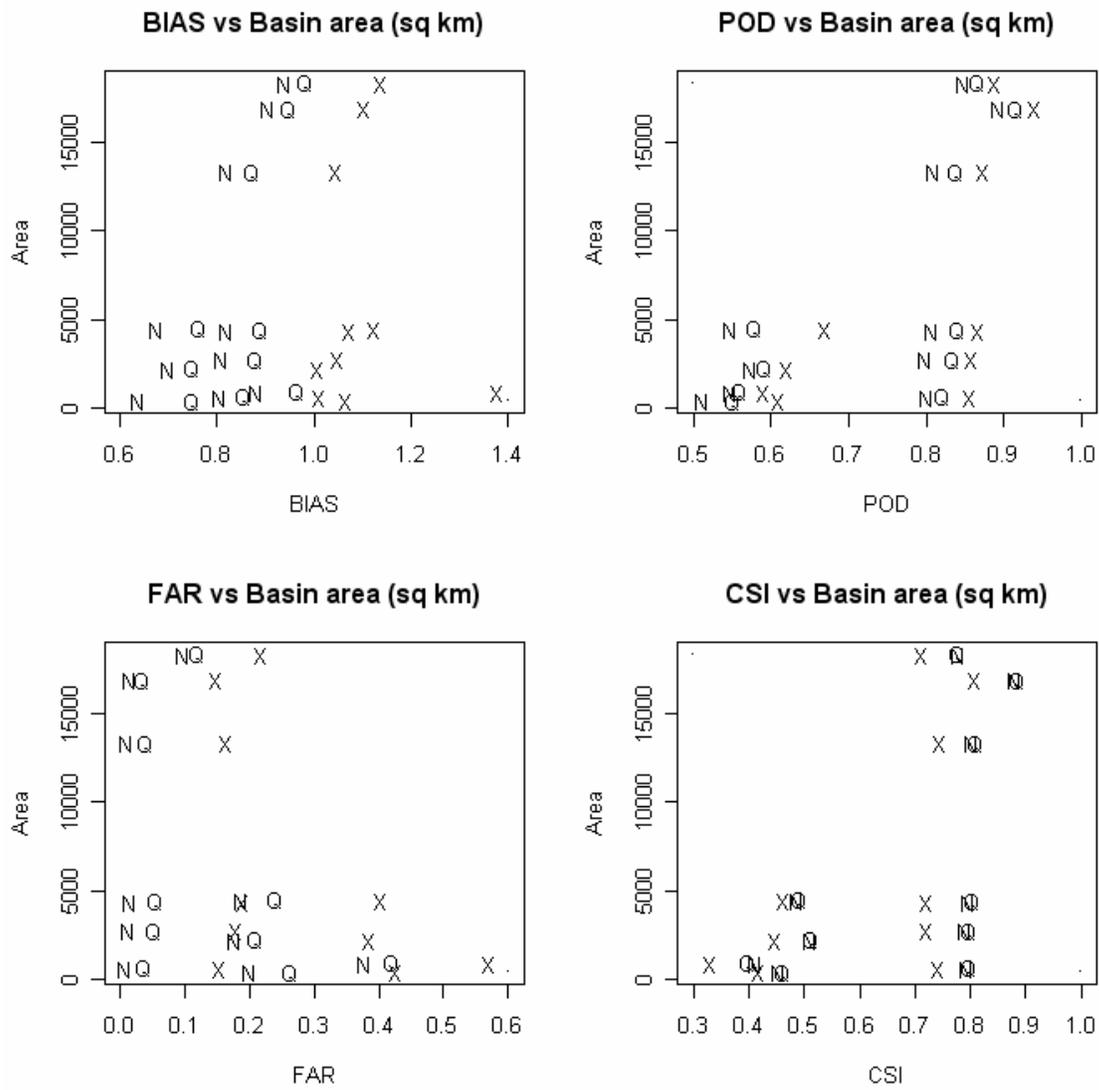


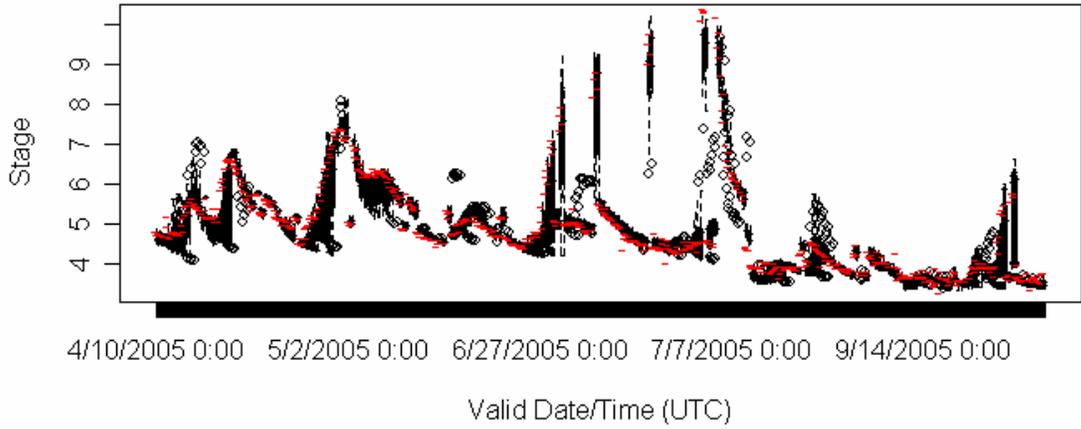
Figure 10: Plots of Statistical measures against basin area.

Q = Modeled w/ HQPF
 X = Modeled w/ QPFCX
 N = Modeled w/ QPFCN

Table 2: Correlation coefficients for BIAS vs basin area and CSI vs basin area

	Correlation (Pearson)			Correlation (Spearman)			
