MEDIUM-RANGE ENSEMBLE PRECIPITATION AND STREAMFLOW FORECASTING FOR THE UPPER TRINITY RIVER BASIN IN TEXAS VIA THE NWS HYDROLOGIC ENSEMBLE FORECAST SERVICE

by

HOSSEIN SADEGHI

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE IN CIVIL ENGINEERING

THE UNIVERSITY OF TEXAS AT ARLINGTON

December 2015

Acknowledgements

I would like to express my sincere gratitude to my advisor Dr. Dong-Jun Seo for the continuous support of my Master's study and related research and for his patience. His guidance helped me at all times of the research and writing of this thesis.

I would like to thank the rest of my thesis committee: Dr. Nick Fang and Dr. Habib Ahmari for their insightful comments and encouragements.

A very special thank-you goes out to Dr. Sunghee Kim; without her technical support, it would not have been possible to conduct this research.

My sincere thanks also go to Dr. Haksu Lee, Dr. James Brown, Dr. Hank Herr, Dr. Limin Wu, Mr. Bob Corby, and Mr. Frank Bell for help at various stages of this research. I would like to acknowledge Mr. Andrew Philpott for his help and also for his support with providing the data.

I thank my lab mates in the Hydrology and Water Resources Laboratory (HWRL) for the stimulating discussions, for the days and nights when we were working together, and for all the fun we have had in the last two and a half years.

Last but not least, I would like to thank my family. This thesis would not have been possible without their love and support. I owe my deepest gratitude to my wife, Elaheh, for her love, help, and support; to my parents for teaching me the importance of learning and all of their sacrifices for me; to my sisters, Mahsan and Parisa, for their emotional support; to my brother- and sister-in-law, Ali and Azade, and their lovely daughter, Arta, for their love and support during this journey. .

November 23, 2015

Abstract

MEDIUM-RANGE ENSEMBLE PRECIPITATION AND STREAMFLOW FORECASTING FOR THE UPPER TRINITY RIVER BASIN IN TEXAS VIA THE NWS HYDROLOGIC ENSEMBLE FORECAST SERVICE

HOSSEIN SADEGHI, MS

The University of Texas at Arlington, 2015

Supervising Professor: Dong-Jun Seo

Compared to forecasts of short-term precipitation accumulations (daily or shorter) at lead times larger than a few days, those of longer-term accumulations (3-daily or longer) are significantly more skillful owing to the larger temporal scale of aggregation. If one can utilize this skill present in medium-range precipitation forecast in hydrologic prediction, it is very likely that the lead time of hydrologic forecasts, in particular, of streamflow and soil moisture may be extended. Though forecasts of longer-term accumulations of precipitation are more skillful than those of shorter-term accumulations, precipitation forecasts in general are too uncertain to be used as deterministic, or single-valued, input.

The main goal of this study is to increase forecast lead time of streamflow forecasts by using medium range ensemble precipitation forecasts. A premise for this study is that, in the ensemble paradigm, forecasting of precipitation and streamflow provides extending forecast lead time with improved forecast skill. To utilize forecast skill in medium range precipitation forecasts in the ensemble paradigm, this study uses Hydrologic Ensemble Forecast Service (HEFS).

In the HEFS, the Meteorological Ensemble Forecast Processor (MEFP) was used to generate ensemble precipitation hindcasts using the Global Ensemble Forecast

System (GEFS) reforecast data. Raw streamflow hindcasts were generated via the Community Hydrologic Prediction System (CHPS) using the Sacramento Soil Moisture Accounting model (SAC-SMA) and unit hydrograph. To reduce biases and uncertainties in the hydrologic model results, raw streamflow ensembles were post-processed by the Ensemble Postprocessor (EnsPost). The precipitation, raw and post-processed streamflow ensembles were verified using the Ensemble Verification System (EVS) to assess the quality of hindcasts. Ensemble hindcasts of precipitation and streamflow were generated using the HEFS for a 26-year period between 1986 and 2011. The study area consisted of five headwater basins located upstream of the Dallas-Fort Worth (DFW) metropolitan area in the Upper Trinity River Basin in Texas.

The main findings of this study include: (1) adjusting modulation canonical events is a very effective way to improve predictive skill in ensemble forecasts of precipitation, raw, and post-processed streamflow forecasts: (2) GEFS-forced medium-range precipitation hindcasts for the study area have valuable skill in 1-, 3-, 5-daily, weekly, and biweekly-aggregated hindcasts; (3) in the ensemble paradigm, forecast skill in medium-range precipitation forecasts can be effectively utilized to improve the quality of streamflow forecasts in extended forecast lead time via HEFS.

This study used the HEFS successfully, demonstrating the HEFS's portability in the Unix/Linux environment outside of National Weather Service (NWS). This study also showed that the HEFS is an effective tool for generating skillful forecasts of precipitation and streamflow ensembles. This study would provide water resources managers with improved streamflow forecasts for the extended forecast lead time to effectively manage water resources and to mitigate water-related hazards.

Table of Contents

List of Illustrations

List of Tables

Chapter 1

Introduction

Medium-range (~ 2 weeks) forecasting of precipitation is critical to meeting different types of user needs for operational hydrology and water resources management (Yuan et al., 2014). For example, some users in the Eastern US may be interested in river forecasts with a lead time of three to seven days to manage and mitigate the impact of flooding (Adams and Ostrowski, 2010) whereas those in the Western US may be interested in weekly or longer-lead forecast of inflow into water supply reservoirs. Skillful medium-range forecasting of precipitation is particularly important in areas that are prone to extreme events such as floods and droughts. In Texas, for example, a severe drought which lasted for four and a half years since 2011 ended with a historical flooding event in May 2015 which resulted in 28 fatalities (https://www.climate.gov/news-features/event-tracker/flood-disaster-texas-and-oklahoma).

For short-range forecasting of precipitation, the National Weather Service (NWS) West Gulf River Forecast Center (WGRFC) produces and use quantitative precipitation forecast (QPF) in the current practice. The WGRFC QPF is single-valued precipitation forecasts with 72-hour forecast lead time. Operationally, only the first 6-hour of QPF is used to generate river forecasts, although the entire 72-hour forecasts may be used, depending on specific weather events (WGRFC, 2015). The reason for this practice is that, due to the very limited predictive skill in QPF particularly for convective precipitation, inputting longer-lead single-valued QPF generally results in longer-lead single-valued river stage forecasts with unacceptably large errors (Regonda, 2013).

Compared to forecasts of short-term precipitation accumulations (daily or shorter) at lead times larger than a few days, those of longer-term accumulations (3-daily or longer) are significantly more skillful owing to the larger temporal scale of aggregation

(Brown at el., 2014b). If one can utilize this skill present in medium-range precipitation forecast in hydrologic prediction, it is very likely that the lead time of hydrologic forecasts, in particular, of streamflow and soil moisture may be extended. Though forecasts of longer-term accumulations of precipitation are more skillful than those of shorter-term accumulations, precipitation forecasts in general are too uncertain to be used as deterministic, or single-valued, input. If, on the other hand, precipitation forecasts are expressed as ensembles or in probabilistic terms, one may produce ensemble or probabilistic hydrologic forecasts, with which the users can make risk-based decisions (Demargne et al., 2014; Seo et al., 2010).

There are many sources of uncertainty in streamflow forecasts: errors in meteorological input, structural errors in the hydrologic model, parametric errors in the hydrologic model, errors in the hydrologic model initial conditions, and human control and alternations of flow and the hydrologic cycle. The uncertainties arising from these sources of error propagate through the modeling system to degrade the quality of hydrologic forecasts (Brown and Heuvelink, 2005). Because such uncertainty information cannot be conveyed in a deterministic, or single-valued, forecast, it is necessary to use probabilistic forecasting methods. Toward that end, ensemble forecasting has gained great popularity in many disciplines because it may the only practical methodology available today that is general enough for operational hydrologic forecasting (Seo et al., 2006; Cloke and Pappenberger, 2009; Nester et al., 2012; Demargne et al., 2014).

To operationalize ensemble hydrologic forecasting, the U.S. National Weather Service Office of the Hydrologic Development (NWS/OHD, now the NWS/National Water Center) has recently developed the Hydrologic Ensemble Forecast Service (HEFS; Demargne et al., 2014; Seo et al., 2010). The HEFS ensemble generation package includes the Meteorological Ensemble Forecast Processor (MEFP), the MEFP Parameter

Estimator (MEFPPE), the Hydrologic Ensemble Processor (HEP), the Ensemble Post Processor (EnsPost), and the EnsPost Parameter Estimator (EnsPostPE) (see Figure 1.1).

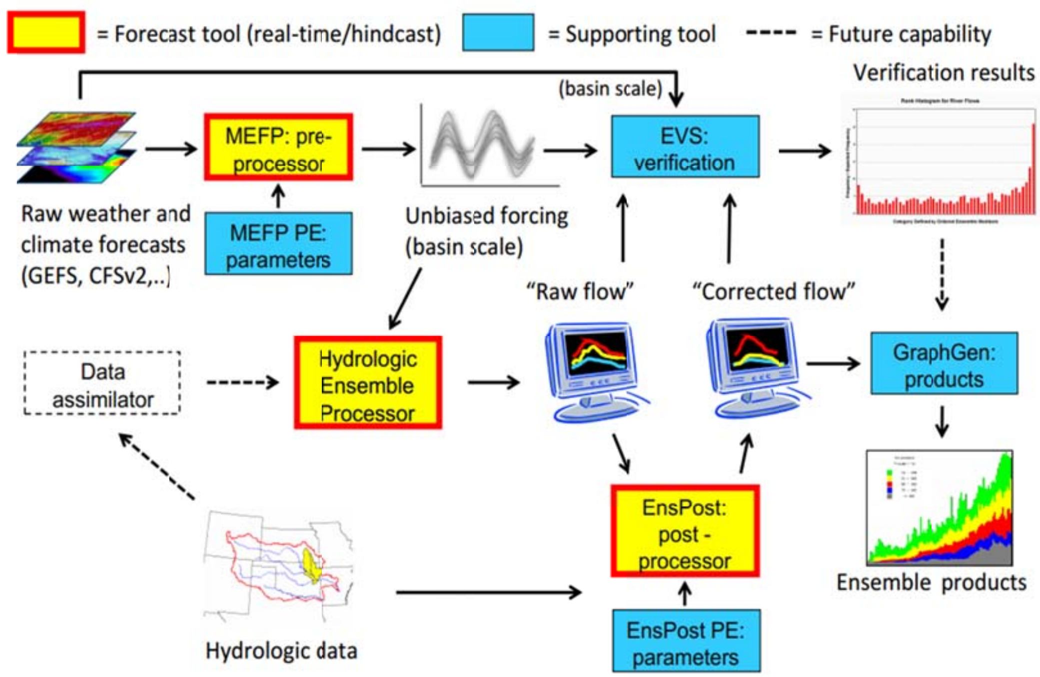


Figure 1.1 Schematic of the Hydrologic Ensemble Forecast Service (Demargne et al., 2014)

The HEFS models meteorological, or input, uncertainty in precipitation and temperature forecasts and hydrologic uncertainty in simulated streamflow separately and integrates the two numerically to estimate the predictive uncertainty in the streamflow forecast. The MEFPPE models input uncertainty to produce the MEFP parameters. The MEFP generates precipitation and temperature ensemble forecasts that are bias-corrected and account for input uncertainty (Krzysztofowicz, 1999; Schaake et al., 2007; Wu et al., 2011) based on single-valued quantitative precipitation and temperature forecasts (QPF, QTF). The HEFS also uses the ensemble post-processor, or EnsPost, to bias-correct and account for hydrologic uncertainty (Krzysztofowicz, 1999; Seo et al., 2010) in the raw ensemble streamflow forecast that reflects only the input uncertainty. Because the quality of ensembles that the MEFP and the EnsPost produce depend heavily on the goodness of the MEFP and the EnsPost parameters estimated by the MEFPPE and the EnsPostPE, it is very important that the MEFP and the EnsPost parameters are estimated carefully to maximize skill in the ensemble forecasts. Yet, this relationship between the forecast accuracy and the parameter quality has not been fully investigated.

For medium-range forecasting of precipitation and temperature, the HEFS currently uses the forecasts from the Global Ensemble Forecast System (GEFS; Demargne et al., 2014). While GEFS produces ensemble forecasts of precipitation and temperature along with many other variables (Hamill at el., 2013), such “raw” ensemble forecasts are generally biased in the mean and in higher-order moments (Wu et al., 2011). As such, it is generally necessary to remove or reduce biases by statistical means once the real-time forecasts become available. Also, it is well known that the ensemble spread in these raw forecasts is not capable of capturing flow-dependent predictability (Wu et al., 2011). Also, while generally skillful in capturing central tendencies, the ensemble mean of raw forecast tends to be biased unconditionally and/or conditionally (Hamill at el., 2013). For these reasons, significant efforts have been made in recent years (Gneiting at el., 2007; Hamill et al., 2008; Hamill at el, 2013) to bias-correct GEFS ensemble mean forecasts and to model the uncertainties associated with the resulting single-valued forecasts statistically. For reliable statistical bias correction and uncertainty modeling, however, a large amount of historical forecasts and verifying observations are necessary. For this purpose, the GEFS reforecast dataset has been developed by

National Centers for Environmental Prediction (NCEP) which provides a consistent NWP model output with a long period of record, from which statistical relationships for bias correction and uncertainty modeling may be derived (Schaake et al., 2007; Wu et al., 2010). The MEFP and the MEFPPE in the HEFS have been designed and developed with the above considerations in mind.

This study assesses the value of utilizing the HEFS by assessing short-range ensemble forecasts generated by the HEFS MEFP using a single-valued WGRFC QPF. The GEFS reforecast dataset is used to assess the value of medium-range ensemble precipitation forecasts to ensemble streamflow forecasting, and to evaluate sensitivity of the MEFP and the EnsPost parameters to the quality of ensemble precipitation and streamflow hindcasts.

The above objectives require careful and rigorous forecast verification. According to MetEd (https://www.meted.ucar.edu/), there are multiple reasons to verify forecasts:

1) Monitor forecast quality by measuring the agreement between forecasts and verifying observations,
2) Improve forecast quality by learning the forecast system, and
3) Compare one forecast system with another.

In this study, the Ensemble Verification System (EVS; Brown et al., 2010) developed by NWS/OHD is used to forecast verification.

The main goal of this study is to increase lead time of streamflow forecasts by using medium range ensemble precipitation forecasts. A premise for this study is that, in the ensemble paradigm, forecasting of precipitation and streamflow provides extended forecast lead time with improved forecast skill. To utilize forecast skill in medium range precipitation forecasts in the ensemble paradigm, the HEFS is used in this study.

The specific objectives of this study are as follows:

1) Evaluate the value of utilizing the HEFS for generating ensemble precipitation forecasts by assessing the predictive skill of the WGRFC QPF-forced short-range ensemble precipitation forecasts generated by the MEFP,
2) Assess the predictive skill of GEFS-forced medium-range ensemble precipitation forecasts generated by the MEFP,
3) Advance understanding of the sensitivity of the MEFP and the EnsPost parameters to the quality of ensemble precipitation and streamflow hindcasts,
4) Evaluate the impact, in terms of extending lead time and improving accuracy, of utilizing GEFS-forced medium-range ensemble precipitation forecast in Objective 1 on the raw ensemble streamflow forecast from HEP, and
5) Assess the value of EnsPost in post-processing the raw streamflow ensemble forecast in Objective 3.

This research makes the following new contributions:

1) It is the first time to utilize HEFS in CHPS outside of NWS.
2) It is the first time to evaluate the value of utilizing the HEFS as an ensemble generator outside of NWS
3) It is the first time to demonstrate possible benefits for decision makers in the water resource field by utilizing GEFS-forced medium-range ensemble forecasts of precipitation and streamflow

The organization of this thesis is as follows. Chapter 1 describes the objectives of the study. Chapter 2 describes the tools used. Chapter 3 describes the study area and the methodology used. Chapter 4 presents the results. Chapter 5 summarizes the conclusions and future research recommendations.

Chapter 2

Tools used

This study uses the HEFS for ensemble generation and verification. This chapter describes the HEFS and its components as they pertain to this research.

2.1 Hydrologic Ensemble Forecast Service (HEFS)

The HEFS was developed to address a wide range of needs for risk-based decision making in operational hydrology and water resources management (Hartman et al., 2007). The HEFS operates on the Community Hydrologic Prediction System (CHPS) which is a hydrologic forecasting shell that provides an open environment for collaborative development and data sharing among diverse stakeholders and users (Demargne et al., 2009; Demargne et al., 2014). The HEFS ensemble generation package consists of statistical models to model and quantify input and hydrologic uncertainties and to generate ensembles via conditional simulation (Deutsch and Journel, 1992). Hydrologic models, such as the Sacramento Soil Moisture Accounting (SAC-SMA) model (Burnash et al., 1995) and unit hydrograph (UHG), are available to the HEFS as CHPS operations. The statistical models in the HEFS require historical forecasts or simulations as well as the verifying observations to model relationships and associated uncertainties. Implicit in HEFS's statistical modeling is the stationarity assumption; the relationships do not change over time and hence may be applied to future events (Brown et al., 2014b; NWS OHD, 2015b). The HEFS includes two statistical models and their parameter estimation tools: the Meteorological Ensemble Forecast Processor (MEFP) for bias correction and accounting of input uncertainty and the Ensemble Post-processor (EnsPost) for bias correction of accounting of hydrologic uncertainty (Figure 1.1). The MEFP generates precipitation and temperature ensembles which provide ensemble input to hydrologic models to generate raw streamflow ensembles. EnsPost bias-corrects the raw streamflow ensembles to produce post-processed streamflow ensembles.

2.1.1 Meteorological Ensemble Forecast Processor (MEFP)

The MEFP generates ensemble forecasts of precipitation and temperature given the conditioning single-valued forecasts. The MEFP can use multiple sources of forcing forecast over different lead times (short, medium-range and long) to produce a single ensemble forecast (see Table 2.1).

Table 2.1 Forecast data used in the MEFPPE (NWS OHD, 2015b)

Forecasting range		Forecast data	Generator/Developer agency
Short range	Up to 5 days	Single-valued quantitative precipitation forecasts (QPF)	NWS River Forecast Centers (RFC)
		Single-valued quantitative precipitation forecasts	NWS Weather Prediction Center (WPC)
Medium range	Up to 15 days	Ensemble forecasts of the Global Ensemble Forecast System (GEFS)	National Centers for Environmental Prediction (NCEP)
Long range	Up to 9 months	Single-valued forecasts of the Climate Forecast System version 2 (CFSv2)	National Centers for Environmental Prediction
	Up to 1 year	Climatology ensembles	Re-sampled by MEFP

The MEFP models the bivariate distribution of the forecast and verifying observation. The bivariate distribution model is used to generate ensemble members given the single-valued forcing forecast. The final ensemble forecast is generated based on the above conditional simulation and Schaake shuffle (Wu et al., 2011; NWS OHD, 2015b) which “shuffles” the ensemble members generated specifically for each time step into the naturally occurring patterns of temporal variability (Clark et al., 2004). Figure 2.1shows the schematic of the MEFP methodology.

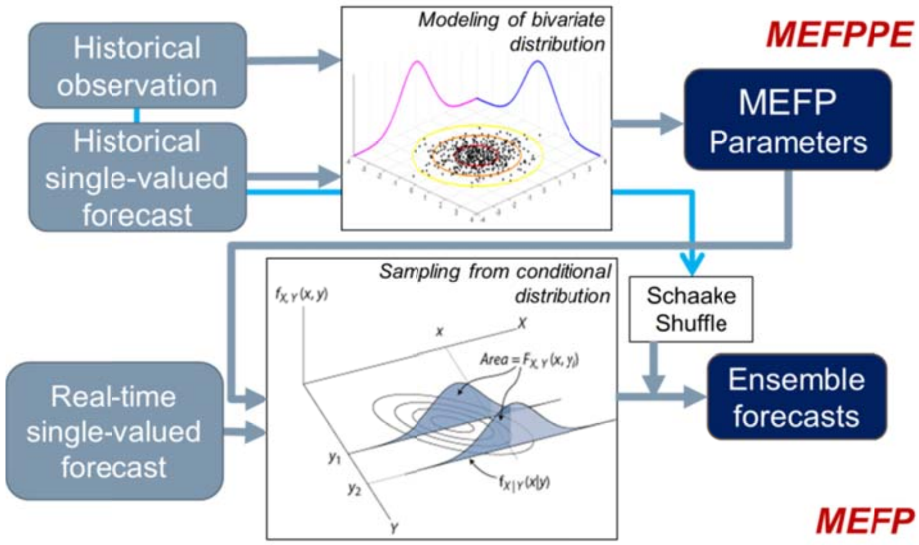


Figure 2.1 Schematic of the MEFP methodology

The MEFP Parameter Estimator, MEFPPE, estimates the MEFP parameters. To reduce sampling uncertainty, the MEFP pools the pairs of forecast and verifying observation over the user-defined time window centered on each Julian day, referred as the sampling window. Sampling uncertainty decreases as the sampling window increases but at the expense of reduced temporal specificity of the parameters in capturing seasonality. Another factor affecting the MEFP parameter estimation is the canonical event set. Canonical events are the time scales for bivariate modeling that vary over the forecast horizon (Brown et al., 2014). A canonical event set may include base and modulation events. The base events are non-overlapping events that fill the entire forecast horizon. The modulation events contain some base events, can overlap, and may cover part of or the entire forecast horizon (NWS OHD, 2015b). The combination of base and modulation events defines temporal windows within which a statistical model is formed from the forecast-observation pairs available. According to Brown (personal communication, 2015), the statistical models are applied in sequence according to the lowest-through-highest correlation/regression coefficients in normal space. For example, when the overall model comprises a sequence of 6-hour base events and 3-day modulation events, each 6-hour period within the forecast horizon is adjusted by the statistical model formulated based on samples pooled from the corresponding 6-hour base events. Then, the statistical model from the 3-day modulation events adjusts each 3-day period.

2.1.2 Ensemble Post-Processor (EnsPost)

The EnsPost removes or reduces biases in streamflow simulated by a suite of hydrologic models called Hydrologic Ensemble Processor (HEP) in the HEFS and models hydrologic uncertainty. Figure 2.2 illustrates how the input and hydrologic uncertainties are numerically integrated under the assumption that each ensemble member is equally likely to yield an estimate of predictive uncertainty in ensemble streamflow forecast. The EnsPostPE estimates the EnsPost parameters. The EnsPost uses the autoregressive-1 model with a single exogenous variable, or ARX (1, 1), in the bivariate normal space (Seo et al. 2006). Because streamflow is non-normal, the simulated and observed streamflow are normal quantile-transformed (NQT) (Bogner et al., 2012) before ARX (1, 1) modeling. Once ARX (1, 1) is applied, the predictand is back-transformed into the original space. Figure 2.3 illustrates the EnsPost methodology. For details, the reader is referred to Seo et al. (2006).

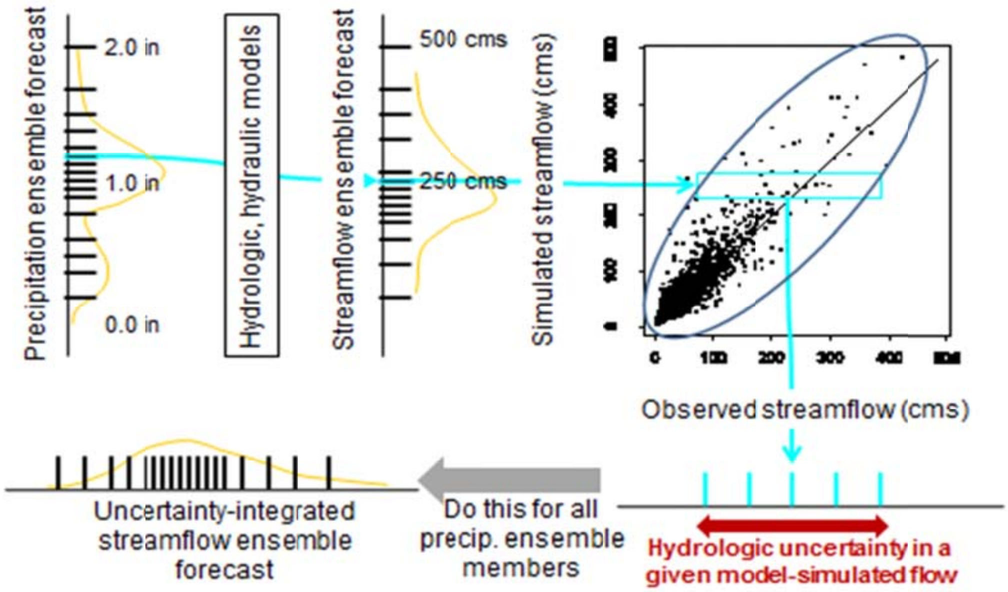


Figure 2.2 Input and hydrologic uncertainties in hydrologic ensemble forecasts (NWS OHD, 2015a)

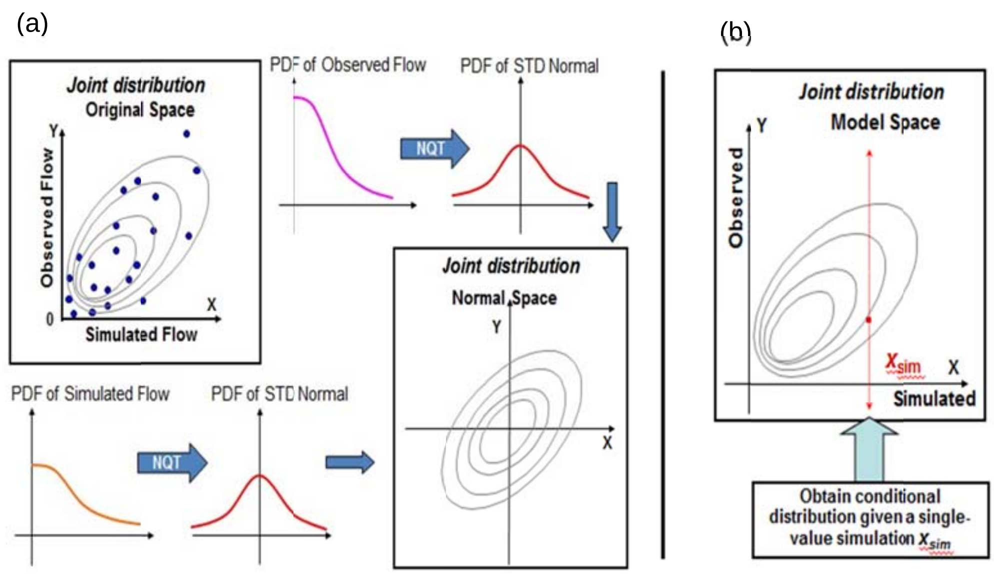


Figure 2.3 Schematic of EnsPost: (a) NQT, and (b) conditional probability distribution (NWS OHD, 2015a)

2.2 Ensemble Verification System (EVS)

The EVS (Brown et al., 2010) is a Java-based tool for verification of forcing and hydrologic ensembles. The EVS includes a comprehensive set of metrics for verification of both single-valued and ensemble forecasts. The EVS can evaluate probabilistic forecast attributes such as reliability, resolution, sharpness and discrimination. In this study, the ensemble mean results are first verified using correlation, which measures the strength of linear association between the ensemble mean forecast and the verifying observation. To verify ensemble forecasts, this study uses box plots, Continuous Ranked Probability Skill Score (CRPSS) and Relative Operating Characteristic (ROC) Score to examine distribution of forecast errors visually, lumped errors, and discrimination, respectively. Below the EVS metrics used in this study are briefly described. For further details, the reader is referred to Brown et al. (2014a&b).