	HQPF	QPFCX	QPFCN	HQPF	QPFCX	QPFCN	
BIAS	0.56526	-0.01674	0.633958	BIAS	0.563636	0.357576	0.650459
CSI	0.590765	0.594267	0.594173	CSI	0.547115	0.50303	0.527273

CIDI4 Modeled w/QPF



CIDI4 Modeled w/95% Max QPF

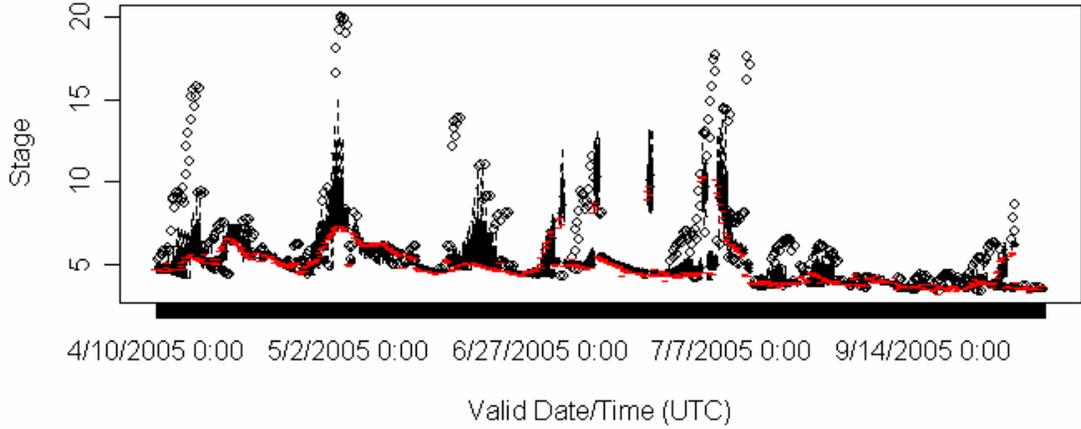
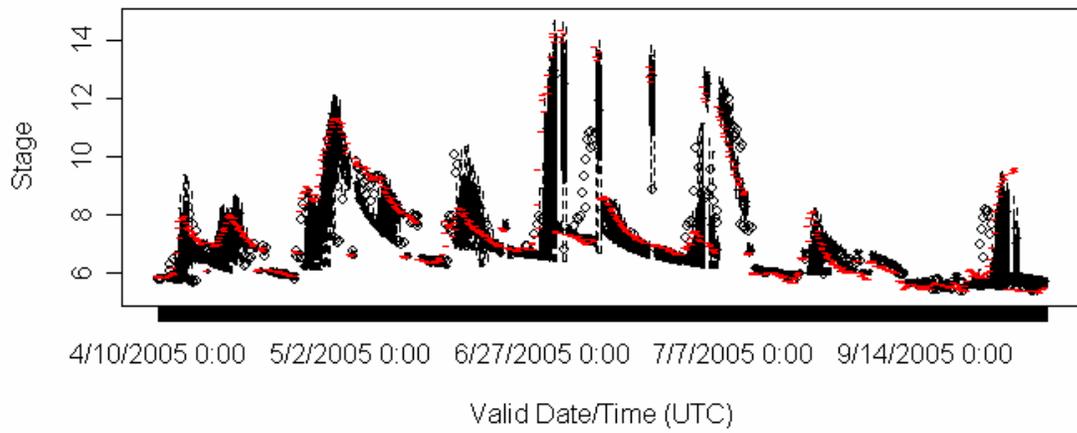