The Continuous Ranked Probability Score (CRPS) represents the integral squared difference between the cumulative distribution functions (CDF) of the predicted variable, $F_Y(q)$, and the corresponding CDF of the observed variable, $F_X(q)$,

$$CRPS = \int \{(F_Y(q) - F_X(q)\}^2 \, dy \quad (2\text{-}1)$$

$$\overline{CRPS} = \frac{1}{n} \sum CRPS \quad (2\text{-}2)$$

where n denotes the number of pairs of forecasts and observations. The Mean CRPSS measures the performance of a forecast system relative to climatology. Because a perfect forecast has a CRPS of zero, the perfect value for CRPSS is unity:

$$\overline{CRPSS} = \frac{\overline{CRPS}_{clim.} - \overline{CRPS}}{\overline{CRPS}_{clim.}} \quad (2\text{-}3)$$

The ROC Score estimates the ability of a forecast system to predict that an event will occur (Probability of Detection or PoD) while avoiding predicting an incorrect event that does not occur (False Alarm Rate or FAR). The area enclosed by the diagonal line and the ROC curve generated by (FAR, PoD) is known as the ROC area. The ROC Score is obtained by multiplying the ROC area by two. The ROC area for a perfect forecast system is 0.5. The ROC score of a perfect system is therefore 1.

The box plot is a widely used tool for visual inspection of data distribution. In this study, the box plot is used to plot box-and-whisker at various quantiles of the forecast error as a function of ascending order of observed values. The plot readily shows the conditional bias of the ensemble mean forecasts, if any.

Chapter 3

Methodology

3.1 Study Area

The area selected for this study is the five headwater basins located upstream of the Dallas-Fort Worth (DFW) Metroplex in the Upper Trinity River Basin. They are the Big Sandy Creek near Bridgeport (BRPT2), Denton Creek near Justin (DCJT2), Elm Fork of the Trinity River near Gainesville (GLLT2), West Fork of the Trinity River near Jacksboro (JAKT2), and Clear Creek near Sanger (SGET2). Figure 3.1 and

Table 3.1 show the locations and physiographic and fluvial characteristics of the basins, respectively. According to the Texas Water Development Board (TWDB), more than 95% of the water used in the Upper Trinity River Basin is surface water and yet Dallas has the highest per-capita water use in Texas. DFW is the largest inland population center and one of the fastest growing urban areas in the U.S. and hence is under a great stress for water resources. The DFW region is very vulnerable to the impacts of urbanization on water sustainability due to the warmer climate conditions, rapid land conversion, high degree of impervious surfaces, and dependence on surface water. Therefore, skillful precipitation and streamflow forecasts for these basins are of great value to flood warning, water supply, reservoir operations, and water quality management.

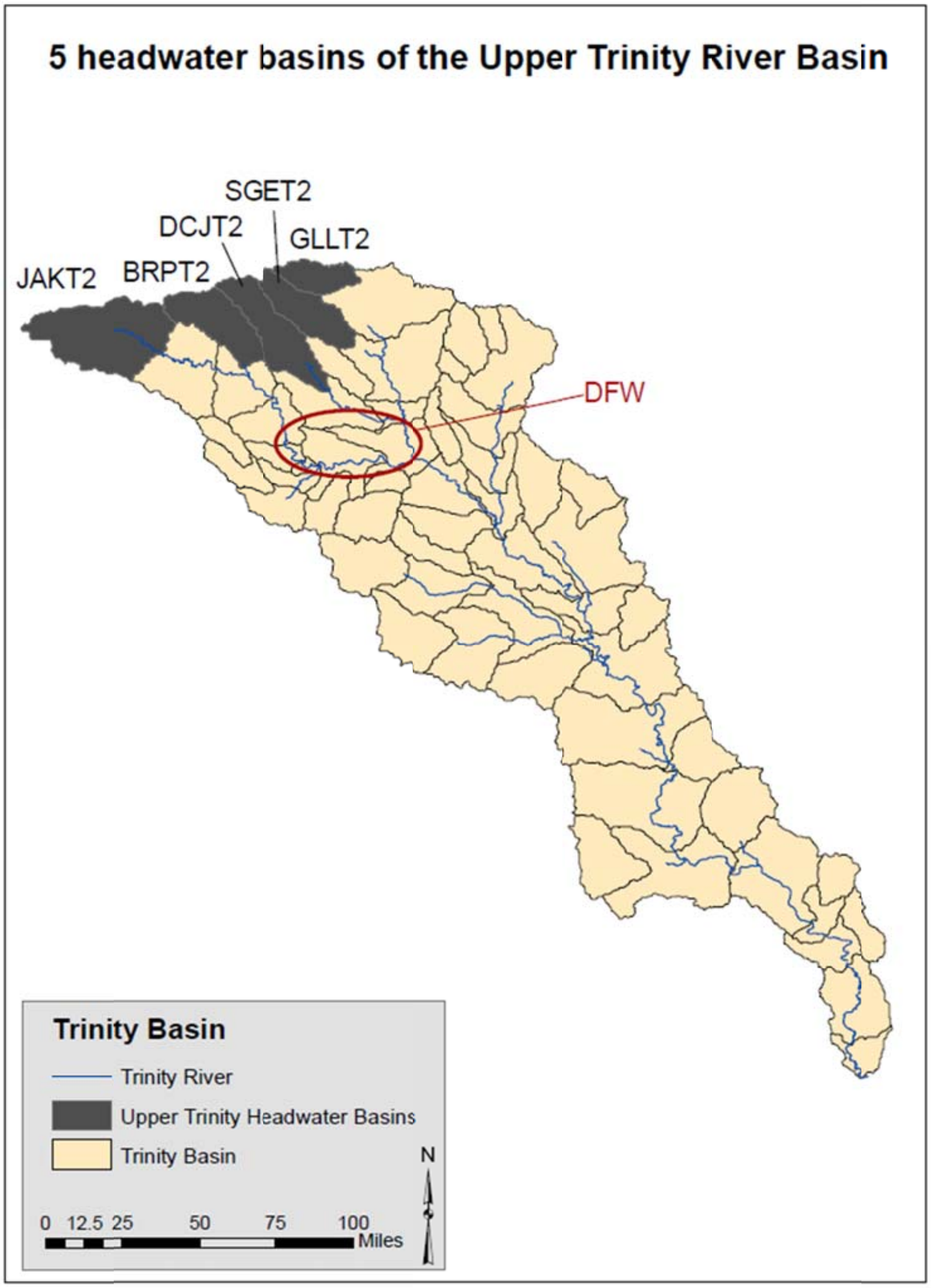


Figure 3.1 The five headwater basins of the Upper Trinity River Basin

Table 3.1 Characteristics of the five headwater basins in the Upper Trinity River Basin (Saharia, 2013)

Characteristics	BRPT2	DCJT2	GLLT2	JAKT2	SGET2
Latitude	33.23	33.12	33.62	33.29	33.34
Longitude	-97.69	-97.29	-97.15	-98.08	-97.18
Area (km^2)	862.47	1036.00	450.66	1769.00	764.05
Mean Elev. (m)	229	197	227	279	193
Action Stage (m)	3.35	2.74	6.10	5.49	7.01
Flood Stage (m)	3.65	3.05	6.71	6.10	7.62
Time to Peak (hours)	24	12	12	24	12

3.2 Data used in this study

Several data sets were used to generate precipitation and streamflow hindcasts for the five headwater basins (Table 3.2). As showed in Figure 2.1, the historical mean areal precipitation (MAP) time series and the GEFS reforecast dataset are used as input to the MEFPPE to estimate the MEFP parameters, and to generate ensemble precipitation forecasts from the MEFP. The observed mean daily flow (QME) and the simulated mean 6-hr flow (SQIN) are used for the EnsPostPE to estimate the EnsPost parameters (Figure 2.3).

Table 3.2 Data sets used in this study

Name	Period of record	Description
Quantitative precipitation forecasts (QPF)	Mar 2004 to Oct 2014	6-hourly Single-valued forecasts Provided by WGRFC
Mean Areal Precipitation (MAP)	Oct1959 to Apr 2015	Observed 6-hour accumulated Provided by WGRFC
Mean daily streamflow (QME)	Oct1959 to Apr 2015	Observed mean daily produced by USGS Provided by WGRFC
GEFS	Jan1985 to Jul 2012	Mean ensemble precipitation forecasts Provided by NWS
SQIN	Oct1959 to Apr 2015	Simulated mean 6-hour streamflow Provided by WGRFC

The GEFS forecasts are 6-hourly precipitation amounts generated at 0Z for forecast horizons of 1 to 16 days (Hamill at el, 2013). Because an ensemble mean of a GEFS reforecast is estimated to be valid at 12Z, the GEFS precipitation reforecasts are available up to 15 days into the future for hindcasting experiments.

3.3 Hindcasting experiments

To address the research questions posed in Chapter 1, hindcasting experiments using the HEFS were designed and carried out (see Figure 3.2). Using ensemble mean of the GEFS reforecast and the verifying observed MAP, the MEFP parameters were estimated by the MEFPPE, which were then used to generate ensemble precipitation hindcasts conditional on the GEFS ensemble mean hindcast via the MEFP. Using the MEFP-generated precipitation ensemble forecasts, raw ensemble streamflow hindcasts are then generated using hydrologic models built in CHPS.

The EnsPostPE uses simulated streamflow (SQIN) from the hydrologic models and the verifying observed streamflow (QME) to estimate the EnsPost parameters. The EnsPost is then used to produce post-processed streamflow hindcasts from the raw streamflow hindcasts using the parameters estimated by the EnsPostPE. The hindcasts are generated every day for 26 years (1986-2011). The resulting large-sample ensemble precipitation, raw, and post-processed streamflow hindcasts are verified using the EVS.

For the sensitivity analysis of the MEFP and the EnsPost parameters to the quality of ensemble precipitation and streamflow hindcasts, the ensemble hindcasts generated by the MEFP and the EnsPost are verified using various the MEFPPE- and the EnsPostPE-estimated parameters. In the HEFS, parameter estimation is controlled by user-defined environmental variables such as the sampling window. In this study, five different combinations of user-defined environmental variables were compared (see Table 3.3). The user-defined environmental variables for MEFPPE include the canonical events and sampling window. Those for EnsPostPE include the sampling period (e.g., monthly, semi-annual, or annual).

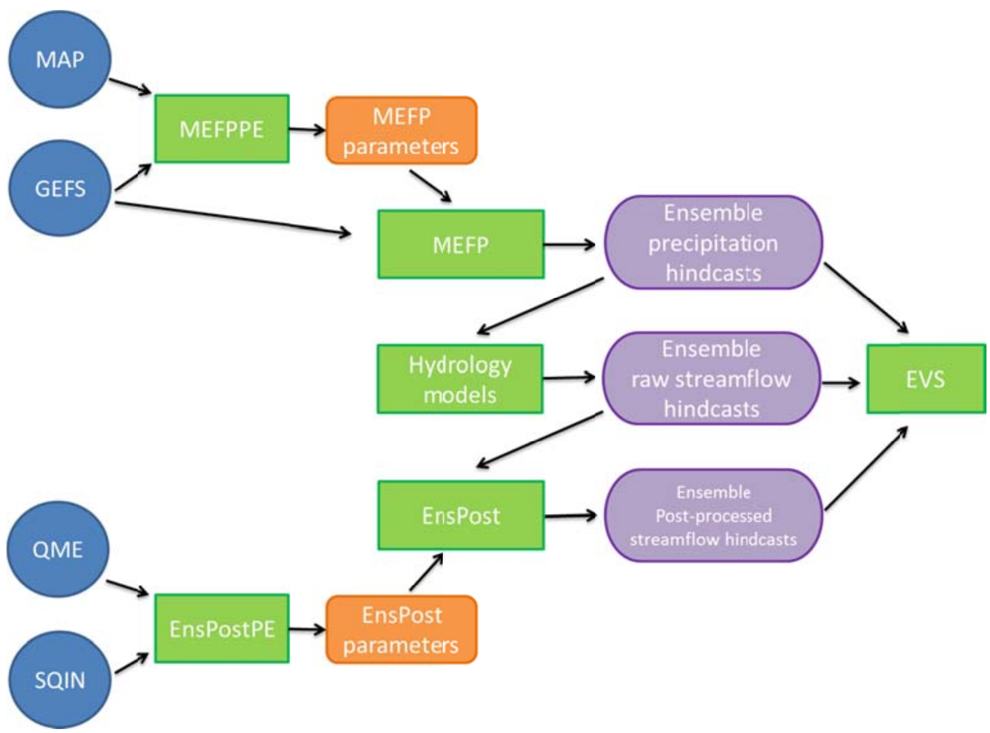


Figure 3.2 Schematic of the hindcasting processes

Table 3.3 Different combinations of environmental variables

<table>
<tr><th rowspan="2">Environmental variable</th><th colspan="2">MEFPPE</th><th>EnsPostPE</th></tr>
<tr><th>Canonical events</th><th>Sampling window (days)</th><th>Sampling period (months)</th></tr>
<tr><td>Case 1</td><td rowspan="3">CE1</td><td rowspan="2">61
(Default value recommended by NWS)</td><td>monthly</td></tr>
<tr><td>Case 2</td><td>semi-annual
wet period: Mar, Apr, May, Jun, Sep, and Oct
dry period: Jan, Feb, Jul, Aug, Nov, and Dec</td></tr>
<tr><td>Case 3</td><td rowspan="3">91[1]</td><td rowspan="3">monthly</td></tr>
<tr><td>Case 4</td><td>CE2</td></tr>
<tr><td>Case 5</td><td>CE3</td></tr>
</table>

Note 1: The choice of 91 days is based on the practice at the River Forecast Centers (RFC) when the default window does not produce large enough sample size.

Four hindcasting experiments (see Table 3.5) were designed and carried out using the five different combinations of the MEFPPE and the EnsPostPE environmental variables (see Table 3.3). Depending on the combination of the environmental variables, the MEFPPE and the EnsPostPE generate different sets of parameters for the MEFP and the EnsPost for use in producing ensemble precipitation and post-processed forecasts, respectively. For all experiments, the study period was 26 years (1986 – 2011). The number of ensemble members generated was 55 corresponding to the number of historical years available for Shaake Shuffle. Each experiment evaluates the effect of different environmental variable-controlled parameters on predictive skill in ensemble forecasts.

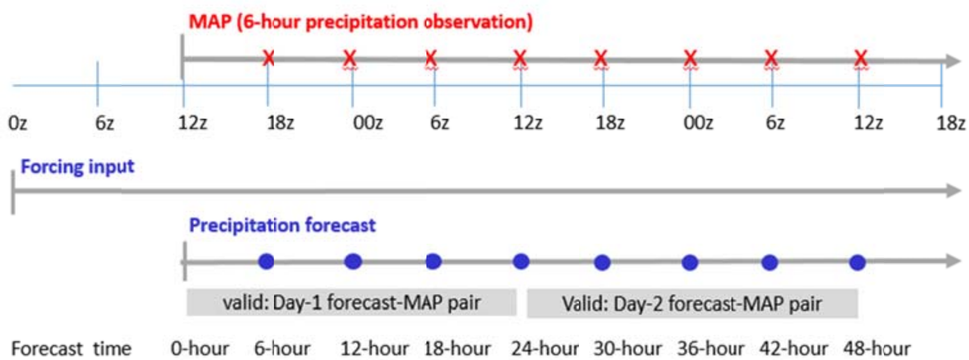
Table 3.4 Hindcasting experiments

Experiment	Comparison (from Table 3.3)	Parameter Estimator in HEFS	Assessment
1	Case 1 vs Case 2	EnsPostPE	aggregation time
2	Case 1 vs Case 3	MEFPPE	sampling windows
3	Case 3 vs Case 5	MEFPPE	canonical events (base)
4	Case 4 vs Case 5	MEFPPE	canonical events (modulation)

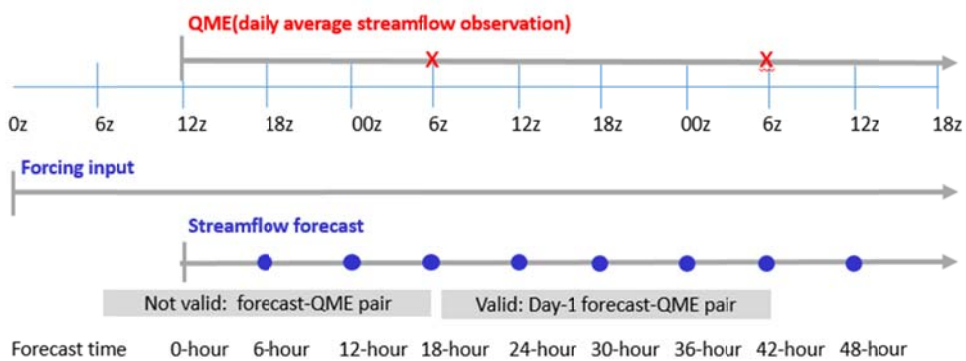
Table 3.5 Different combinations of canonical events

<table>
<tr><th rowspan="2">lead time (hour)</th><th>CE1</th><th colspan="6">CE2</th></tr>
<tr><th>base</th><th>base (CE3)</th><th colspan="5">modulation</th></tr>
<tr><td>6</td><td>1</td><td>1</td><td rowspan="4">1</td><td rowspan="12">4</td><td rowspan="20">7</td><td rowspan="28">10</td><td rowspan="60">21</td></tr>
<tr><td>12</td><td>2</td><td>2</td></tr>
<tr><td>18</td><td rowspan="2">3</td><td>3</td></tr>
<tr><td>24</td><td>4</td></tr>
<tr><td>30</td><td rowspan="2">4</td><td>5</td><td rowspan="4">2</td></tr>
<tr><td>36</td><td>6</td></tr>
<tr><td>42</td><td rowspan="2">5</td><td>7</td></tr>
<tr><td>48</td><td>8</td></tr>
<tr><td>54</td><td rowspan="2">6</td><td>9</td><td rowspan="4">3</td></tr>
<tr><td>60</td><td>10</td></tr>
<tr><td>66</td><td rowspan="2">7</td><td>11</td></tr>
<tr><td>72</td><td>12</td></tr>
<tr><td>78</td><td rowspan="2">8</td><td>13</td><td rowspan="4">5</td><td rowspan="12">9</td></tr>
<tr><td>84</td><td>14</td></tr>
<tr><td>90</td><td rowspan="2">9</td><td>15</td></tr>
<tr><td>96</td><td>16</td></tr>
<tr><td>102</td><td rowspan="2">10</td><td>17</td><td rowspan="4">6</td></tr>
<tr><td>108</td><td>18</td></tr>
<tr><td>114</td><td rowspan="2">11</td><td>19</td></tr>
<tr><td>120</td><td>20</td></tr>
<tr><td>126</td><td rowspan="4">12</td><td rowspan="4">21</td><td rowspan="8">8</td><td rowspan="20">14</td></tr>
<tr><td>132</td></tr>
<tr><td>138</td></tr>
<tr><td>144</td></tr>
<tr><td>150</td><td rowspan="4">13</td><td rowspan="4">22</td><td rowspan="12">12</td></tr>
<tr><td>156</td></tr>
<tr><td>162</td></tr>
<tr><td>168</td></tr>
<tr><td>174</td><td rowspan="4">14</td><td rowspan="4">23</td><td rowspan="8">11</td><td rowspan="28">20</td></tr>
<tr><td>180</td></tr>
<tr><td>186</td></tr>
<tr><td>192</td></tr>
<tr><td>198</td><td rowspan="8">15</td><td rowspan="4">24</td></tr>
<tr><td>204</td></tr>
<tr><td>210</td></tr>
<tr><td>216</td></tr>
<tr><td>222</td><td rowspan="4">25</td><td rowspan="8">13</td><td rowspan="12">16</td></tr>
<tr><td>228</td></tr>
<tr><td>234</td></tr>
<tr><td>240</td></tr>
<tr><td>246</td><td rowspan="8">16</td><td rowspan="4">26</td><td rowspan="20">19</td></tr>
<tr><td>252</td></tr>
<tr><td>258</td></tr>
<tr><td>264</td></tr>
<tr><td>270</td><td rowspan="4">27</td><td rowspan="8">15</td></tr>
<tr><td>276</td></tr>
<tr><td>282</td></tr>
<tr><td>288</td></tr>
<tr><td>294</td><td rowspan="12">17</td><td rowspan="4">28</td><td rowspan="12">18</td></tr>
<tr><td>300</td></tr>
<tr><td>306</td></tr>
<tr><td>312</td></tr>
<tr><td>318</td><td rowspan="4">29</td><td rowspan="8">17</td></tr>
<tr><td>324</td></tr>
<tr><td>330</td></tr>
<tr><td>336</td></tr>
<tr><td>342</td><td rowspan="4">30</td><td rowspan="4"></td></tr>
<tr><td>348</td></tr>
<tr><td>354</td></tr>
<tr><td>360</td></tr>
</table>

The ensemble precipitation and streamflow hindcasts produced from the above experiments are verified at the daily scale via EVS. The EVS pairs streamflow forecasts and verifying daily observations with consideration of local timing of streamflow observation (Brown, 2014). For precipitation forecasts, the pairing process is rather straightforward. As illustrated in Figure 3.3, daily accumulation for the first pair of precipitation forecast and MAP occurs between 0 and 24 hours of forecast lead time (Figure 3.3 (a)) whereas the first daily pair of streamflow forecast and QME occurs between 18 and 42 hours of forecast lead time (Figure 3.3 (b)).



(a)



(b)

Figure 3.3 Daily aggregation for forecasts evaluation in the EVS

Chapter 4

Results

The ensemble precipitation and streamflow hindcasts generated for the 26-year period (1986 - 2011) were verified against MAP and QME, respectively. The results are presented as a function of forecast lead time with respect to different thresholds chosen based on observed precipitation and streamflow for the entire hindcast period of 26 years. To evaluate the value of utilizing the HEFS for generating ensemble precipitation forecasts, this study assessed forecast skill in short-range ensemble precipitation hindcasts generated using WGRFC QPF during 2004-2011 for JAKT2 via HEFS. For the assessment of medium range GEFS-forced precipitation hindcasts, the hindcasts were pooled over all five basins to increase the sample size. All the results from the four experiments are presented at a daily scale unless specified otherwise. Throughout this chapter, precipitation hindcasts are in millimeters, and streamflow hindcasts are in cubic meters per second.

4.1 WGRFC QPF-forced short-range ensemble precipitation hindcasts

In order to evaluate in a single-valued forecast sense, the correlation coefficient is used. Figure 4.1 shows correlation coefficients between the ensembles mean QPF and observed precipitation as a function of lead time for JAKT2. As expected, a significant correlation coefficient for Day-1 decreases slowly up to Day-3 and abruptly drops off to negligible levels after that, showing no skill. The plot indicates that ensemble hindcasts generated via HEFS allows forecast skill to last longer than the first 6 hours offered by a single-valued QPF for general use as mentioned in Chapter 1. Each curve represents a different threshold of precipitation value expressed as percentiles of observed precipitation in the 26-yr period (25th, 70th, 90th, 95th, and 99th percentiles). For higher thresholds, the sample size is rather small and sampling uncertainties exist in the verification statistics.

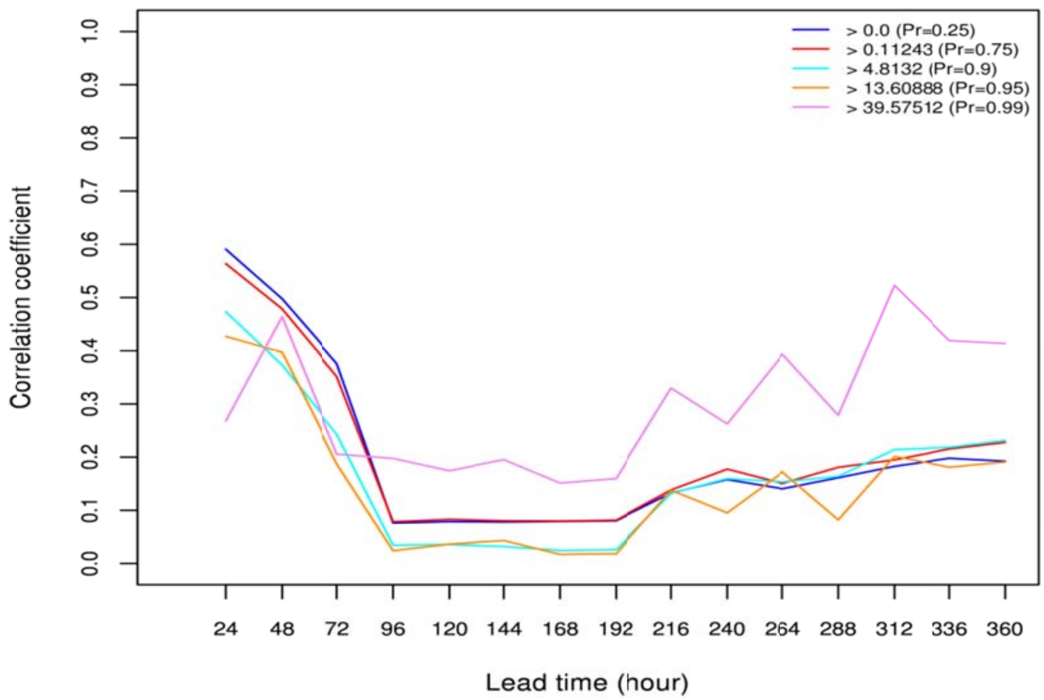


Figure 4.1 Correlation Coefficient of ensemble mean precipitation forecasts generated by MEFP using a single-valued WGRFC QPF and corresponding observations for JAKT2

Figure 4.2 and 4.3 show the CRPSS and ROC score of ensemble precipitation hindcasts as a function of forecast lead time, respectively. The pattern of change in forecast skill as a function of forecast lead time is the same as that shown in Figure 4.1 for correlation coefficient: CRPSS and ROC scores for Day-1 gradually declines as lead time increases up to Day-3, indicating that forecast skill in ensemble precipitation hindcasts forced by WGRFC QPF via MEFP exists for longer forecast lead time than a single-valued WGRFC QPF.

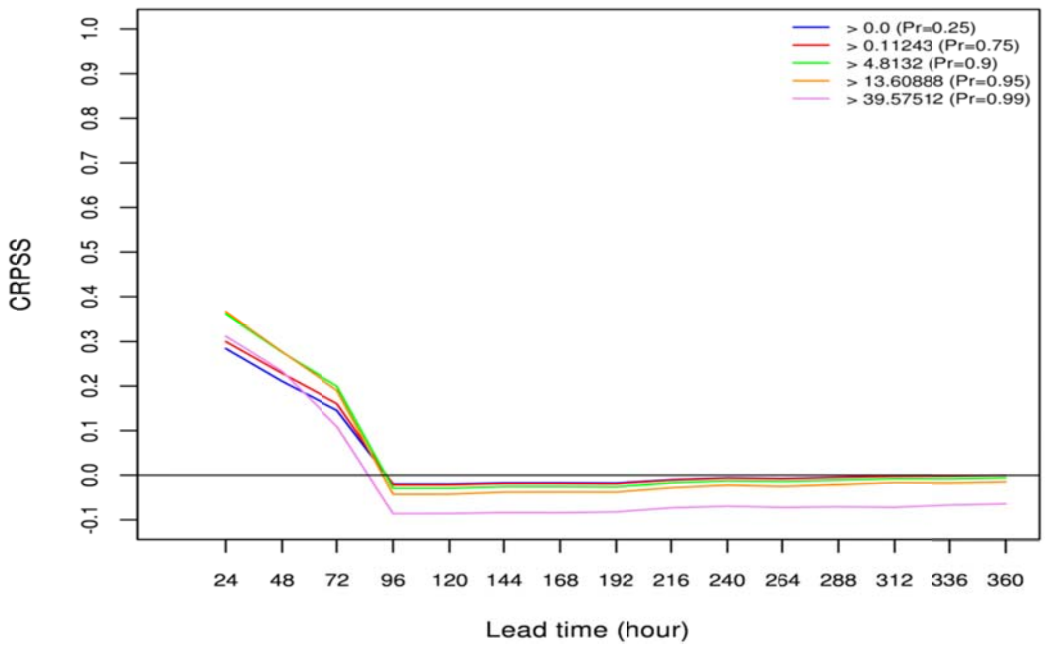


Figure 4.2 CRPSS of ensemble precipitation hindcasts generated by MEFP using a single-valued WGRFC QPF for JAKT2

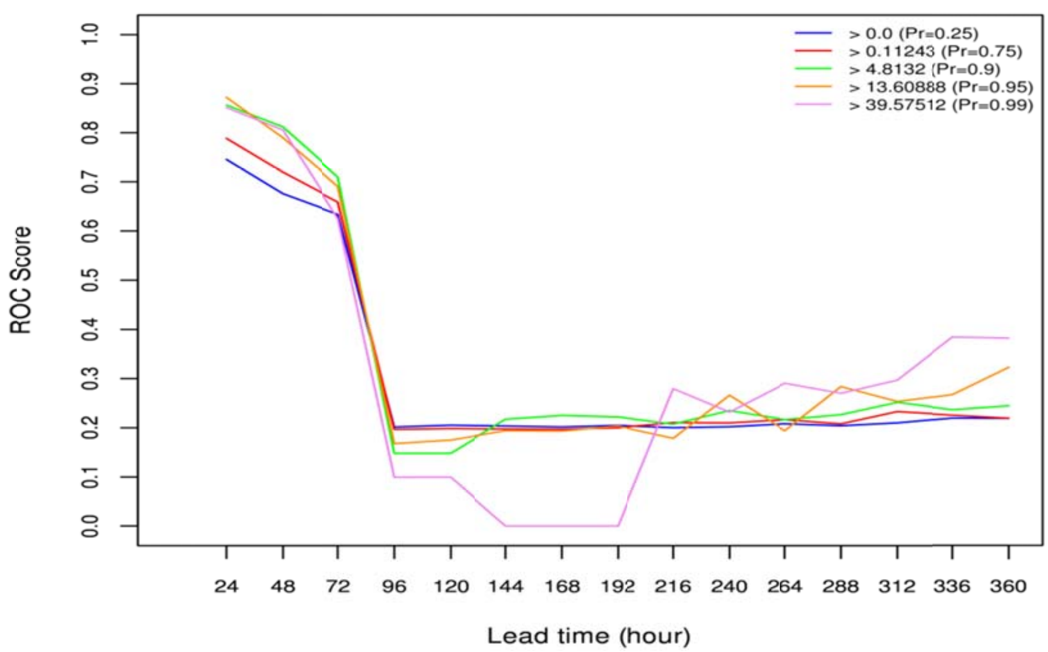


Figure 4.3.ROC score of ensemble precipitation hindcasts generated by MEFP using a single-valued WGRFC QPF for JAKT2

Due to lack of forecast skill in WGRFC QPF-forced ensemble precipitation hindcasts after Day-3, the assessment has not been carried out further. The value of utilizing HEFS is obviously seen when this study presents that forecast skill in ensemble QPF generated using the single-valued WGRFC QPF via HEFS shows for the extended forecast lead time (6 hour vs Day-3). To assess ensemble precipitation and streamflow hindcasts, the GEFS data were used.

4.2 Medium-range GEFS-forced ensemble precipitation hindcasts

4.2.1 Sample size

In this study, a sample refers to a pair between an ensemble mean hindcast and the verifying observation. Sampling uncertainty increases as the sample size decreases with increasing thresholds (Table 4.1). For the 97.5^{th} and 99^{th} percentile thresholds, the sample size is 238 and 95 for each basin, respectively. When the hindcasts are pooled over all 5 basins, the sample size increases to 1187 and 476 for the same thresholds. To reduce sampling uncertainty, verification was carried out by pooling hindcasts over all 5 basins. The EVS metrics for individual basins are presented in Appendix A.

Table 4.1 Sample size corresponding to various thresholds of observed precipitation at Day-1

Basin \ Threshold	All[1]	0 %	75%	90%	95%	97.5%	99%
BRPT2	9484	3013	2373	951	475	238	95
DCJT2	9484	3230	2374	949	475	238	95
GLLT2	9484	2600	2373	950	476	238	95
JAKT2	9484	3014	2384	950	475	238	95
SGET2	9484	3022	2395	950	475	238	95
All	47420	14879	11863	4747	2376	1187	476

Note 1: “All” includes “no-rain”

4.2.2 Correlation Coefficient

Figure 4.4 shows the correlation coefficient of ensemble mean of daily precipitation hindcasts and the verifying observations as a function of forecast lead time for different thresholds. The correlation for Day-1 is as high as 0.6 and decreases with increasing lead time and precipitation thresholds. Forecast skill exists for longer lead time, when compared to the skill in WGRFC QPF-forced ensemble precipitation hindcasts. Since no forecast skill shows after Day-8, the assessment throughout this section is up to Day-8. For high thresholds, correlations are very small to negligible, indicating little skill in predicting large precipitation amounts in the single-valued sense.

Figures 4.5, 4.6, 4.7, 4.8, and 4.9 show the correlation coefficients of ensemble mean precipitation hindcasts and verifying observations for BRPT2, DCJT2, GLLT2, JAKT2, and SGET2, respectively. The correlation coefficients for the individual basins are similar among themselves in that the correlation for Day 1 is approximately 0.6 and decreases as lead time increases. Again, the correlation coefficients are very low for high thresholds. The consistent pattern among all basins indicates that pooling hindcasts from all five basins for verification is reasonable.

Figure 4.10 shows the correlation coefficient between the mean of raw (upper panel) and post-processed (lower panel) streamflow ensembles forced by GEFS-based ensembles and the verifying observations as a function of lead time for different thresholds. In the upper panel, the correlation starts at approximately 0.8 and decreases as forecast lead time increases. The lower panel shows the same correlation pattern, indicating that no improvement in correlation occurs through post-processing. This is not surprising in that post-processing addresses biases, to which correlation is immune.

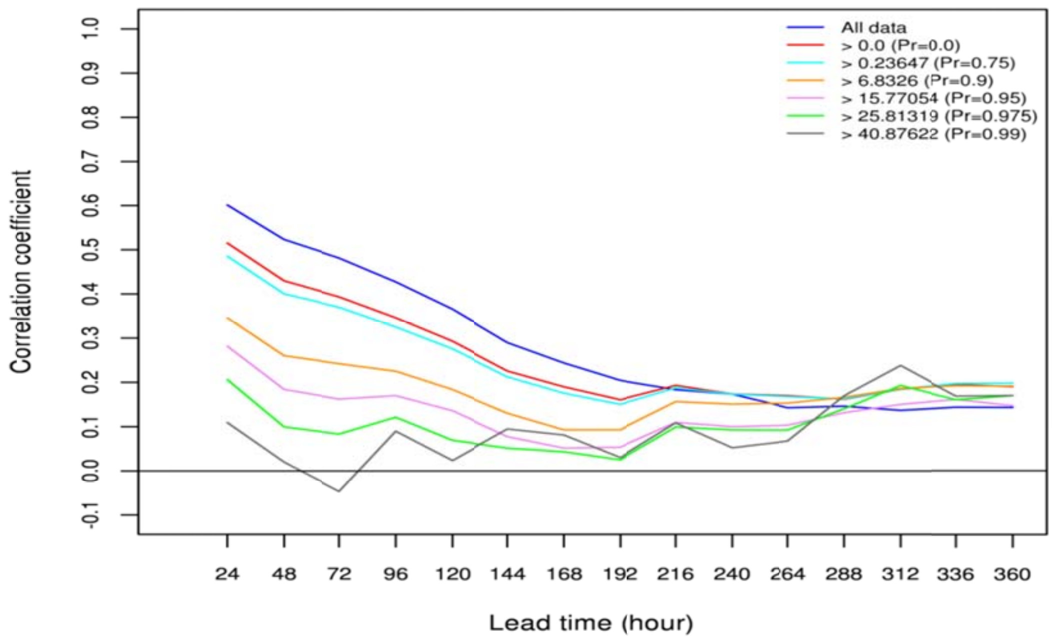


Figure 4.4 Correlation Coefficient of ensemble mean precipitation forecasts and corresponding observations for all basins

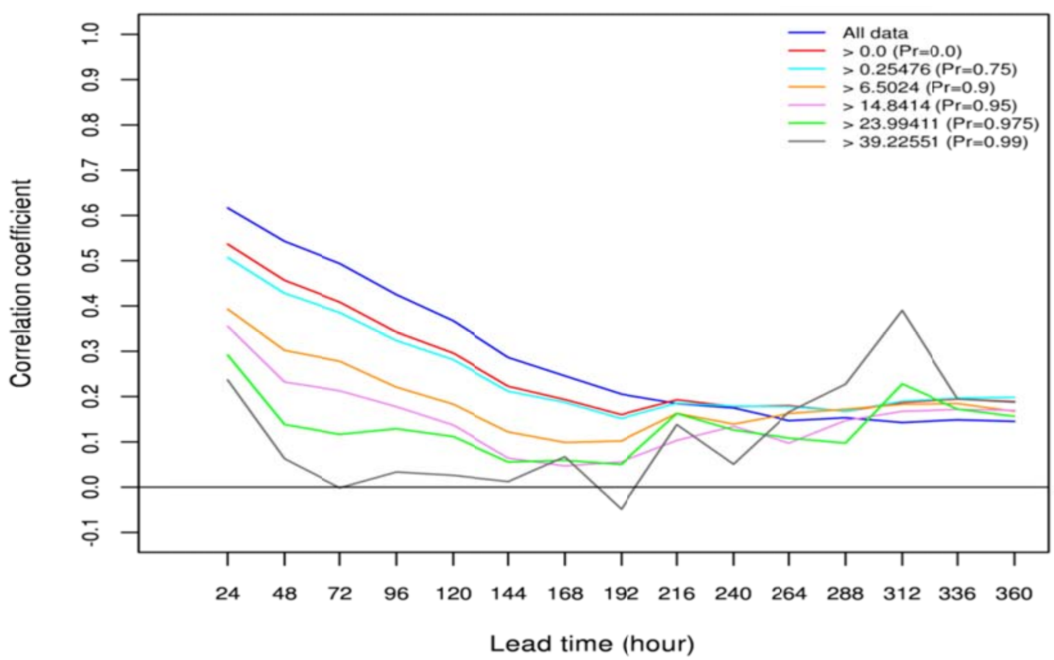


Figure 4.5 Correlation Coefficient of ensemble mean precipitation forecasts and corresponding observations for BRPT2

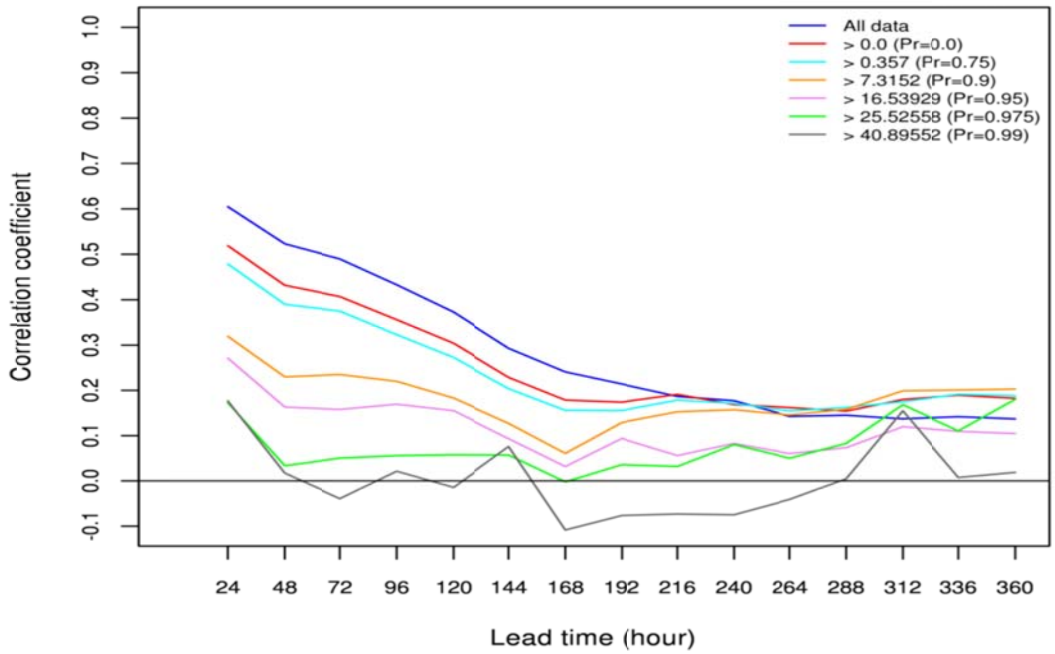


Figure 4.6 Correlation Coefficient of ensemble mean precipitation forecasts and corresponding observations for DCJT2

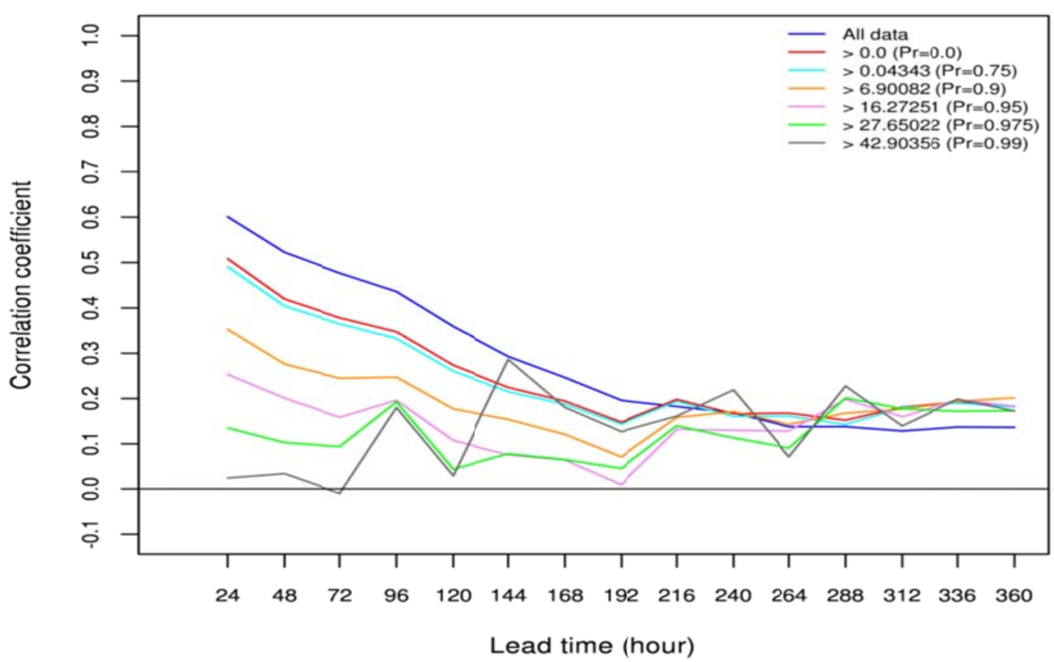


Figure 4.7 Correlation Coefficient of ensemble mean precipitation forecasts and corresponding observations for GLLT2

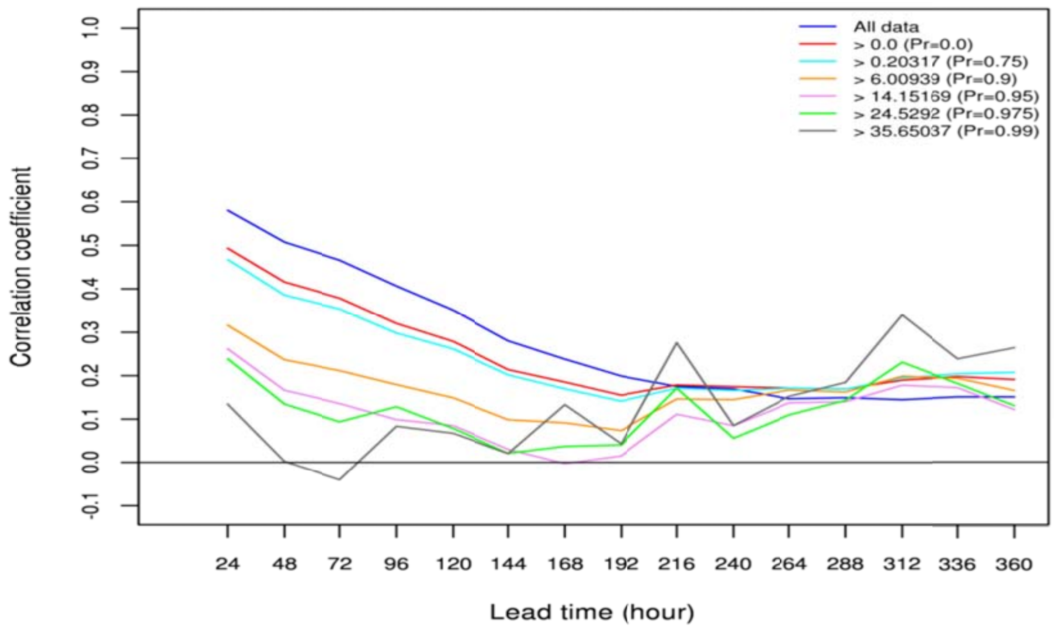


Figure 4.8 Correlation Coefficient of ensemble mean precipitation forecasts and corresponding observations for JAKT2

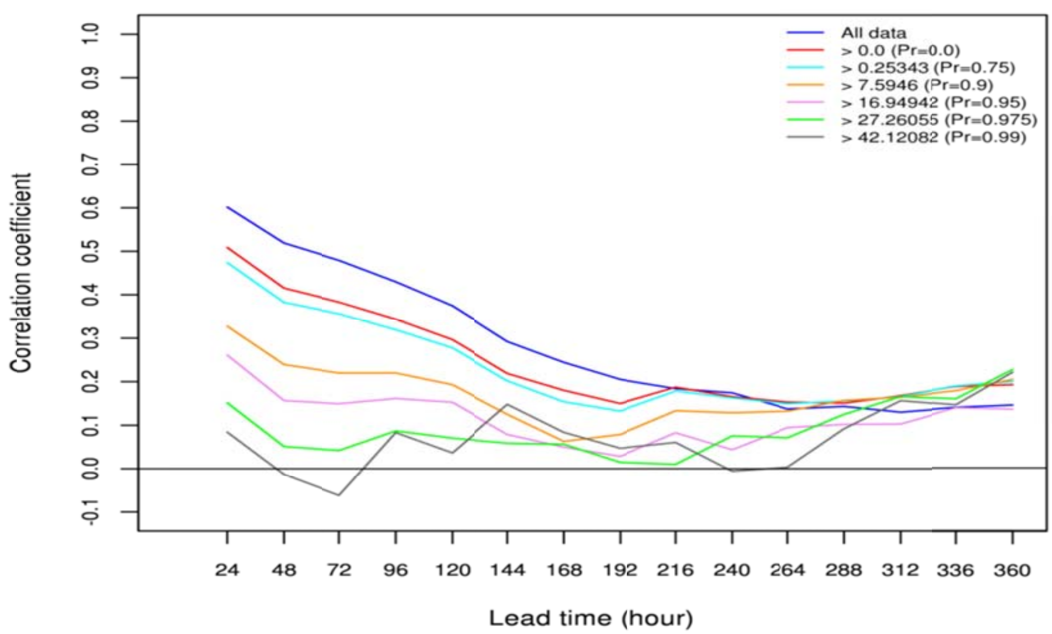


Figure 4.9 Correlation Coefficient of ensemble mean precipitation forecasts and corresponding observations for SGET2

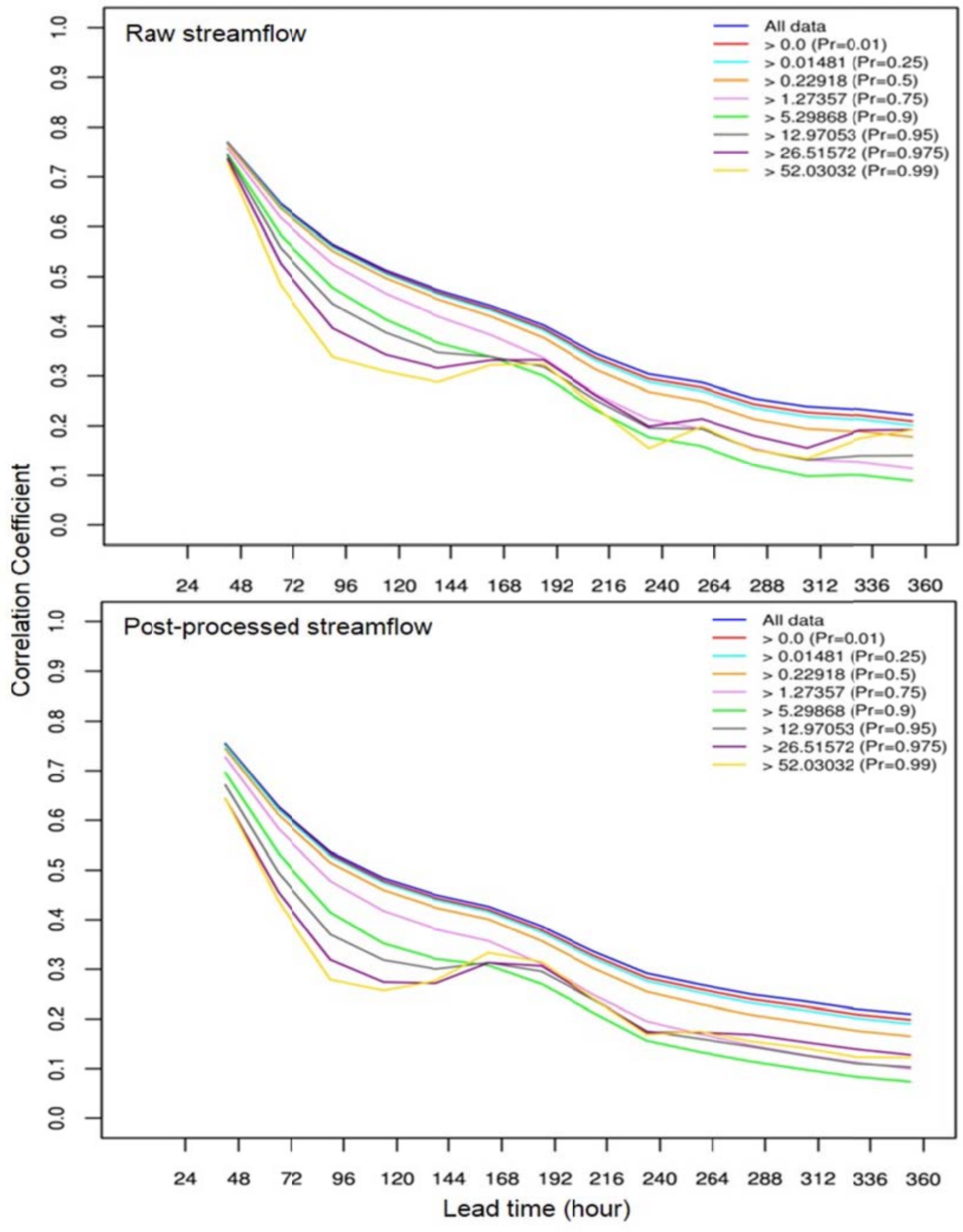


Figure 4.10 Correlation coefficient of raw (upper panel) and post-processed (lower panel) streamflow and verifying observations

4.2.3 Box plot

Boxplots of forecast errors against observed precipitations for up to Day-8 of forecast lead time are presented in Figure 4.11 and 4.12. GEFS-forced ensemble precipitation hindcasts are under-forecasting high precipitation events. Considering that the 99th observed precipitation threshold is less than 40 millimeters for this study area, GEFS-forced ensemble precipitation hindcasts has forecast skill up to Day-8. Figure 4.13 and 4.14 show the box plots of forecast errors of raw streamflow against observed values up to Hour-210. Post-processed ensemble streamflow hindcasts tend to under-forecast high streamflow events, but forecast skill exists up to 50 cfs, 99th percentile observed streamflow events in the study area up to about Day-8 (Figure 4.15 and 4.16). Post-processing raw streamflow hindcasts improves forecast skill for the entire forecast horizon. In general, it can be said that medium-range ensemble streamflow forecasts generated for the 5 headwater basins of the Upper Trinity River Basin has reasonable skill to predict precipitation and streamflow events up to a week.