Figure 11: CIDI4 modeled time-series w/ HQPF and QPFCX overlaid with observations (-)

FNHI4 Modeled w/QPF



FNHI4 Modeled w/95% Max QPF

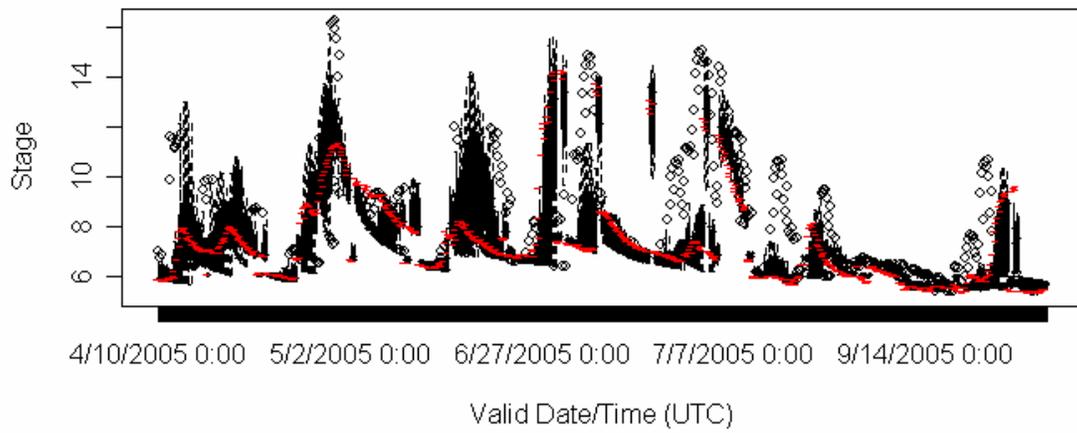
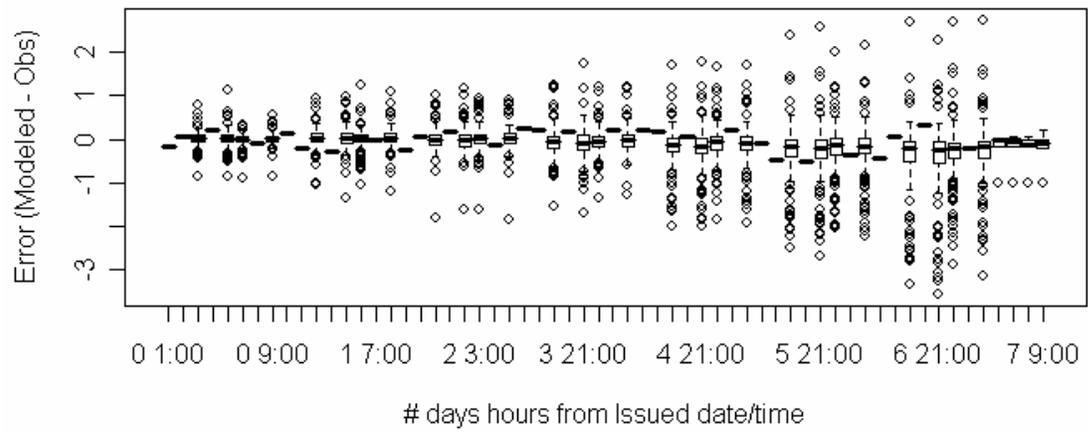


Figure 12: FNHI4 modeled time-series w/ HQPF and QPFCX overlaid with observations (-)

CIDI4 Model Error w/QPF vs Time in future



CIDI4 Model Error w/95% Max QPF vs Time in future

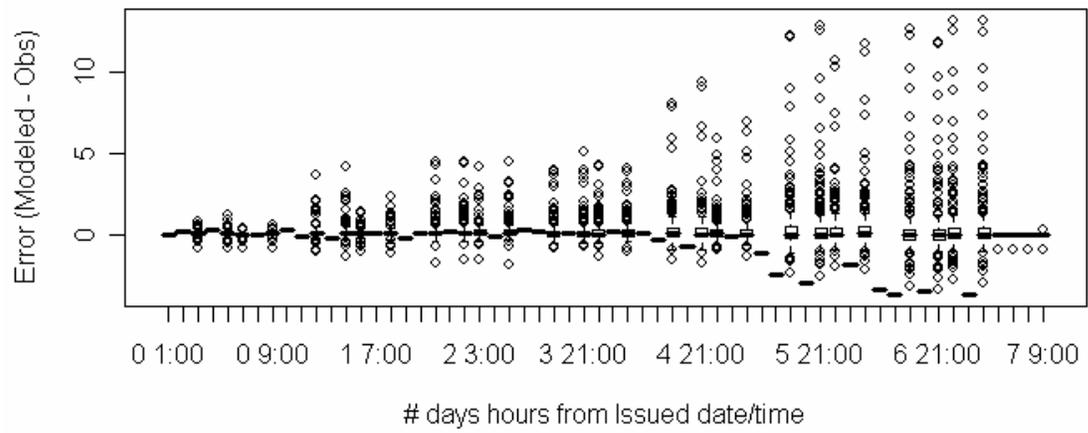
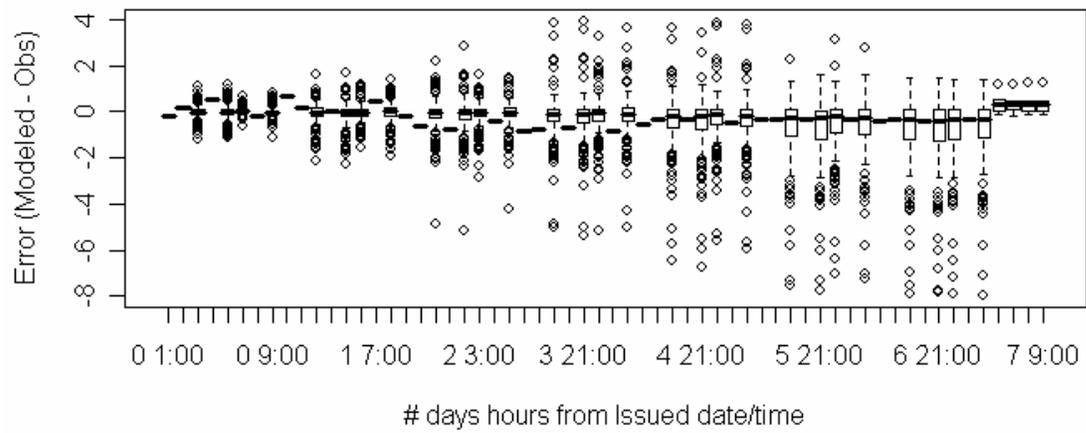


Figure 13: Model Errors plotted against the number of days and hours into the future from the time of model run