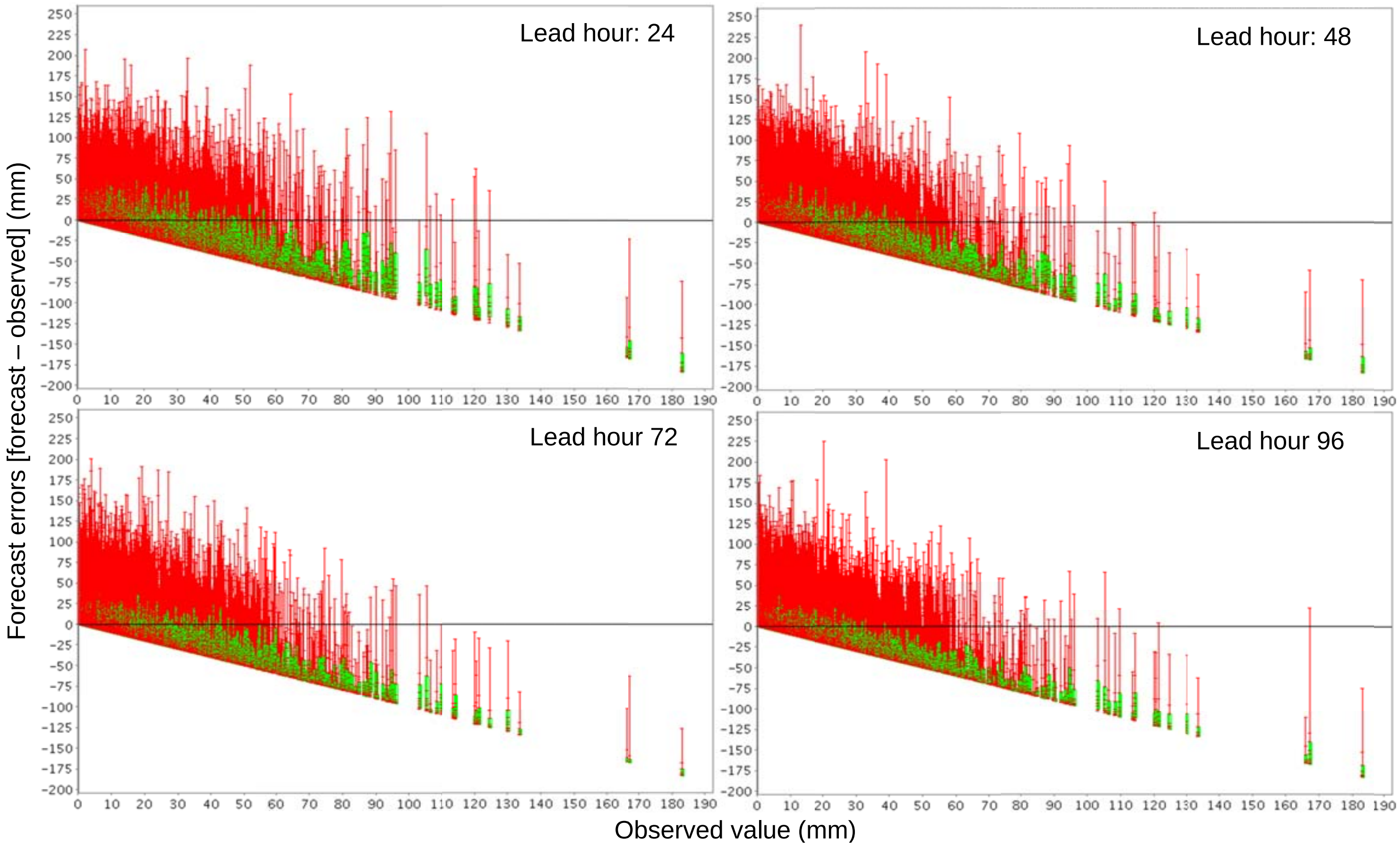


Figure 4.11 Box plot of forecast errors of the ensemble precipitation hindcasts against observed precipitations for the all 5 headwater basins up to Day-4

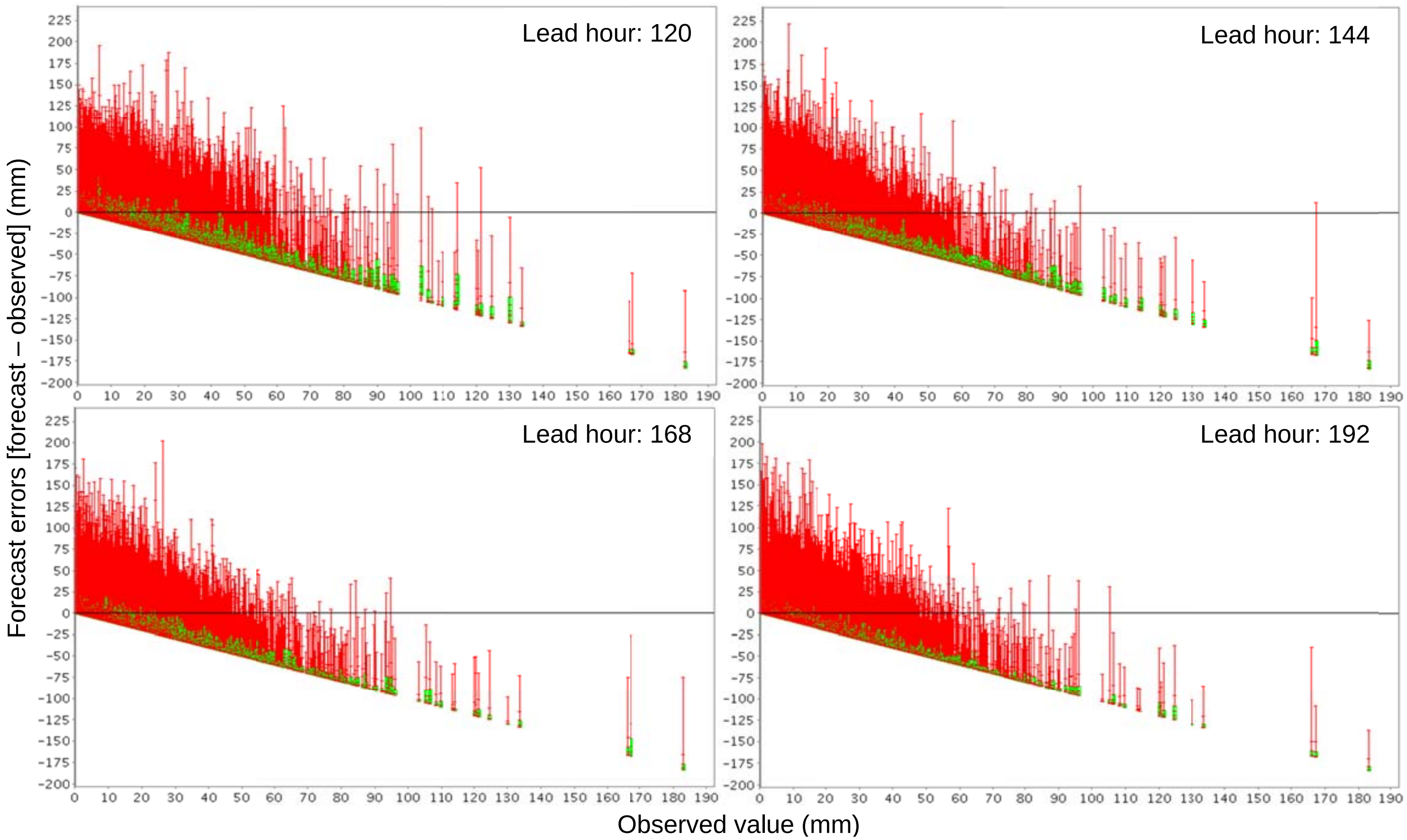


Figure 4.12 Box plot of forecast errors of the ensemble precipitation hindcasts against observed precipitations for the all 5 headwater basins up to Day-8

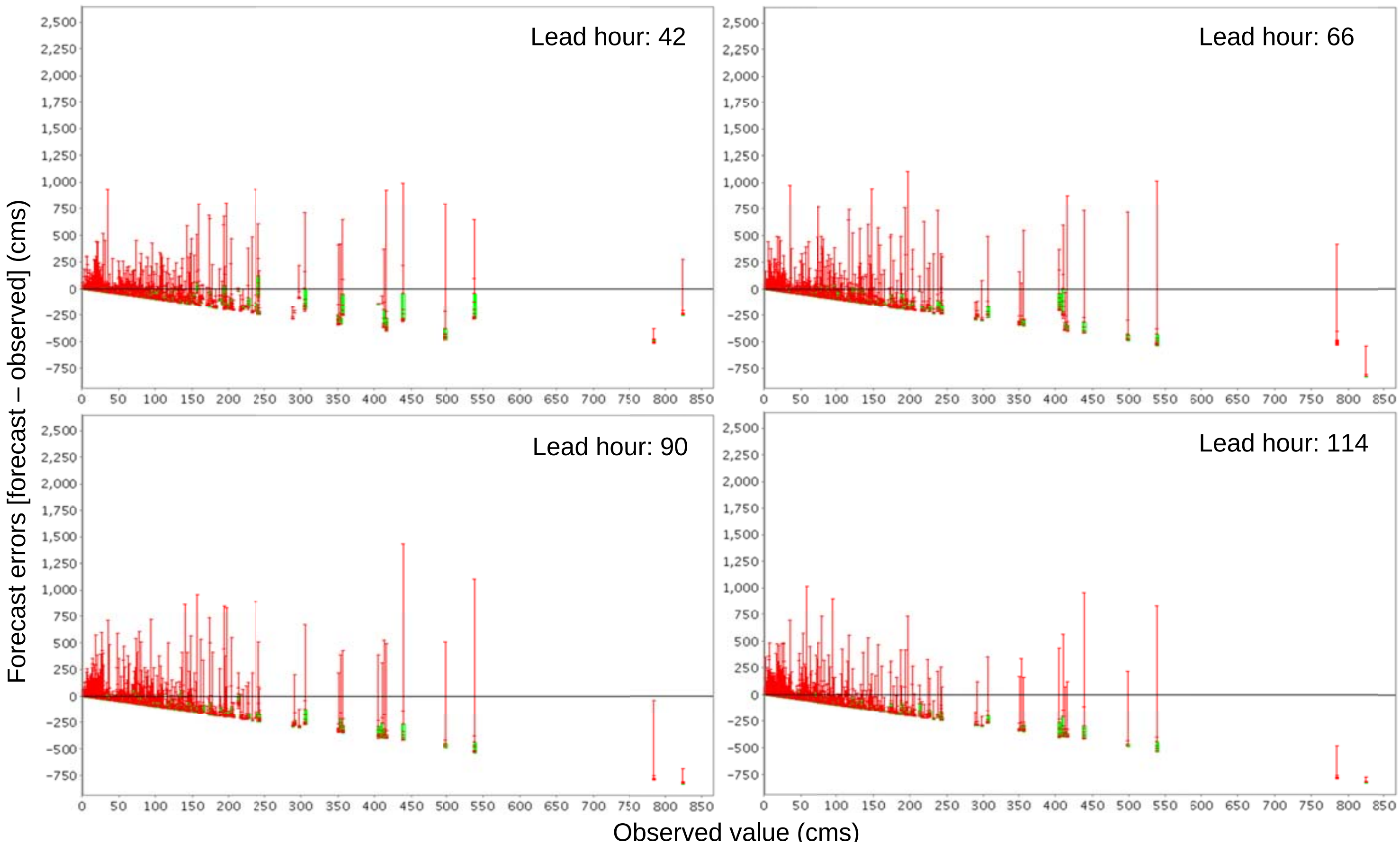


Figure 4.13 Box plot of forecast errors of raw streamflow hindcasts against observed values for the all 5 headwater basins up to Hour-114

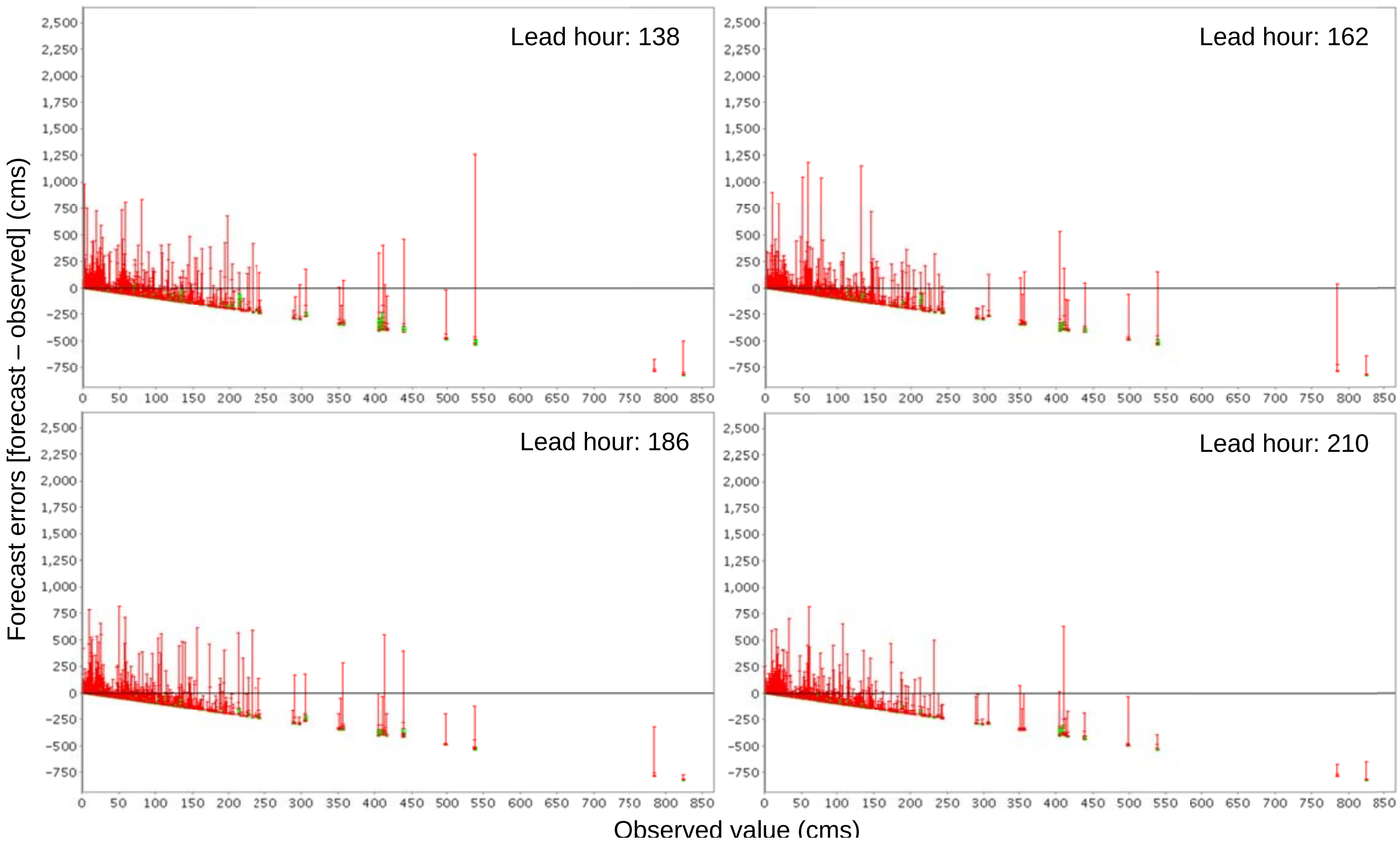


Figure 4.14 Box plot of forecast errors of raw streamflow hindcasts against observed values for the all 5 headwater basins up to

Hour-210

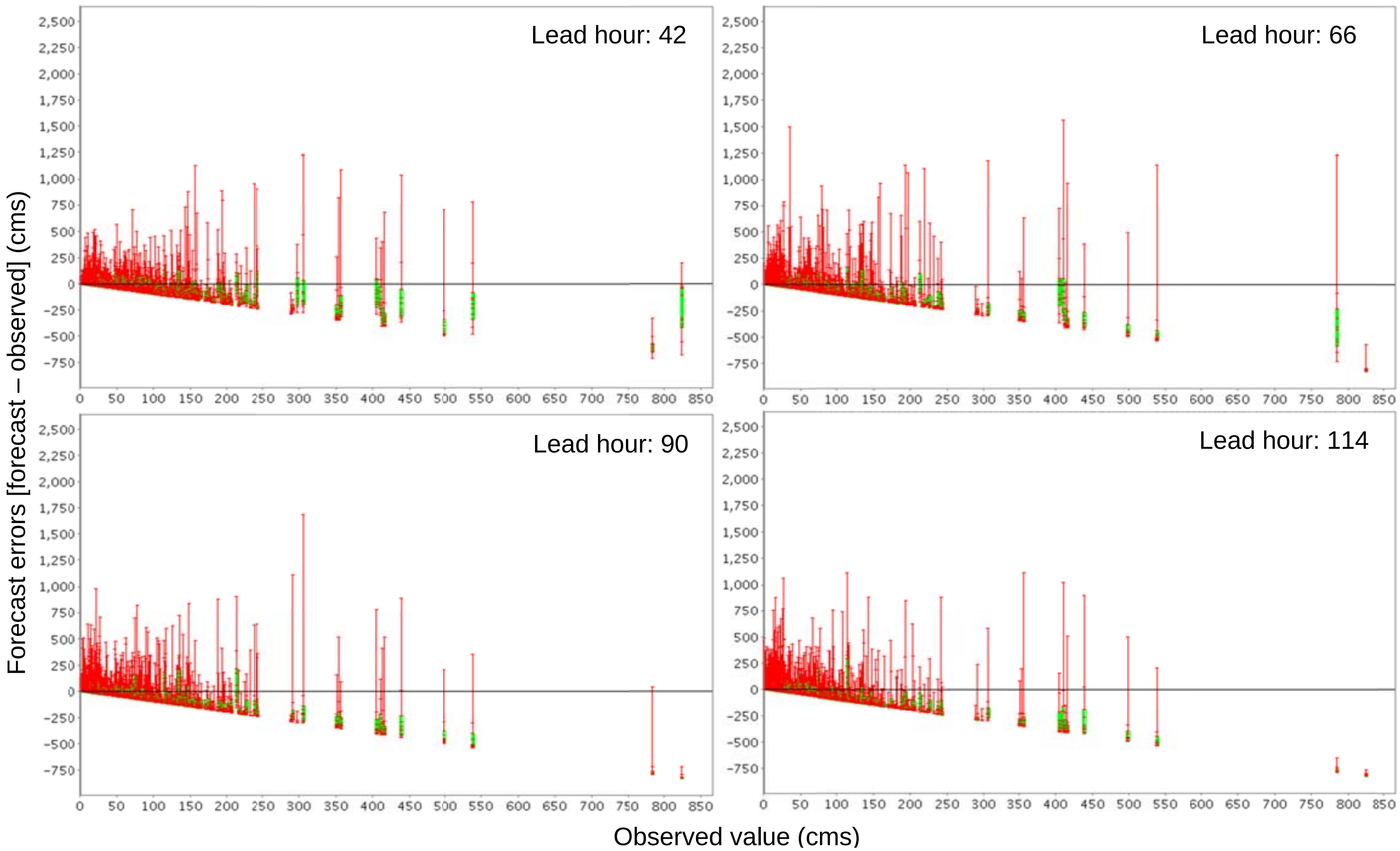


Figure 4.15 Box plot of forecast errors of post-processed streamflow hindcasts against observed values for the all 5 headwater basins up to Hour-114

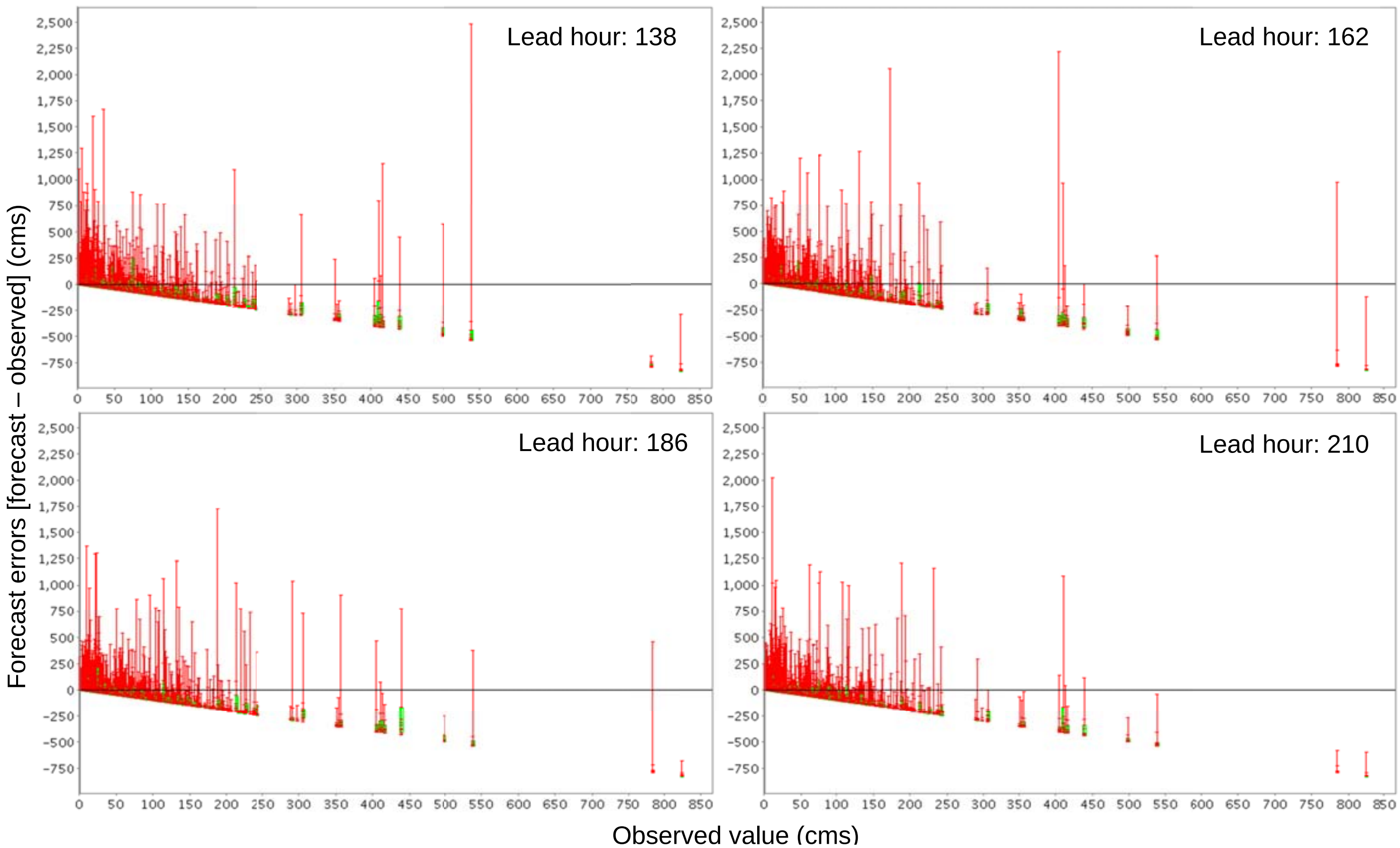


Figure 4.16 Box plot of forecast errors of post-processed streamflow hindcasts against observed values for the all 5 headwater basins up to Hour-210

4.2.4 Continuous Ranked Probability Skill Score (CRPSS)

Figure 4.17 shows the CRPSS of ensemble precipitation hindcasts as a function of forecast lead time. CRPSS for all thresholds is approximately 0.3 for Day 1 and gradually declines as lead time increases. While not very high, but the CRPSS is significantly positive out to several days.