FNHI4 Model Error w/QPF vs Time in future



FNHI4 Model Error w/95% Max QPF vs Time in future

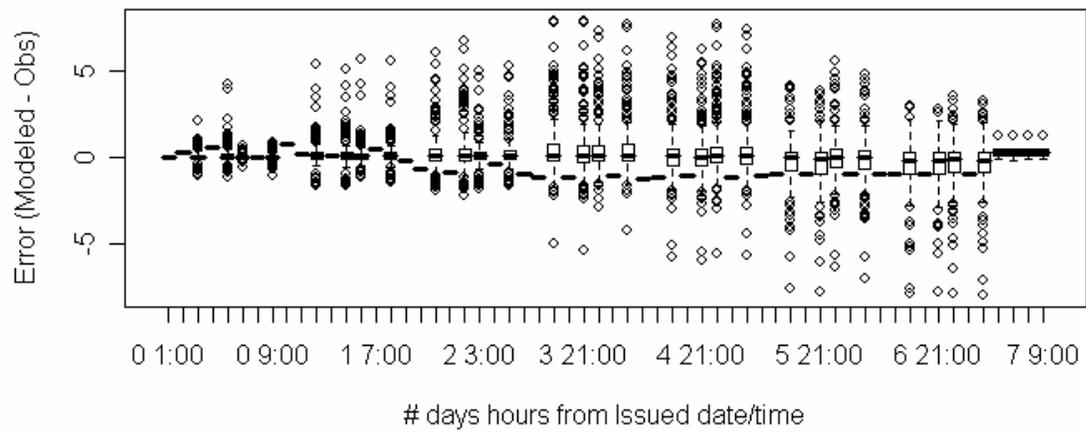


Figure 14: Model Errors plotted against the number of days and hours into the future from the time of model run

simulations. There is also potential for applying filters to the model results to produce bias adjusted results. And as more data are available, analysis of possible seasonal biases will be investigated.

We believe that the results presented here offer sufficient foundation to begin making these products available to our customers, for their use, and look forward to their comments.

6. REFERENCES

Braatz, D. T., J. B. Halquist, R. J Wavrin, J. Ingram, J. J. Feldt, and M. S. Longnecker, 1997; NWS Hydrologic products and services: Moving from the Traditional to the Technically Advanced, *13th International Conference on Interactive Information and Processing System for Meteorology, Oceanography, and Hydrology*, American Meteorological Society, Long Beach, CA, J63-J69

Herr, H., E. Welles, M. Mullusky, L. Wu, J. Schaake, 2002; Simplified Short Term Precipitation Ensemble Forecasts: Theory, Preprints, *Symposium on Observations, Data Assimilation, and Probabilistic Prediction 16th Conference on Probability and Statistics in the Atmospheric Sciences*, Orlando, FL

Im, J.-S., K. Brill, and E. Danaher, 2006: Confidence interval estimation for quantitative precipitation forecasts (QPF) using short-range ensemble forecasts (SREF). *Weather and Forecasting*, in press.

Krzysztofowicz, R. and T. R. Drake, 1992; Probabilistic quantitative precipitation forecasts for river forecasting. Preprints, *Symposium on Weather Forecasting, 72nd Annual Meeting of the American Meteorological Society*, American Meteorological Society, Atlanta, GA, 66-71.

Krzysztofowicz, R., 1993; Probabilistic hydrometeorological forecasting system: A conceptual design. *NOAA Technical Memorandum NWS ER-87*, National Oceanic and Atmospheric Administration, U.S. Department of Commerce

Krzysztofowicz, R., T. E. Adams, and T. Rossi, 1997; Probabilistic river stage forecasting. *Final Report to University Corporation for Atmospheric Research, COMET Cooperative Project, Sub-awards No. UCAR S94-43833 and S96-75677*, 1 January 1994 – 31 December 1996, 15 pp.

Wilks, D. S. 1995; *Statistical Methods in the Atmospheric Sciences: An Introduction*. Academic Press, 467 pp.

APPENDIX 1.

Variable definitions:

H = hits, forecast = observed
F = false alarm, forecasted but not observed
M = missed event, observed but not forecast
Fcst = forecast or modeled value
Obs = observed value

Bias perfect score = 1

$$BIAS = \frac{H + F}{H + M}$$

Probability of Detection (POD) perfect = 1

$$POD = \frac{H}{H + M}$$

False alarm ratio (FAR) perfect = 0

$$FAR = \frac{F}{H + F}$$

Critical success index (CSI) perfect = 1

$$CSI = \frac{H}{H + M + F}$$

Percent correct (PC) perfect = 100

$$PC = \frac{\sum_{i=1}^j H_i}{n} \times 100$$

Mean error (ME) perfect = 0

$$ME = \frac{1}{n} \sum_{i=1}^n (Fcst_i - Obs_i)$$

Mean absolute error (MAE) perfect = 0

$$MAE = \frac{1}{n} \sum_{i=1}^n |Fcst_i - Obs_i|$$

Root mean square error (RMSE) perfect = 0

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Fcst_i - Obs_i)^2}$$

APPENDIX 2.

BIAS			
	QPF	QPFCN	QPFCX
ALOI4	0.870	0.819	1.045
ASTM5	0.852	0.807	1.010
CCYI4	0.879	0.810	1.050
CIDI4	0.946	0.906	1.104
CNEI4	0.982	0.941	1.139
FNHI4	0.748	0.698	1.009
JANI4	0.888	0.819	1.073
LANM4	0.746	0.639	1.066
NHRI4	0.963	0.881	1.379
SHRI4	0.760	0.674	1.124

POD			
	QPF	QPFCN	QPFCX
ALOI4	0.837	0.810	0.873
ASTM5	0.821	0.800	0.856
CCYI4	0.833	0.799	0.859
CIDI4	0.914	0.893	0.940
CNEI4	0.866	0.849	0.890
FNHI4	0.591	0.574	0.620
JANI4	0.840	0.807	0.868
LANM4	0.551	0.511	0.610
NHRI4	0.559	0.548	0.592
SHRI4	0.579	0.548	0.671

FAR			
	QPF	QPFCN	QPFCX
ALOI4	0.038	0.011	0.165
ASTM5	0.037	0.008	0.153
CCYI4	0.052	0.013	0.181
CIDI4	0.034	0.014	0.148
CNEI4	0.118	0.098	0.218
FNHI4	0.209	0.178	0.385
JANI4	0.053	0.015	0.191
LANM4	0.262	0.201	0.427
NHRI4	0.419	0.378	0.571
SHRI4	0.238	0.188	0.403

CSI			
	QPF	QPFCN	QPFCX
ALOI4	0.810	0.803	0.745
ASTM5	0.796	0.795	0.741
CCYI4	0.796	0.791	0.722
CIDI4	0.885	0.882	0.808
CNEI4	0.776	0.777	0.713
FNHI4	0.511	0.511	0.447
JANI4	0.802	0.797	0.720
LANM4	0.461	0.453	0.419
NHRI4	0.398	0.411	0.331
SHRI4	0.490	0.486	0.462

Basin Area (sq km)		
	Local	Total Upstream
ALOI4	318	13360
ASTM5	129	634
CCYI4	1688	2742
CIDI4	1271	16922
CNEI4	1435	18357
FNHI4	402	2217
JANI4	792	4359
LANM5	401	401
NHRI4	901	901
SHRI4	1106	4477

Location Key

- ALOI4 – Cedar R at Waterloo, IA
- ASTM5 – Cedar R at Austin, MN
- CCYI4 – Cedar R at Charles City, IA
- CIDI4 – Cedar R at Cedar Rapids, IA
- CNEI4 – Cedar R at Conesville, IA
- FNHI4 – W Fork Cedar R at Finchford, IA
- JANI4 – Cedar R at Janesville, IA
- LANM5 – Cedar R at Lansing, MN
- NHRI4 – Beaver Ck at New Hartford, IA
- SHRI4 – Shell Rock R at Shell Rock, IA