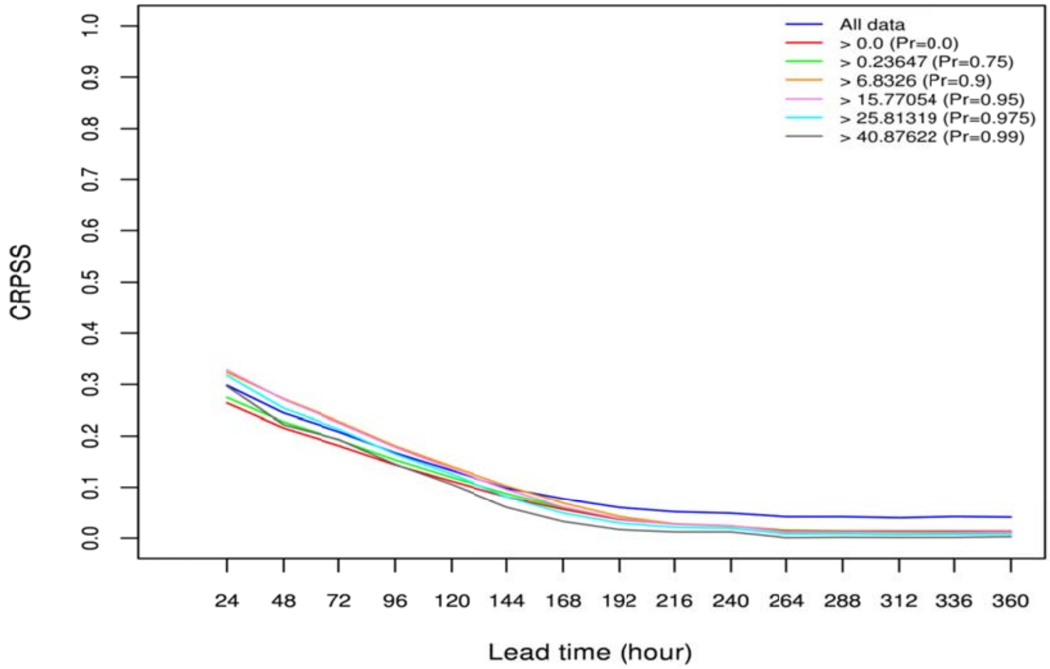


Figure 4.17 CRPSS of ensemble precipitation hindcasts

The CRPSS decreases with increasing lead time and increasing threshold for both raw and post-processed streamflow hindcasts (Figure 4.18). The post-processed streamflow hindcasts show consistently larger predictive skill up to 70%, from 0.28 to 0.48, for Day 1.

Figure 4.19 shows that the improvement from post-processing is significant for all thresholds. This is a reflection that the EnsPost is generally successful in reducing biases and capturing hydrologic uncertainty. That the improvement is larger for short lead times is a reflection of the relatively short hydrologic memory in these basins that controls fast runoff.

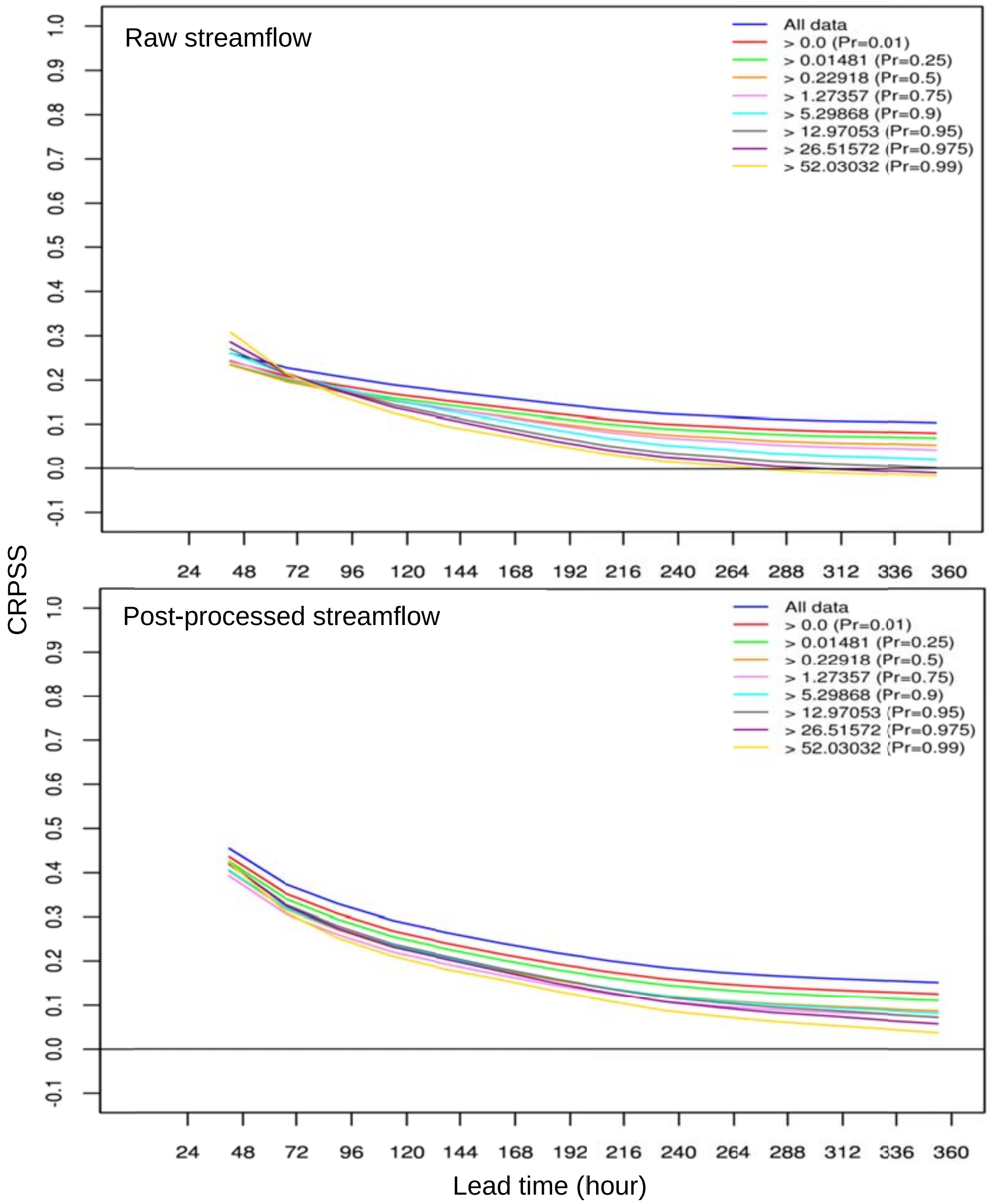


Figure 4.18 CRPSS for ensemble raw streamflow hindcasts (upper panel) and for ensemble post-processed streamflow hindcasts (lower panel)

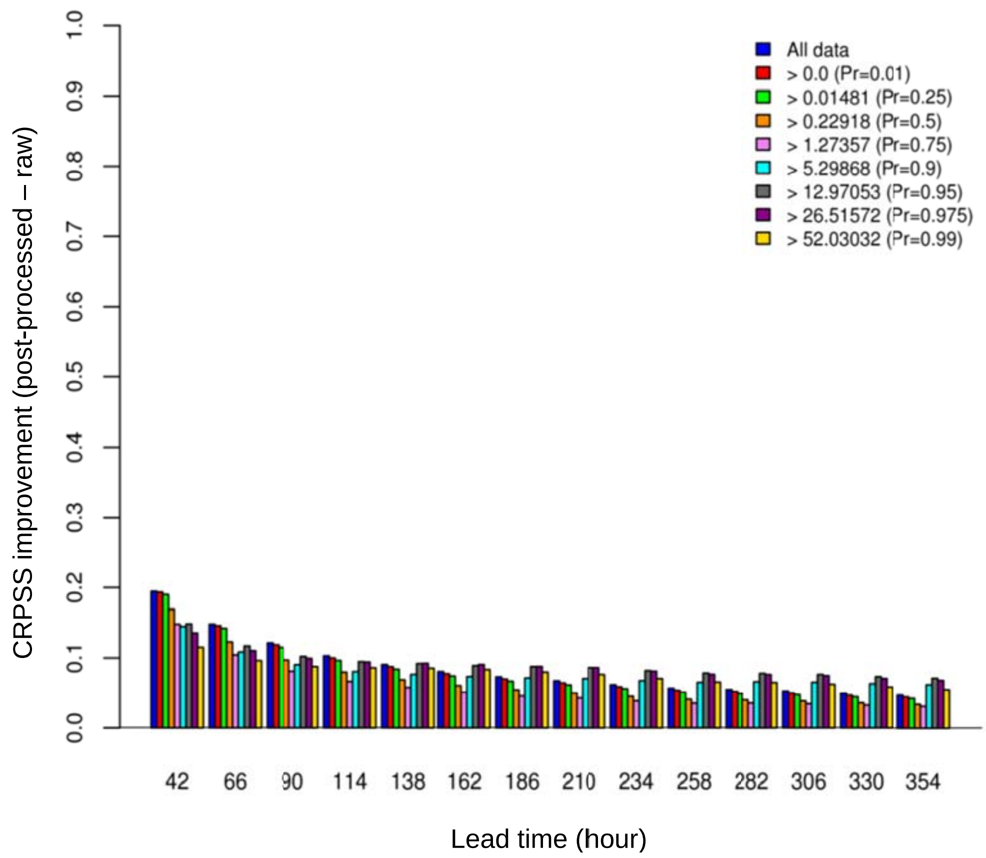


Figure 4.19 CRPSS improvement due to post-processing raw streamflow hindcasts

4.2.5 Relative Operating Characteristics Score (ROC score)

ROC score measures discrimination, i.e., the forecast's ability to tell apart a user-defined event from a non-event. For all thresholds, the ROC score of ensemble precipitation hindcasts exceeds 0.7 for Day 1 and declines over the course of Week 1 as lead time increases (Figure 4.20). Figure 4.21 shows the ROC score of raw (upper panel) and post-processed streamflow hindcasts (lower panel). This figure suggests that post-processing generally improves ROC score. Figure 4.22 shows that the improvement is very large for low flows (as high as 120%) but remains significant for high flows as well.

The large improvement in low flow conditions is a reflection that the EnsPost is successful in removing systematic biases in such conditions.

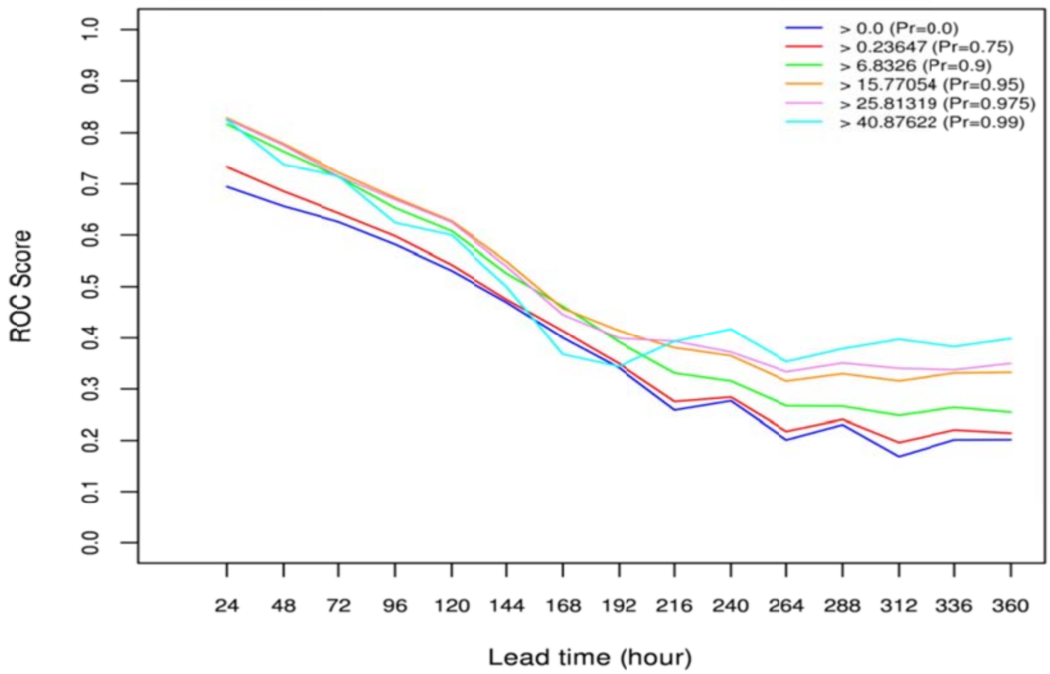


Figure 4.20 ROC Score of ensemble precipitation hindcasts

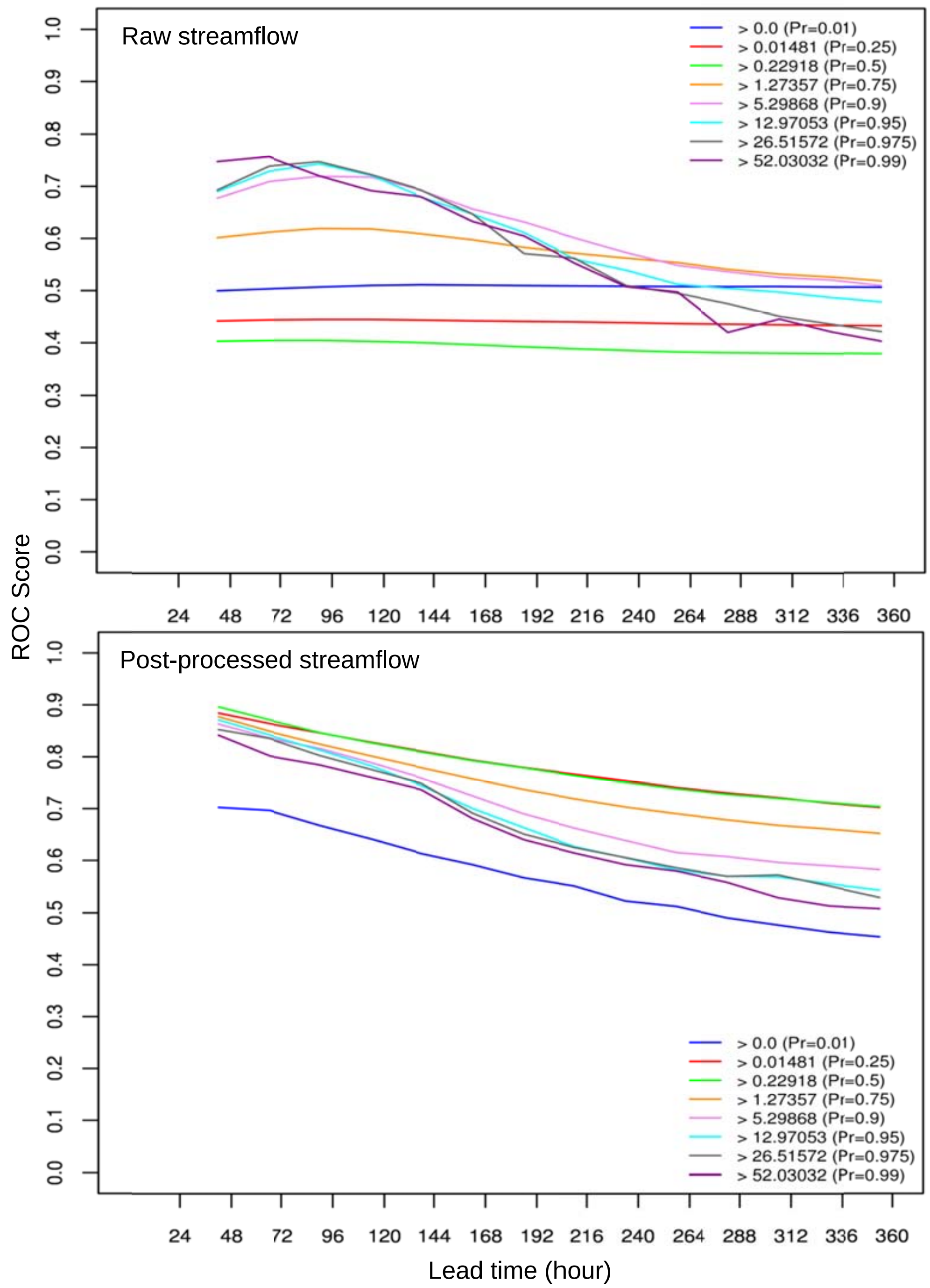


Figure 4.21 ROC Score of streamflow ensemble hindcasts for all 5 headwater basins

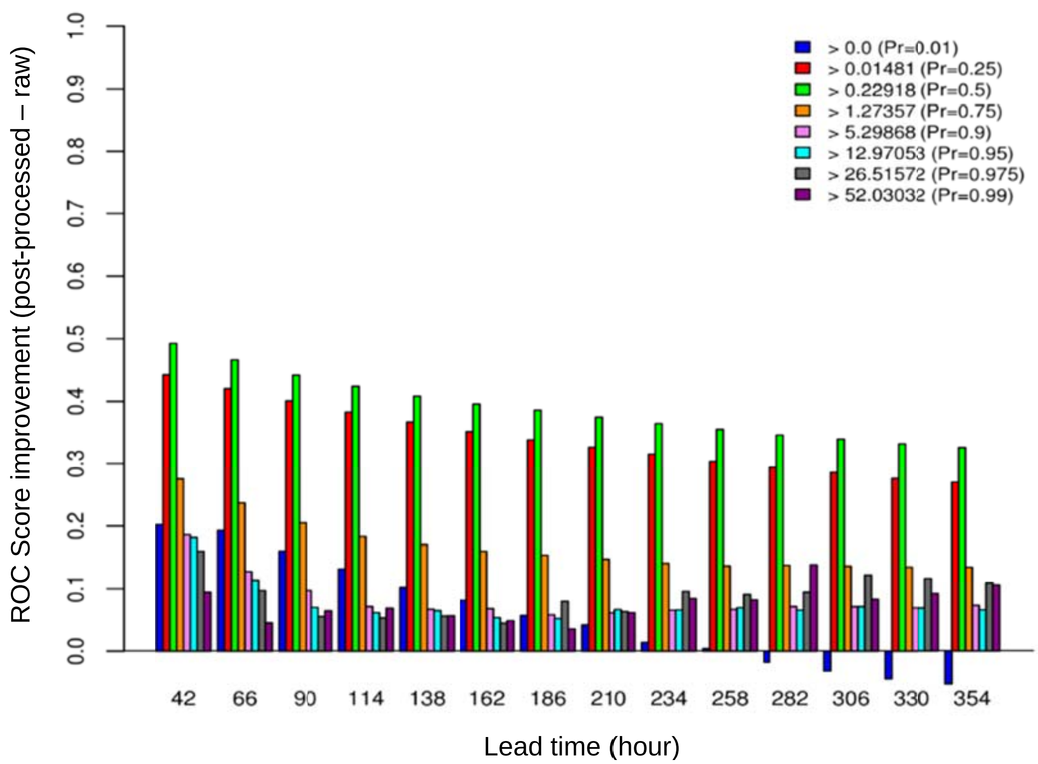


Figure 4.22 ROC Score improvement due to post-processing raw streamflow hindcasts

4.3 Sensitivity analysis

The four hindcast experiments carried out in this study (Table 3.4) to assess the effect of different aspects of the user-defined environmental variables in the MEFPPE and the EnsPostPE on the quality of ensemble forecasts from the MEFP and the EnsPost. Experiment 1 assesses the effect of the sampling period in the EnsPostPE. Experiment 2 assesses the effect of the sampling window in the MEFPPE. Experiments 3 and 4 assess the effect of canonical event definitions in the MEFPPE. For assessing the results from the sensitivity analysis, CRPSS is used throughout in this subsection to provide error statistics which is analogous to the mean absolute error in the single-value sense.

4.3.1 Experiment 1: sampling period for EnsPost parameters

The EnsPost parameters are estimated based on the historical pairs of simulated and observed streamflow values for the user-defined sampling period. In this experiment, monthly (Case 1) and semi-annual (Case 2) periods were used. Figure 4.23 shows differences in CRPSS of post-processed streamflow forecasts generated in Case 1 vs. those in Case 2 by subtracting the CRPSS of Case 2 from that of Case 1. A positive difference hence indicates that the monthly EnsPost parameters generate more skillful hindcasts than the semi-annual parameters. Since the precipitation hindcasts used in both Case 1 and Case 2 are the same, gain in forecast skill results from the EnsPost sampling period only. Although the gain is small (see Figure 4.23), the monthly parameters improve skill up to 10%. This improvement is not surprising in that a high-resolution in seasonality definition in dependent validation amounts to higher-order fit.

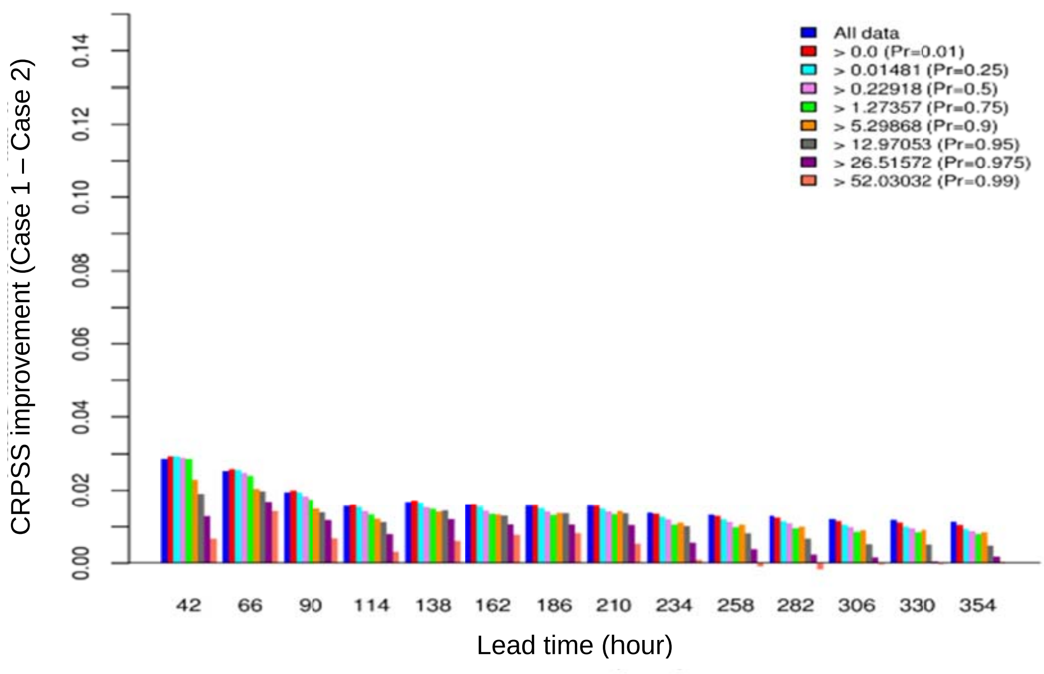


Figure 4.23 Improvement in CRPSS of post-processed streamflow hindcasts generated using Case 1-based and Case 2-based parameters

4.3.2 Experiment 2: sampling window for MEFP parameters

In general, sampling uncertainty decreases as a sampling window increases but at the expense of less day-of-the-year-specific parameters as a wider window can include samples with different seasonality. In this experiment, sampling windows of 61 (Case 1) and 91 days (Case 3) are considered. The results show that the differences in CRPSS between the two cases are negligible in precipitation hindcasts as well as in raw or post-processed streamflow hindcasts (see Figure 4.24, 4.25, and 4.26). Expanding the sampling window does not improve forecast skill because the 26-year period of the GEFS record is long enough to meet the minimum sample size within the sampling window of 61 days to estimate MEFP parameters.

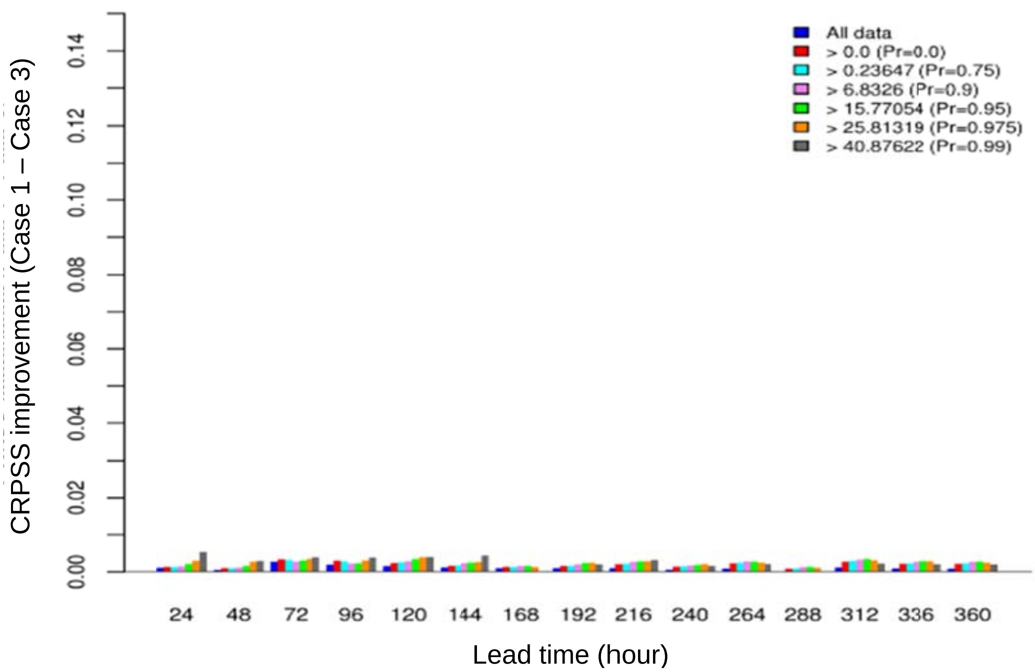


Figure 4.24 Improvement in CRPSS of precipitation hindcasts generated using Case 1-based and Case 3-based parameters

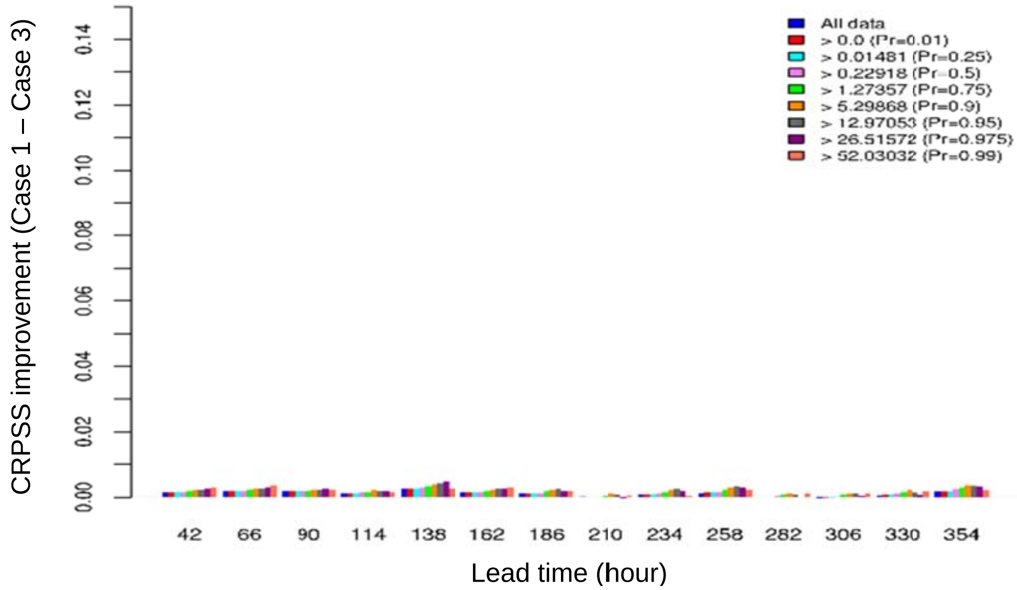


Figure 4.25 Improvement in CRPSS of raw streamflow hindcasts generated using Case 1-based and Case 3-based parameters

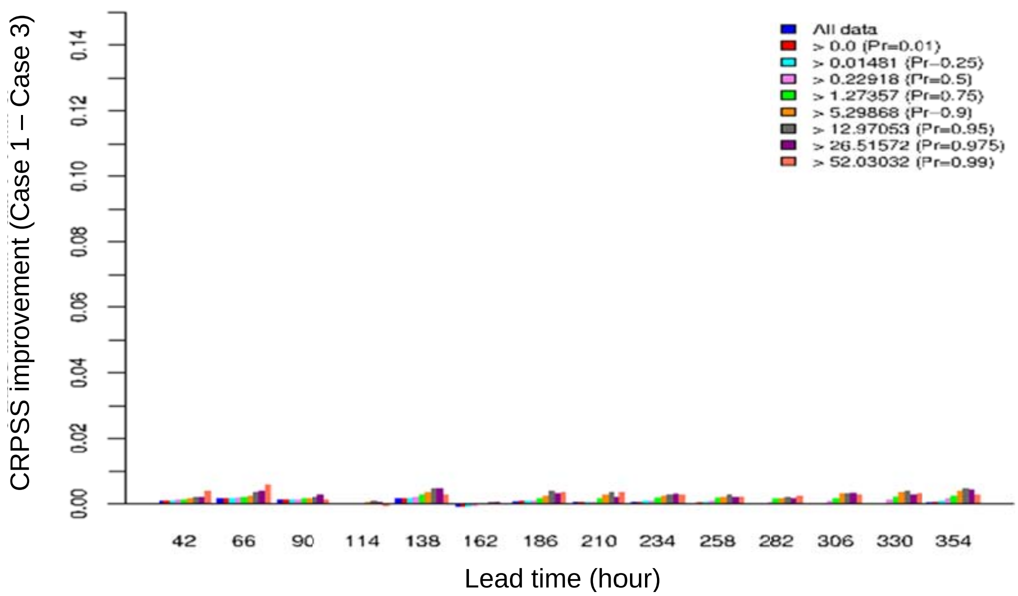


Figure 4.26 Improvement in CRPSS of post-processed streamflow hindcasts generated using Case 1-based and Case 3-based parameters

4.3.3 Experiment 3: Canonical events (base) for MEFP parameters

The main difference in this experiment relative to the default lies in the base events in terms of the number of 6-hour events employed. Case 3 includes only two 6-hour base events at the beginning of the forecast horizon and shifts to 12-hour events up to Day 5, whereas Case 5 includes twenty 6-hour events up to Day-5. Figure 4.27 shows that the CRPSS of Case 3 for precipitation is about 5% higher than that for Case 5 for Day 1 and 10% for Days 2 to 5. There is no gain in Days 6 to 8 because the temporal aggregation scheme in the canonical event definitions is the same over this part of the forecast horizon for both Cases 3 and 5. Starting Day 9, however, gain appears again due to larger temporal aggregation in the canonical event definitions in Case 3 (48 hour- vs. 24 hour-aggregated canonical events). It indicates that aggregating 6-hour events generates more skillful daily precipitation hindcasts. Such a gain, however, does not improve skill in raw or post-processed streamflow hindcasts (see Figure 4.29and 4.31)

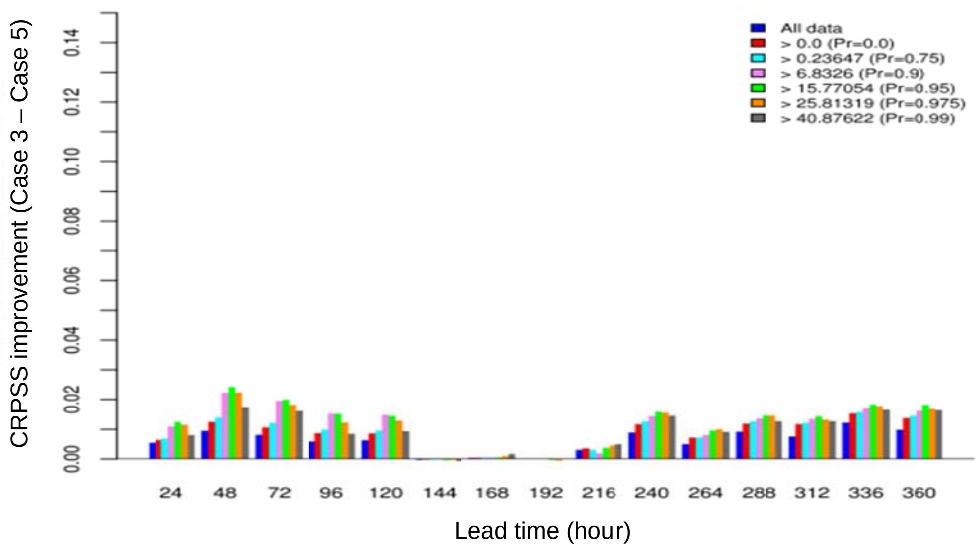


Figure 4.27 Improvement in CRPSS values of precipitation hindcasts generated using Case 3-based and Case 5-based parameters

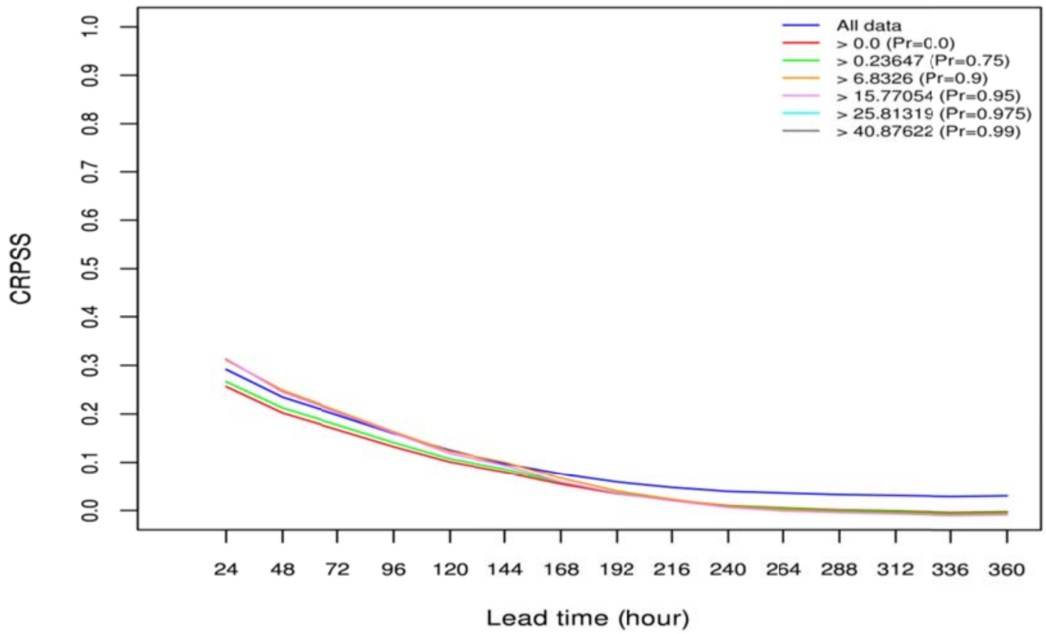


Figure 4.28 CRPSS values of precipitation hindcasts generated using Case 3-based parameters

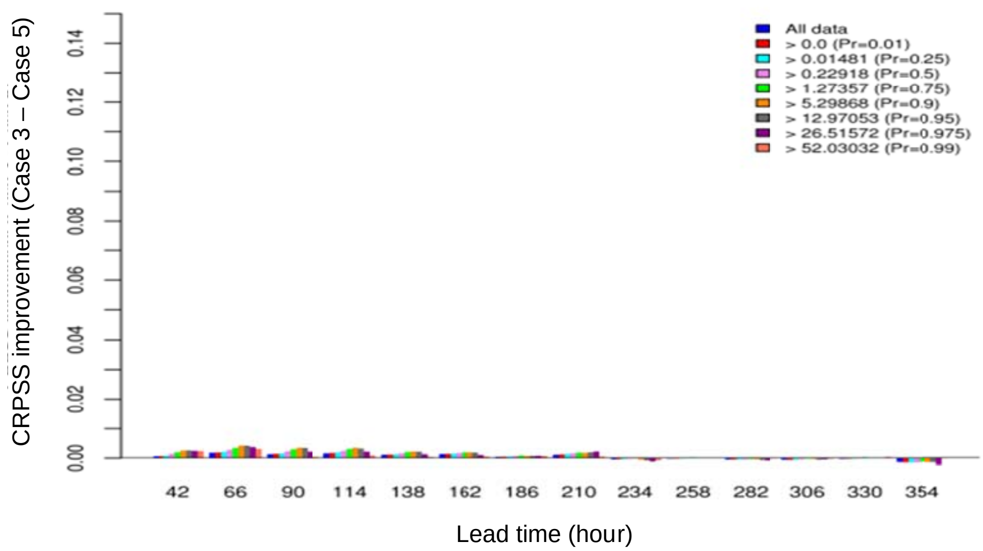


Figure 4.29 Improvement in CRPSS of raw streamflow hindcasts generated using Case 3-based and Case 5-based parameters

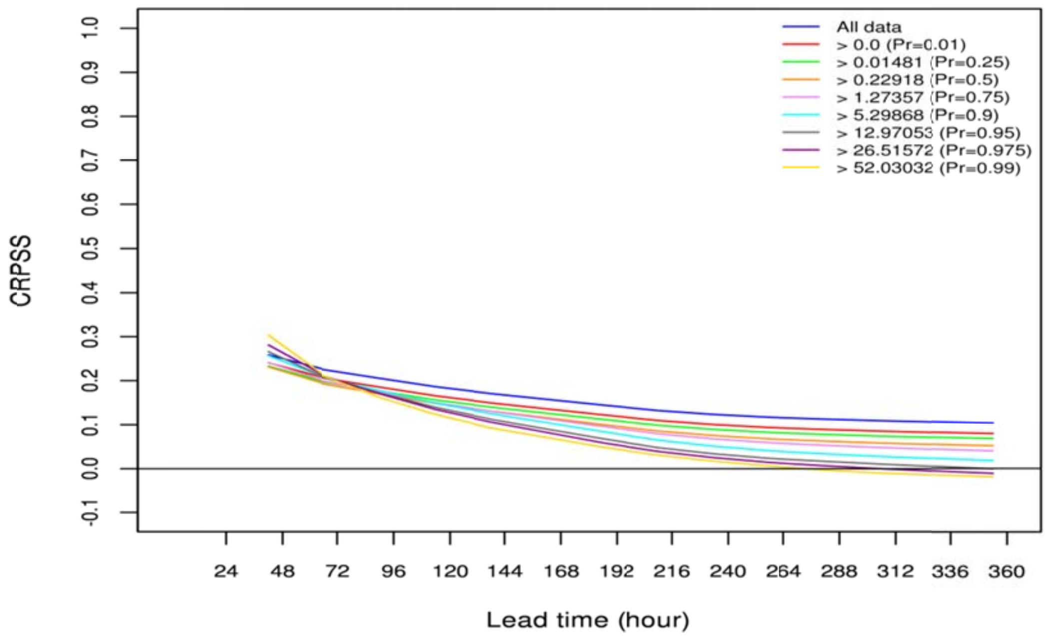


Figure 4.30 CRPSS of raw streamflow hindcasts generated using Case 3-based parameters

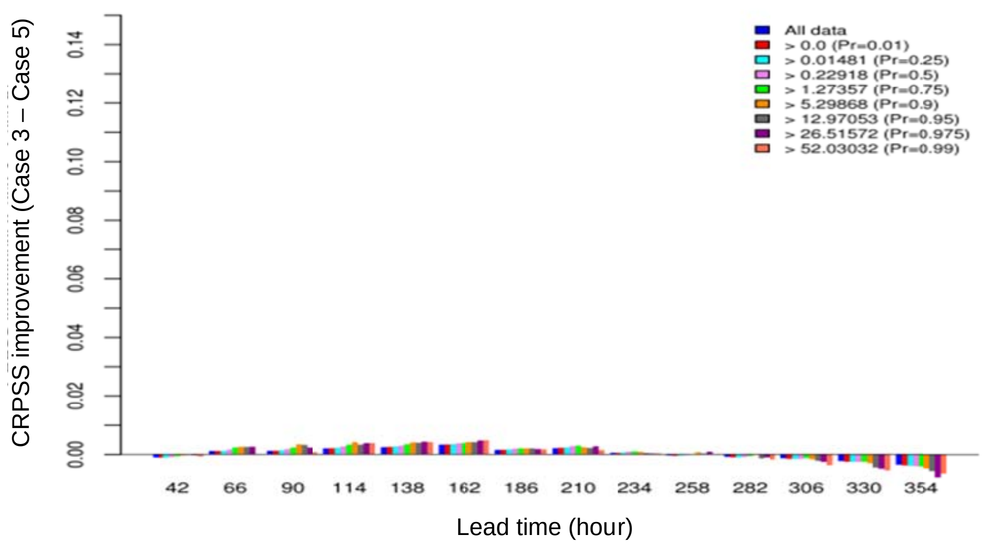


Figure 4.31 Improvement in CRPSS of post-processed streamflow hindcasts generated using Case 3-based and Case 5-based parameters

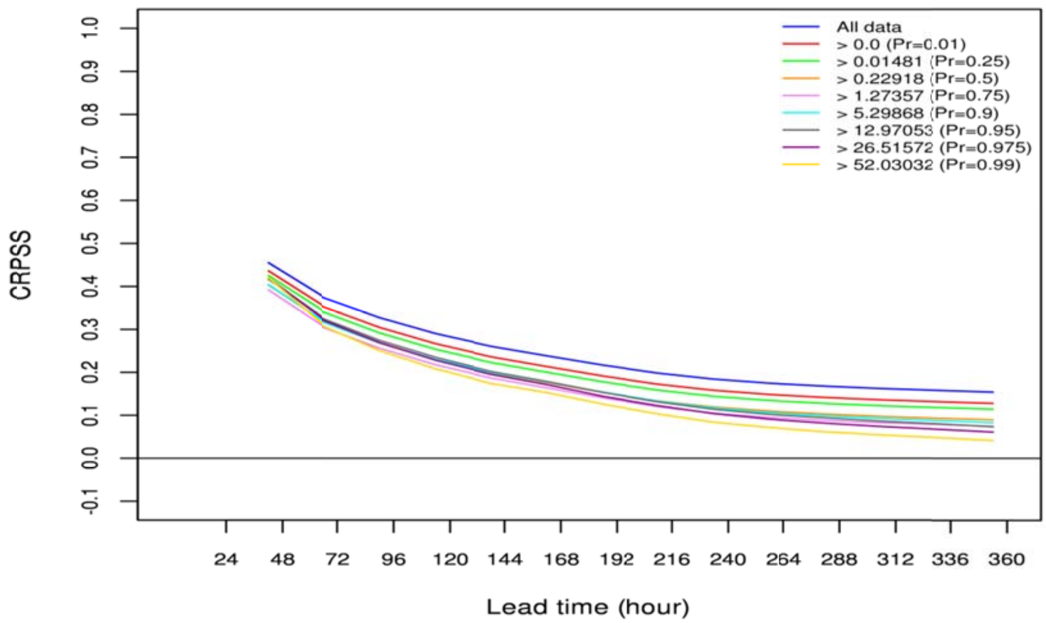


Figure 4.32 CRPSS of post-processed streamflow hindcasts generated using Case 3-based parameters

4.3.4 Experiment 4: Canonical events (modulation) for MEFP parameters

The effect of adding modulation canonical events on MEFP parameters is assessed in this experiment. By adding modulation events (Case 4), the skill improvement in precipitation hindcasts ranges from 5 to 20% for low thresholds (< 90th percentile) and from 15 to 35% for high thresholds (> 90th percentile) (see Figure 4.33). It suggests that, by adjusting modulation canonical events, MEFP can capture the underlying skill of GEFS for larger temporal aggregation period, which may be important in medium range forecasts. Such gains hold up to Day 5 and improve skill in both raw and post-processed streamflow hindcasts past Day 5 (see Figure 4.34 and 4.35). Because the EnsPost parameters were the same for both Cases 3 and 5, the improved skill in the post-processed streamflow hindcasts came from the improved skill in the raw streamflow hindcasts. The above results suggest that the improved skill in precipitation hindcasts improves raw and post-processed streamflow hindcasts more for high flow conditions, a reflection of the fact that significant improvement in skill in precipitation forecast occurs over larger amounts.

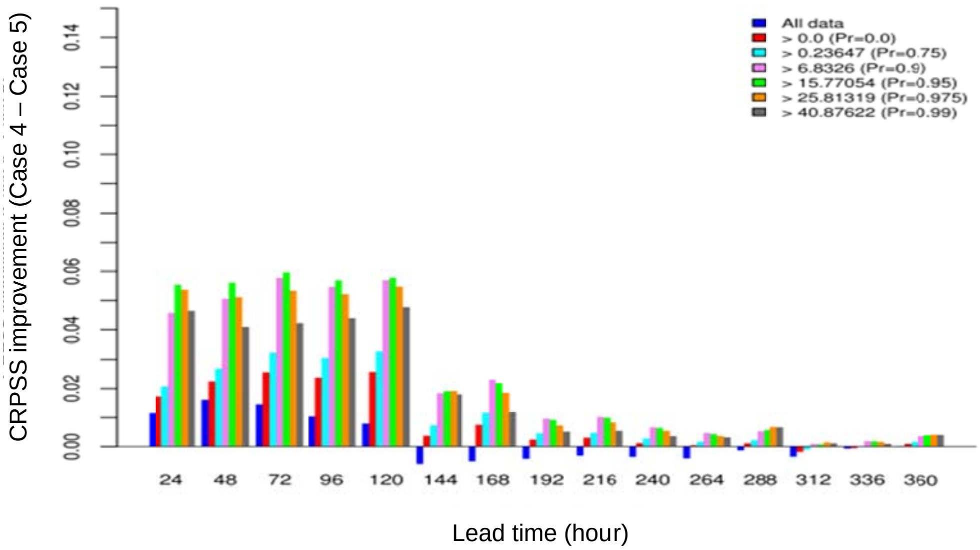


Figure 4.33 CRPSS improvement in of precipitation hindcasts generated using Case 4-based and Case 5-based parameters

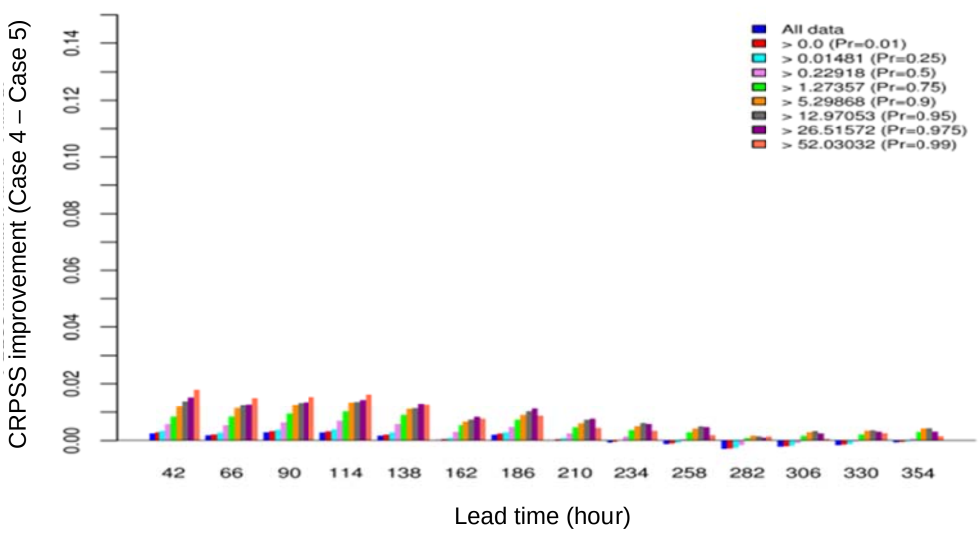


Figure 4.34 Improvement in CRPSS of raw streamflow hindcasts generated using Case 4-based and Case 5-based parameters

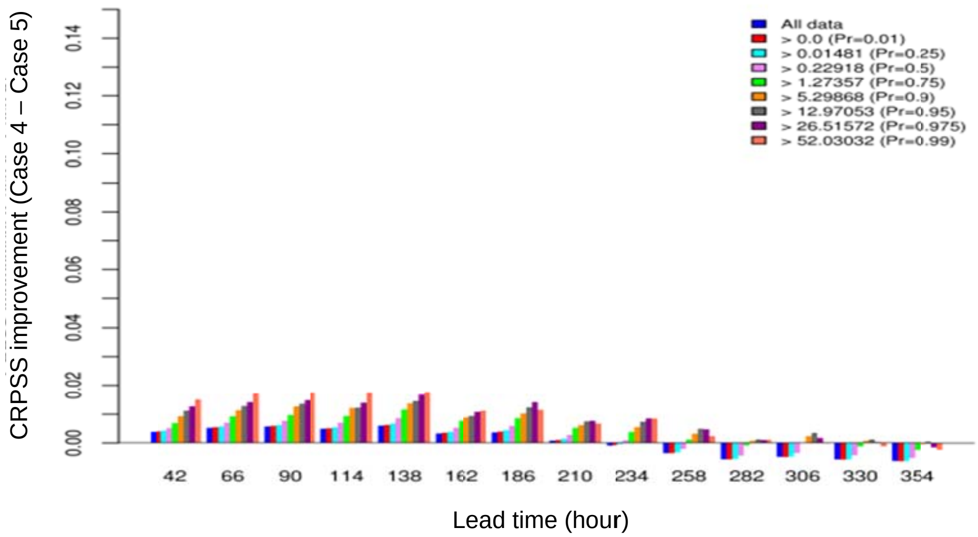


Figure 4.35 Improvement in CRPSS of post-processed streamflow hindcasts generated using Case 4-based and Case 5-based parameters

4.4 Interpretation of ensemble forecasts in a single-valued sense

Various probabilistic attributes existing in ensemble hindcasts are selected and converted into readily useful information to decision makers. Ensemble hindcasts used in this section are the ensemble precipitation, raw streamflow, and post-processed streamflow hindcasts generated using Case 4-based parameters based on the sensitivity analysis in Section 4.3.

4.4.1 Probability of detection at a given false alarm rate

As explained in Chapter 2, ROC expresses PoD and FAR at different levels of exceedance probability. Depending on the risk tolerance of the user, he/she may choose a level which may result in a lower FAR with a lower PoD or a higher PoD with a higher FAR. For example, if an FAR of 3-5% is acceptable for 90th percentile precipitation to the user, the PoD of precipitation for Day 1 is 50% and decreases as lead time increases

(Figure 4.36). For 90th percentile streamflow, PoD of raw streamflow is 50% whereas PoD of post-processed streamflow is 65% for Day 1 (Figure 4.37), which illustrates the value of post-processing in the single-valued sense as well.

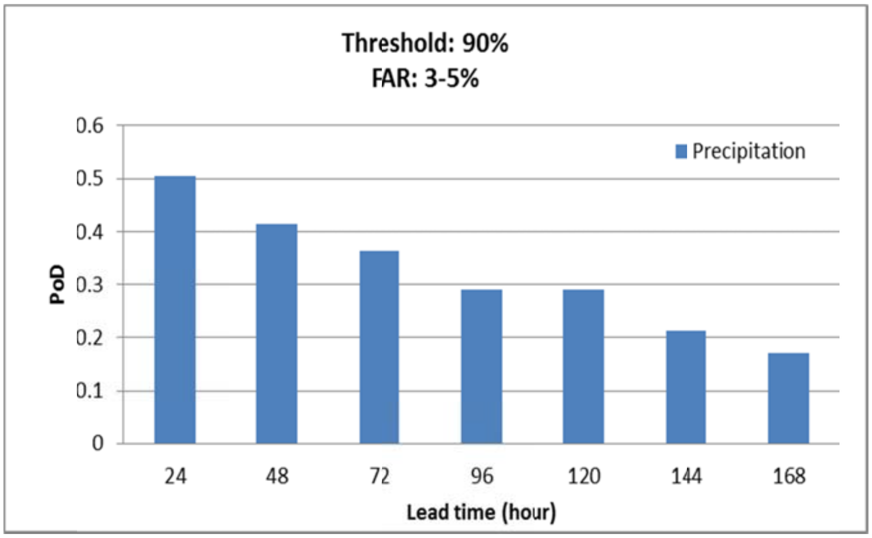


Figure 4.36 PoD vs FAR for precipitation hindcasts

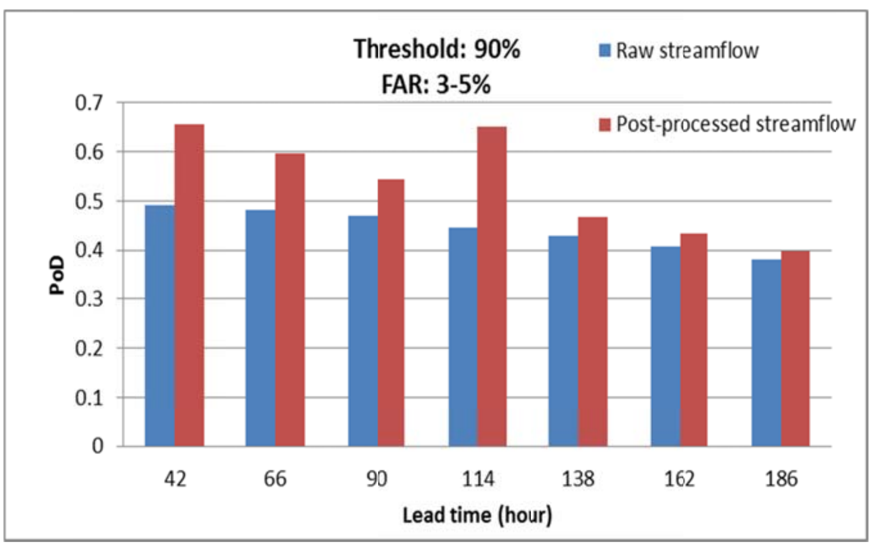


Figure 4.37 PoD vs FAR for streamflow hindcasts

4.4.2 *Temporal aggregation in forecasts*

Figure 4.38 shows that correlation coefficient between precipitation hindcasts and observed precipitation increases as the temporal aggregation period increases up to 14 days. They indicate that GEFS-forced medium-range precipitation forecast for the study area have valuable skill in 1-, 3-, 5-daily, weekly, and biweekly-aggregated forecasts.

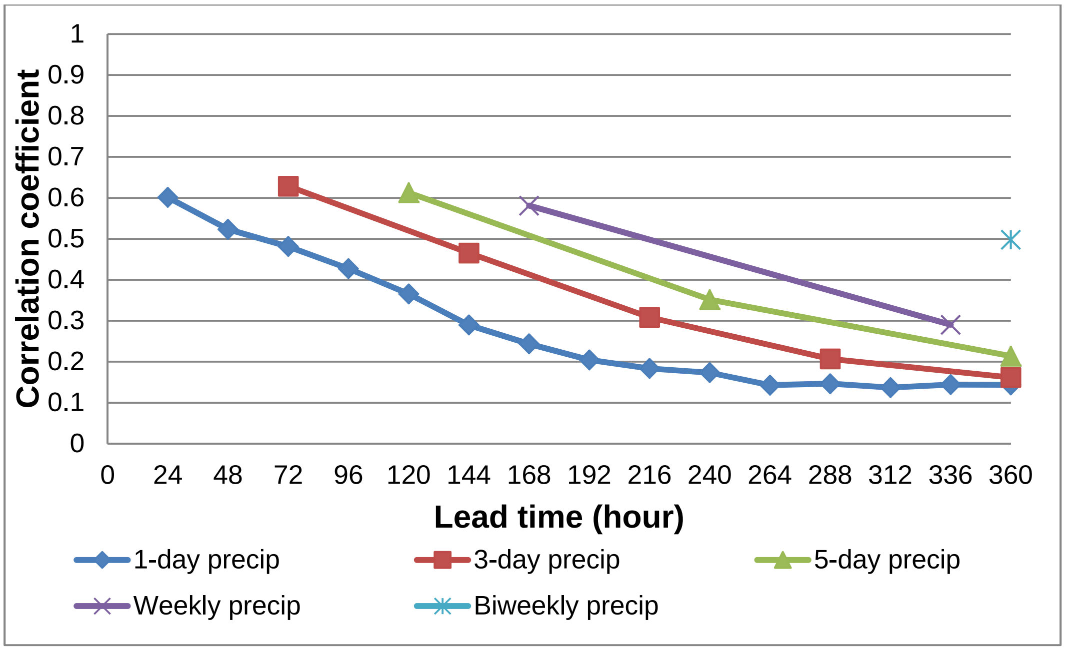


Figure 4.38 Correlation coefficient of precipitation hindcasts and corresponding observations with different temporal aggregation periods

Finally, Figure 4.39 shows the correlation coefficient between post-processed streamflow and observed flow. Note that, with temporal aggregation, there exists very significant skill up to 14 days. They indicate that, with ensemble forecasting, it is possible to effectively utilize the skill in medium-range forecast of precipitation to improve the quality and to increase the lead time of streamflow forecast.

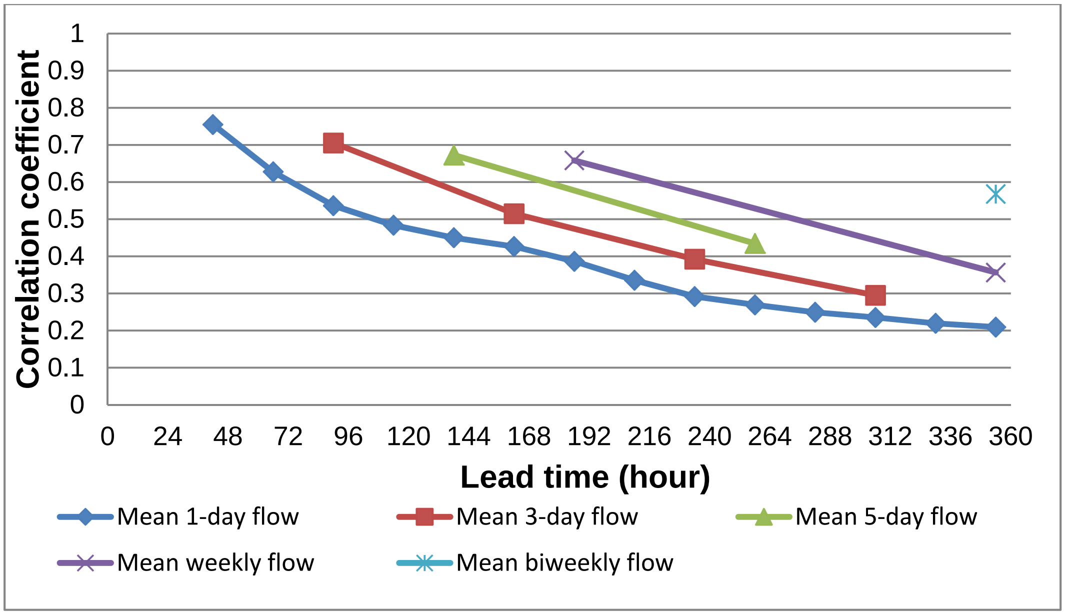


Figure 4.39 Correlation coefficient of post-processed streamflow hindcasts and corresponding observations with different temporal aggregation periods

Chapter 5

Conclusions and future research recommendations

Compared to forecasts of short-term precipitation accumulations (daily or shorter) at lead times larger than a few days, those of longer-term accumulations (3-daily or longer) are significantly more skillful owing to the larger temporal scale of aggregation. If one can utilize this skill present in medium-range precipitation forecast in hydrologic prediction, it is very likely that the lead time of hydrologic forecasts, in particular, of streamflow and soil moisture may be extended. Though forecasts of longer-term accumulations of precipitation are more skillful than those of shorter-term accumulations, precipitation forecasts in general are too uncertain to be used as deterministic, or single-valued, input.

The main goal of this study was to increase forecast lead time of streamflow forecasts by using medium range ensemble precipitation forecasts. A premise for this study is that, in the ensemble paradigm, forecasting of precipitation and streamflow provides extending forecast lead time with improved forecast skill. To utilize forecast skill in medium range precipitation forecasts in the ensemble paradigm, this study used Hydrologic Ensemble Forecast Service (HEFS) developed by the U.S. National Weather Service Office of the Hydrologic Development (NWS/OHD, now the NWS/National Water Center).

In the HEFS, the Meteorological Ensemble Forecast Processor (MEFP) was used to generate ensemble precipitation hindcasts using the Global Ensemble Forecast System (GEFS) reforecast data. Raw streamflow hindcasts were generated via the Community Hydrologic Prediction System (CHPS) using the Sacramento Soil Moisture Accounting model (SAC-SMA) and unit hydrograph. To reduce biases and uncertainties in the hydrologic model results, raw streamflow ensembles were post-processed by the

Ensemble Postprocessor (EnsPost). The precipitation, raw and post-processed streamflow ensembles were verified using the Ensemble Verification System (EVS) to assess the quality of hindcasts.

Ensemble hindcasts of precipitation and streamflow were generated using the HEFS for a 26-year period between 1986 and 2011. The study area consisted of five headwater basins located upstream of the Dallas-Fort Worth (DFW) metropolitan area in the Upper Trinity River Basin in Texas. These study basins offer a tough test for the HEFS, because precipitation is dominated by convection which has very limited predictability. The basins are flashy with fast-rising streamflow when they respond to rainfall but also with periods of no streamflow.

The main findings of this study include:

(1) The ensemble QPF generated from the single-valued WGRFC QPF using the MEFP in the HEFS has forecast skill for long forecast lead time (up to Day-3), when compared to the lead time provided by the single-valued WGRFC QPF used in current practice (6-hour in general).

(2) Medium range GEFS-forced ensemble precipitation hindcasts generated with the MEFP in the HEFS has forecast skill up to more than a week, longer forecast lead time than that offered by the short-range ensemble QPF generated with the MEFP (up to Day-3).

(3) Having monthly sampling period for estimating EnsPost parameters improves forecast skill in post-processed streamflow hindcasts, when compared to semi-annual sampling period. This improvement is not surprising in that a high-resolution in seasonality definition in dependent validation amounts to higher-order fit.

(4) Controlling sampling window (61 vs 91 days) for estimating MEFP parameters does not affect forecast skill in GEFS-forced ensemble precipitation forecasts. Expanding the sampling window does not improve forecast skill because the 26-year period of the GEFS record is long enough to meet the minimum sample size within the sampling window of 61 days to estimate MEFP parameters.

(5) Aggregating 6-hour base canonical events generates more skillful daily precipitation hindcasts. Such a gain, however, does not improve skill in raw or post-processed streamflow hindcasts. The gain in precipitation hindcasts is probably too small to dominate hydrologic uncertainty occurred during the hydrologic process via hydrologic models.

(6) Adjusting modulation canonical events is a very effective way to improve predictive skill in ensemble forecasts of precipitation, raw, and post-processed streamflow forecasts. The skill improvement in precipitation hindcasts ranges from 5% to 35%, holding up to Day 5. It suggests that, by adjusting modulation canonical events, MEFP can capture the underlying skill of GEFS for larger temporal aggregation period, which may be important in medium range forecasts. Such improvement enhances skill in both raw and post-processed streamflow hindcasts past Day 5, more effectively for high flow condition. This indicates that significant improvement in skill in precipitation forecasts occurs over larger amounts. The improved skill in the post-processed streamflow hindcasts came from the improved skill in the raw streamflow hindcasts.

(7) Correlation coefficients between precipitation hindcasts and observed precipitation increase as the temporal aggregation period increases up to 14 days. They indicate that GEFS-forced medium-range precipitation hindcasts for

the study area have valuable skill in 1-, 3-, 5-daily, weekly, and biweekly-aggregated hindcasts.

(8) With temporal aggregation, there exists very significant skill in post-processed streamflow up to 14 days. They suggest that, with ensemble forecasting, it is possible to effectively utilize the skill in medium-range forecast of precipitation to improve the quality and to increase the lead time of streamflow forecasts.

This study used the HEFS successfully, demonstrating the HEFS's portability in the Unix/Linux environment outside of NWS. This study also showed that the HEFS is an effective tool for generating skillful forecasts of precipitation and streamflow ensembles. In the ensemble paradigm, forecast skill in medium-range precipitation forecasts can be effectively utilized to improve the quality of streamflow forecasts in extended forecast lead time via HEFS. This study contributed to the knowledge of providing water resources managers with improved streamflow forecasts for the extended forecast lead time for effective both management of water resource and mitigation of water-related hazards.

The main recommendations for future research are as follows:

(1) Extend the study to a large number of basins for large-sample verification, especially for large events.

(2) Develop and implement the parametric uncertainty processor and the ensemble data assimilator (DA). The current statistical techniques for modeling and reducing hydrologic uncertainty should be upgraded to take account of the dynamics of urbanization and possibly climate change.

Appendix A

Hindcast results of individual headwater basins

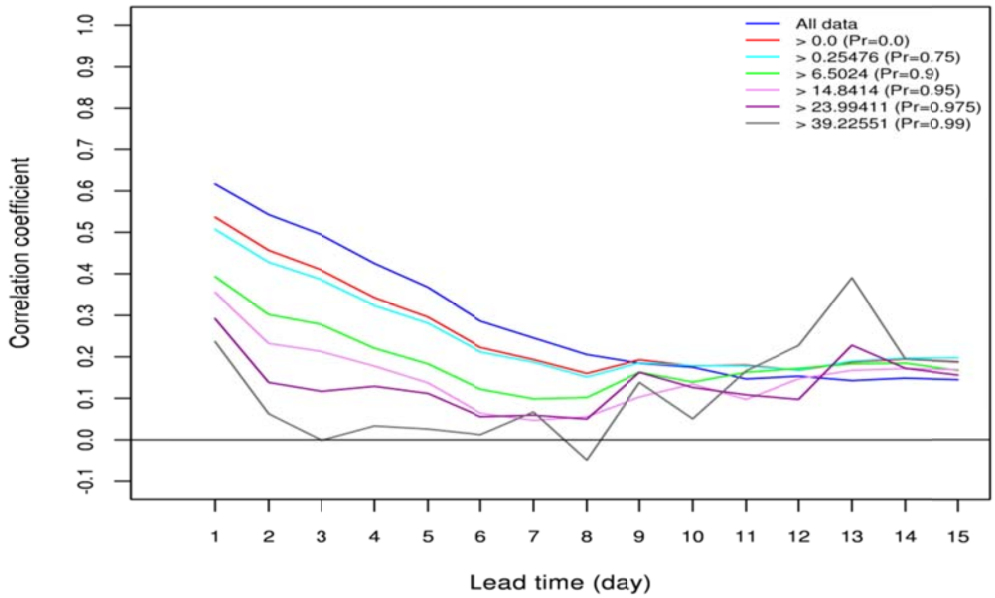


Figure A. 1 Correlation Coefficient of ensemble mean precipitation forecasts and corresponding observed values for the BRPT2

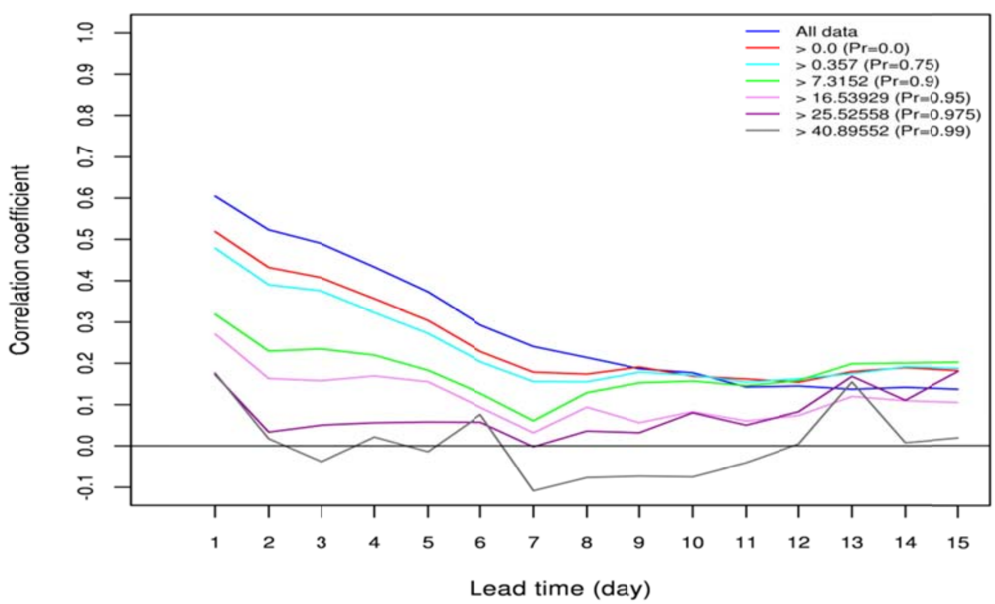


Figure A. 2 Correlation Coefficient of ensemble mean precipitation forecasts and corresponding observed values for the DCJT2

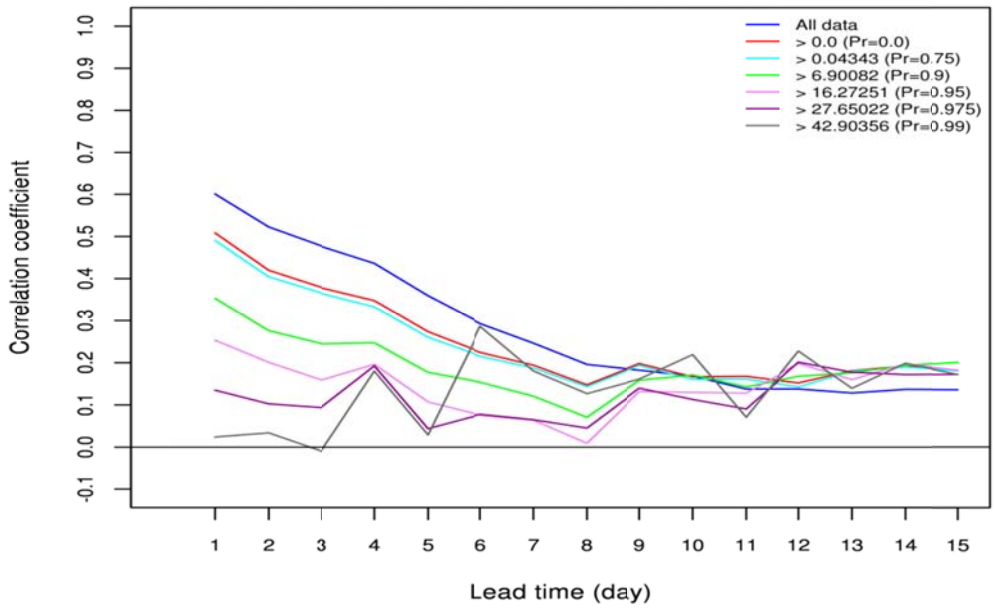


Figure A. 3 Correlation Coefficient of ensemble mean precipitation forecasts and corresponding observed values for the GLLT2

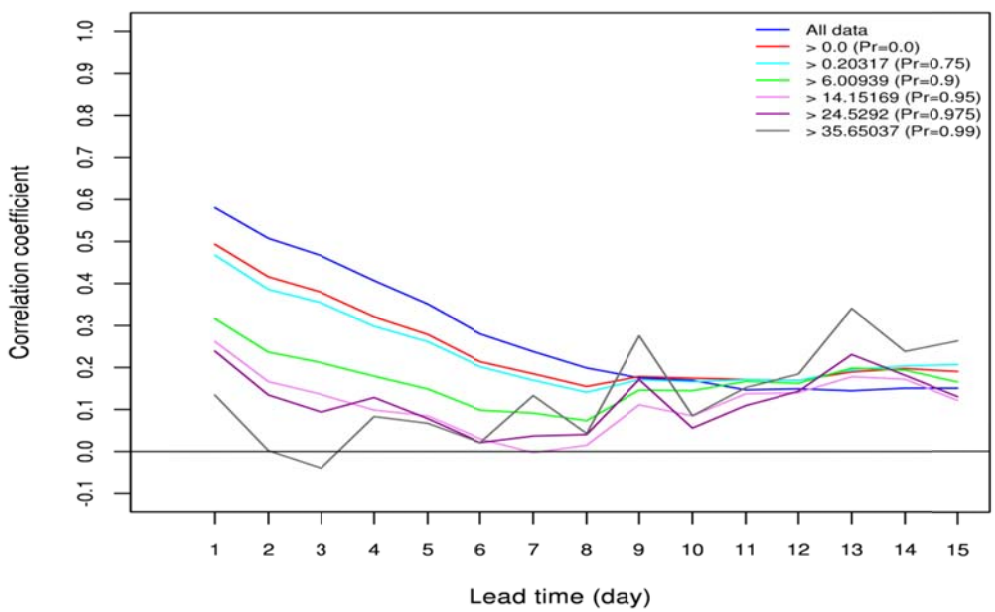


Figure A. 4 Correlation Coefficient of ensemble mean precipitation forecasts and corresponding observed values for the JAKT2

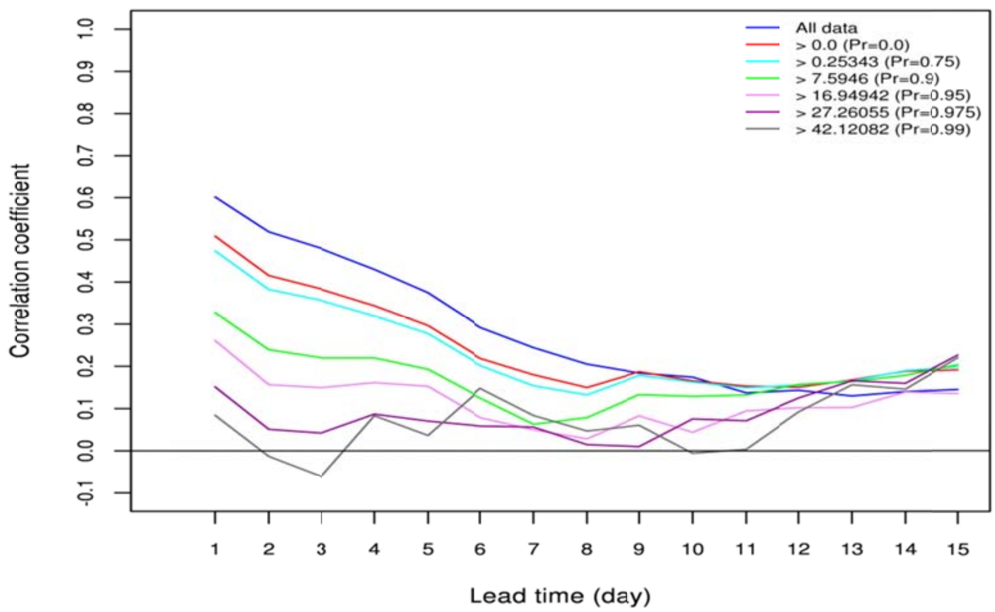


Figure A. 5 Correlation Coefficient of ensemble mean precipitation forecasts and corresponding observed values for the SGET2

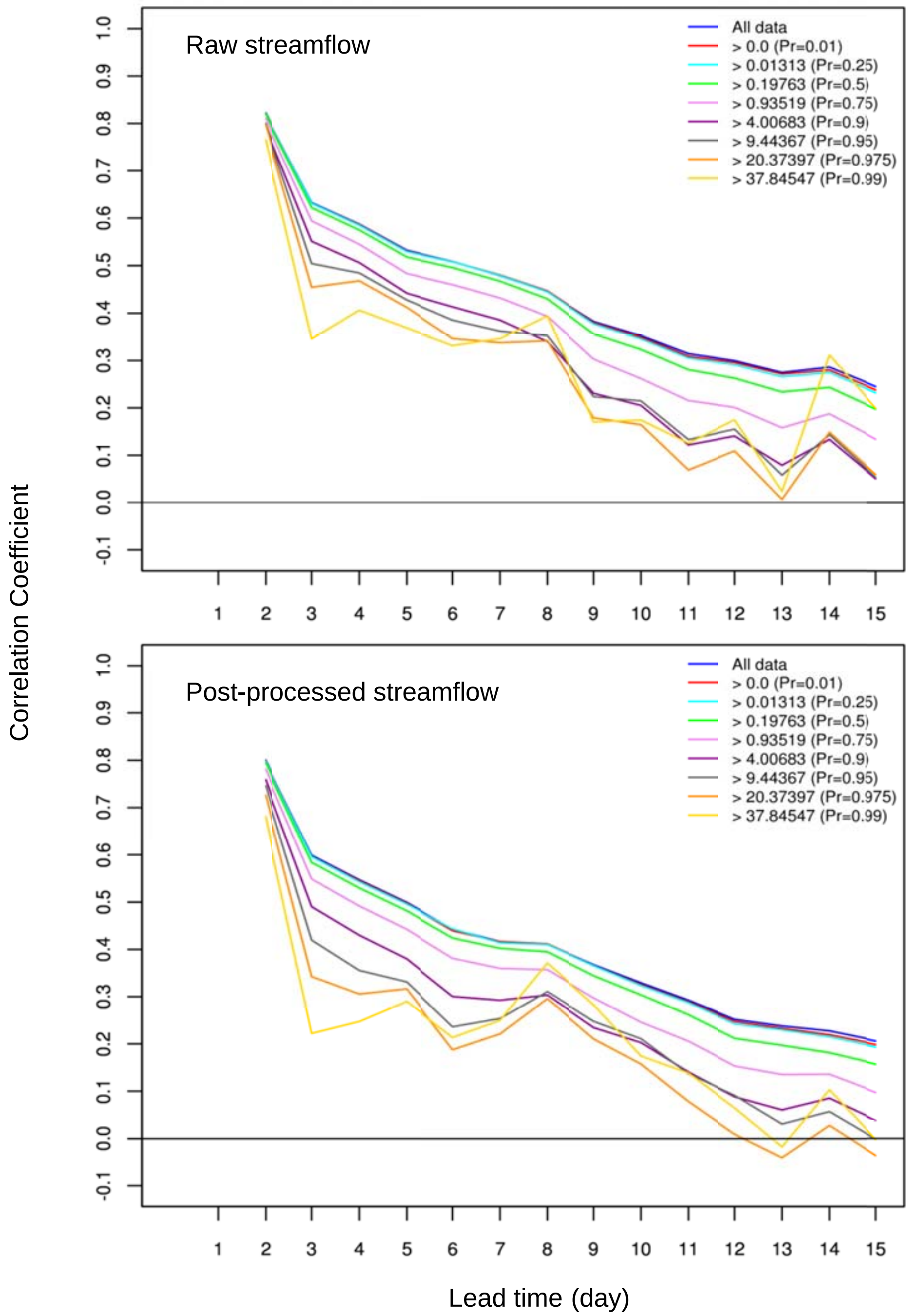


Figure A. 6 Correlation Coefficient of ensemble mean streamflow forecasts and corresponding observed values for the BRPT2

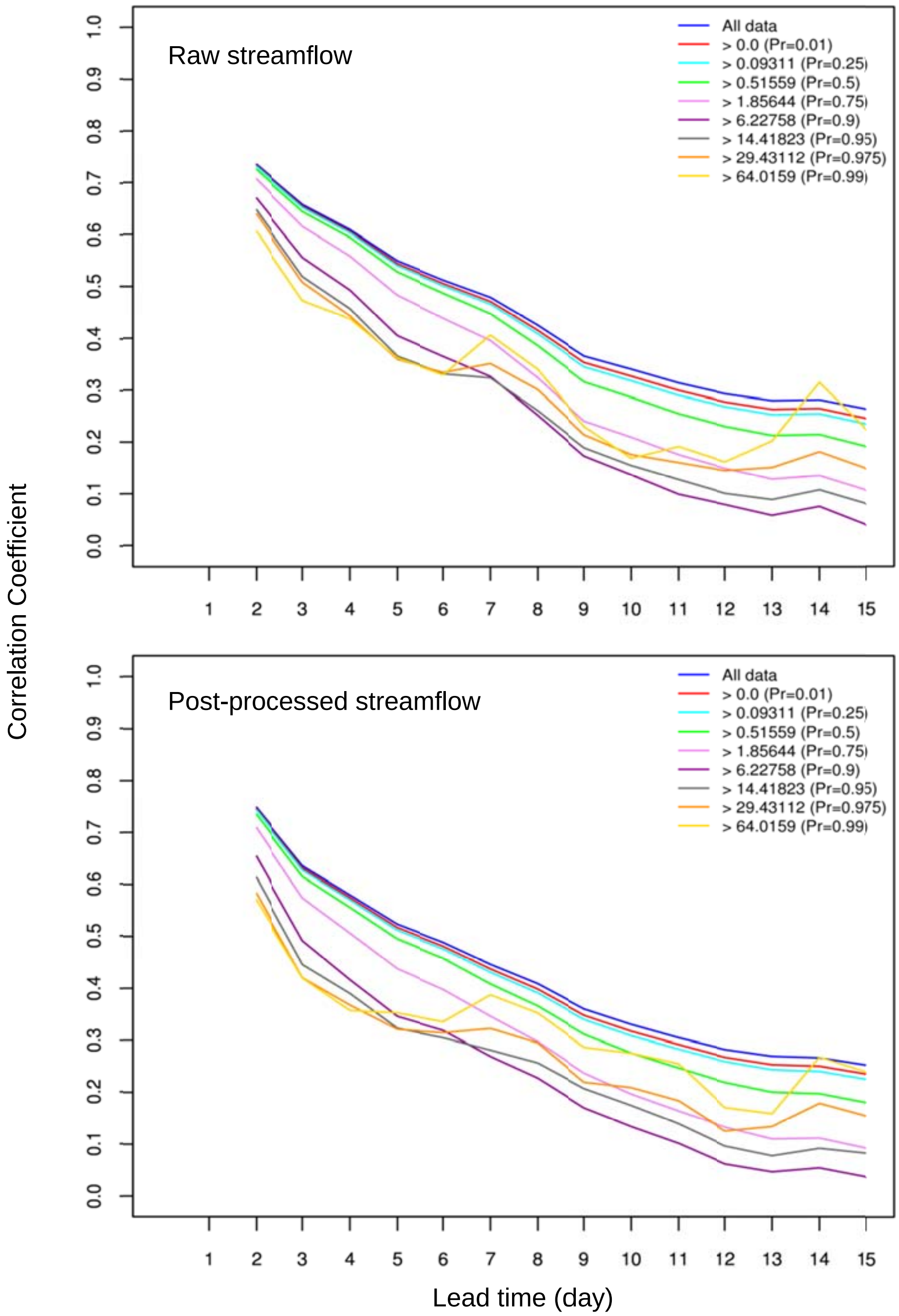


Figure A. 7 Correlation Coefficient of ensemble mean streamflow forecasts and corresponding observed values for the DCJT2

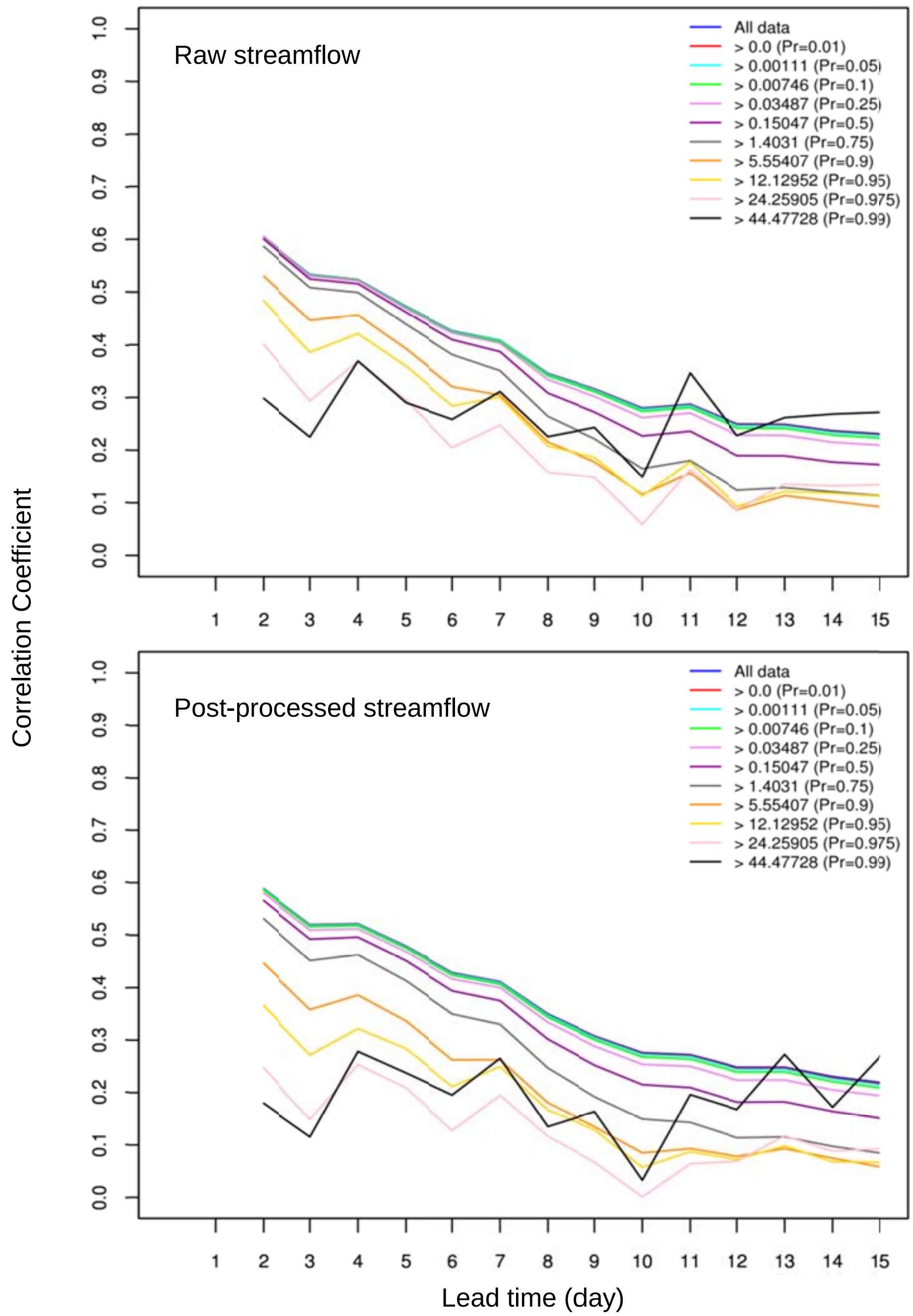


Figure A. 8 Correlation Coefficient of ensemble mean streamflow forecasts and corresponding observed values for the GLLT2

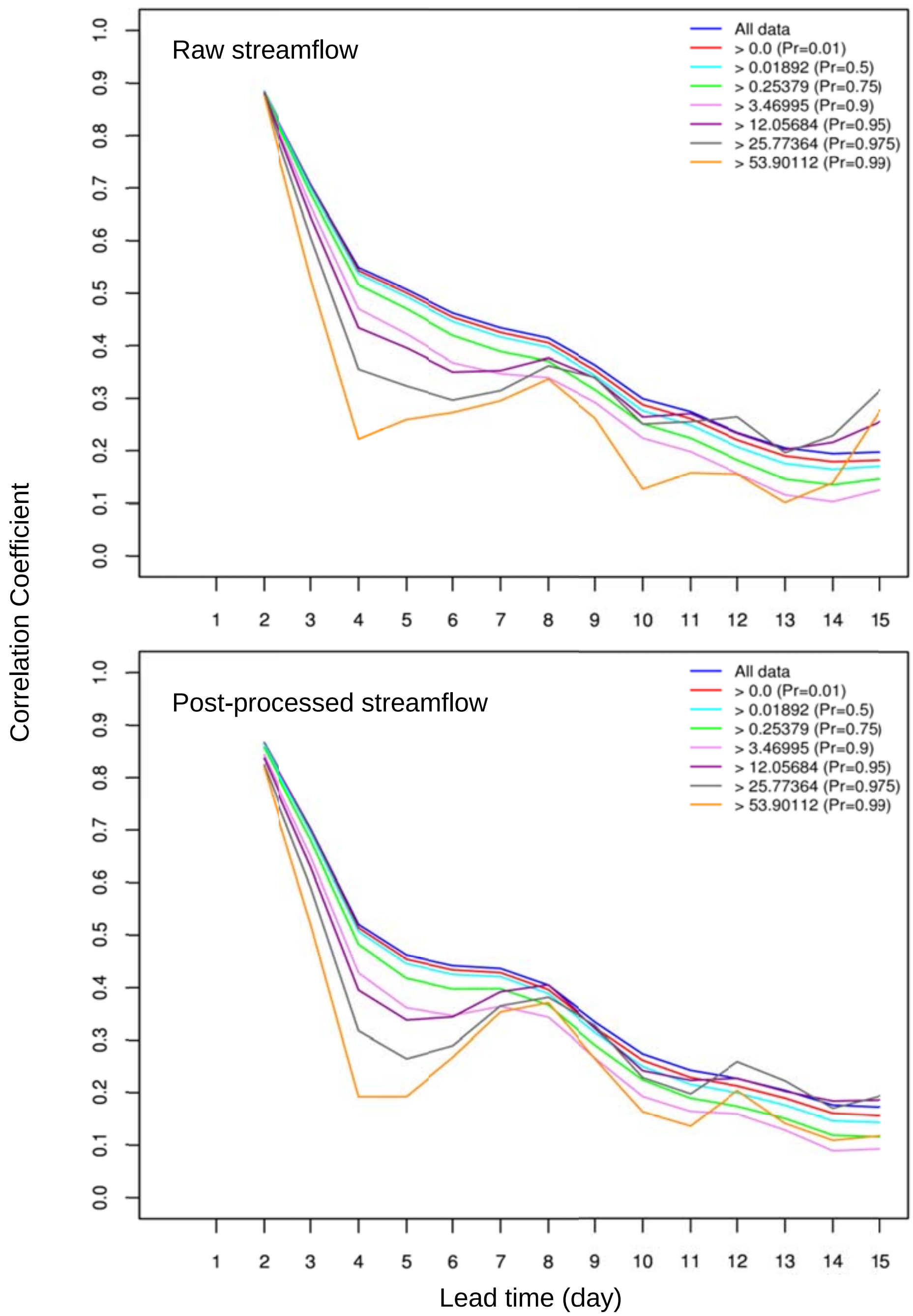


Figure A. 9 Correlation Coefficient of ensemble mean streamflow forecasts and corresponding observed values for the JAKT2

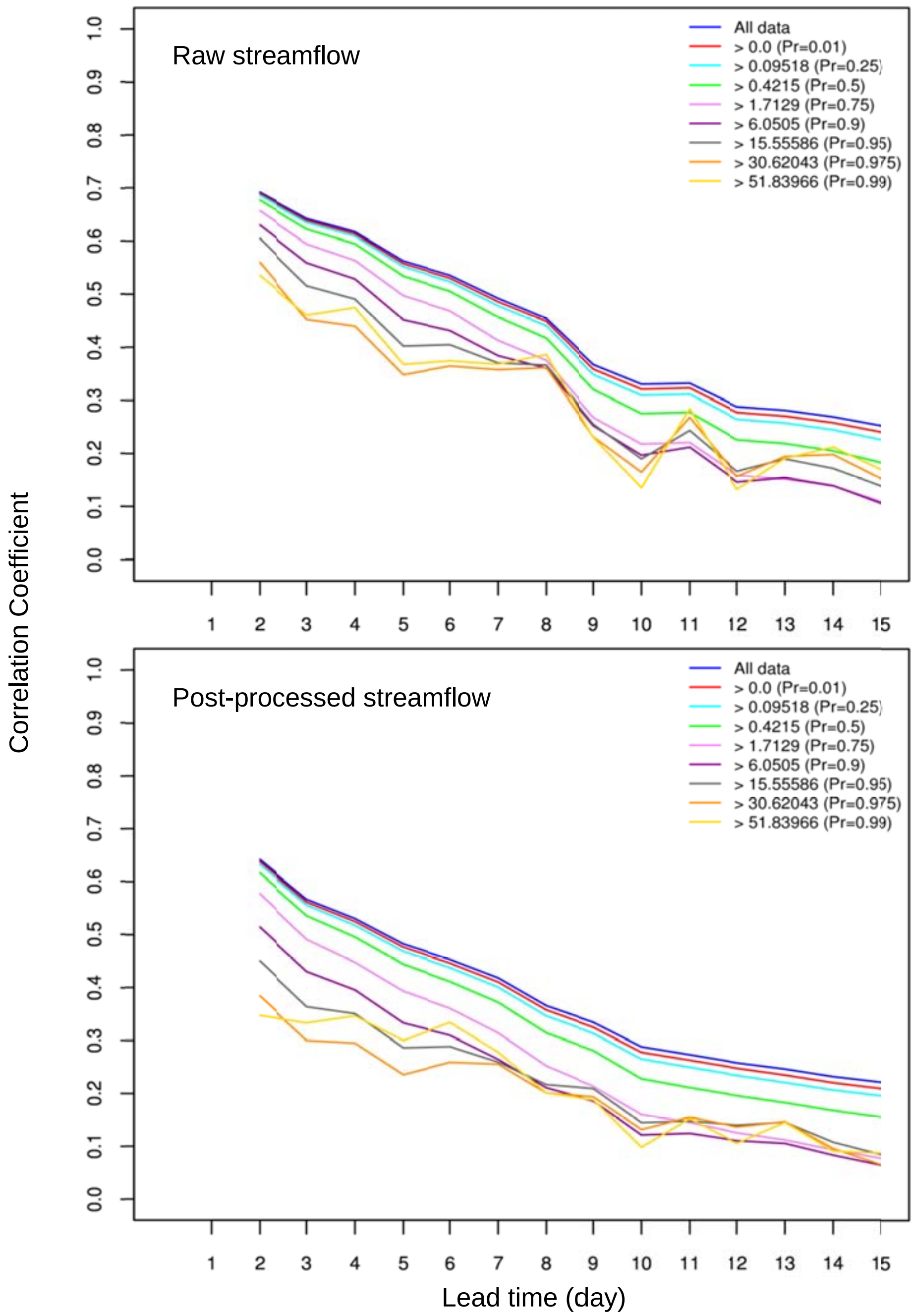


Figure A. 10 Correlation Coefficient of ensemble mean streamflow forecasts and corresponding observed values for the SGET2

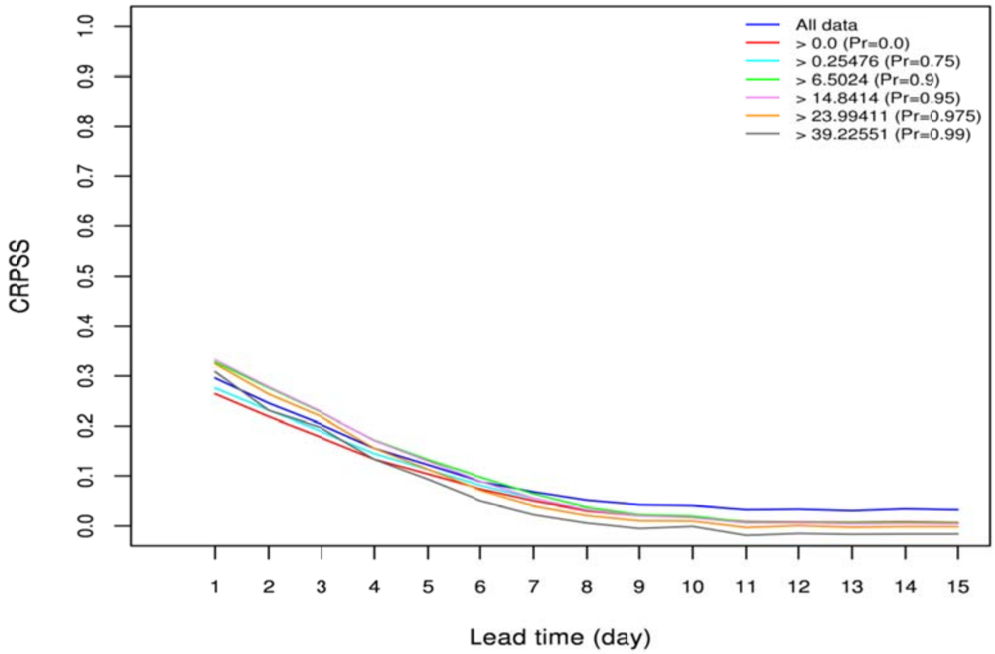


Figure A. 11 CRPSS of daily precipitation ensemble hindcasts for the BRPT2 basin

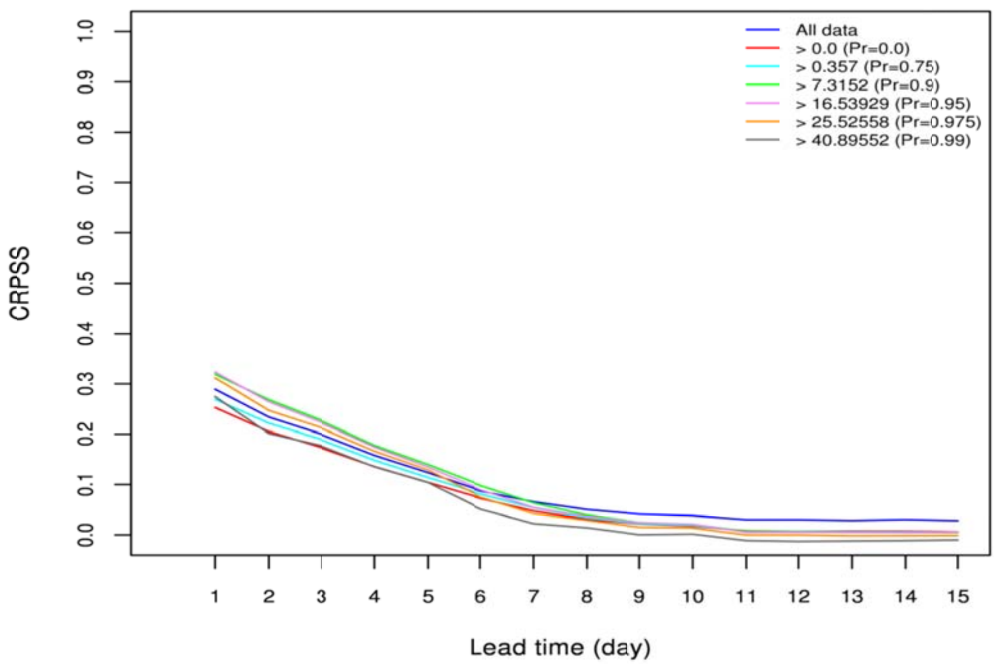


Figure A. 12 CRPSS of daily precipitation ensemble hindcasts for the DCJT2 basin

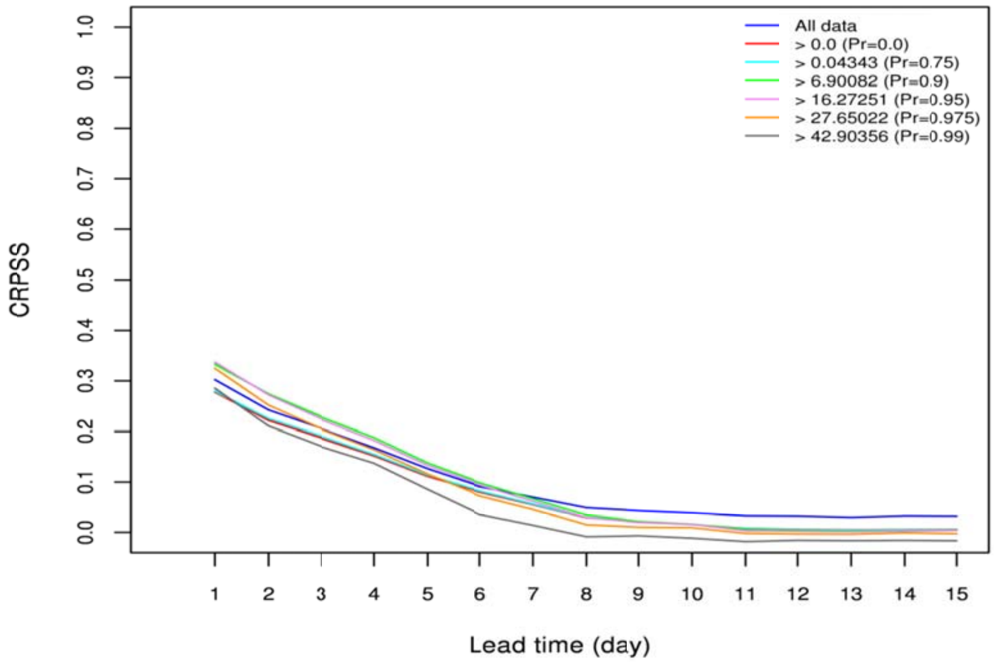


Figure A. 13 CRPSS of daily precipitation ensemble hindcasts for the GLLT2 basin

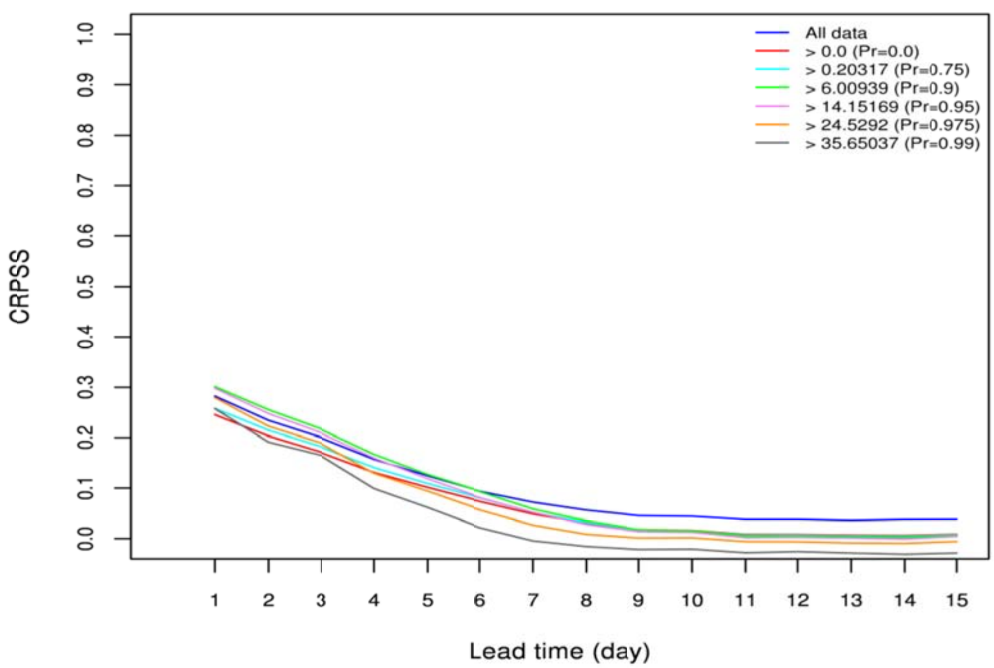


Figure A. 14 CRPSS of daily precipitation ensemble hindcasts for the JAKT2 basin

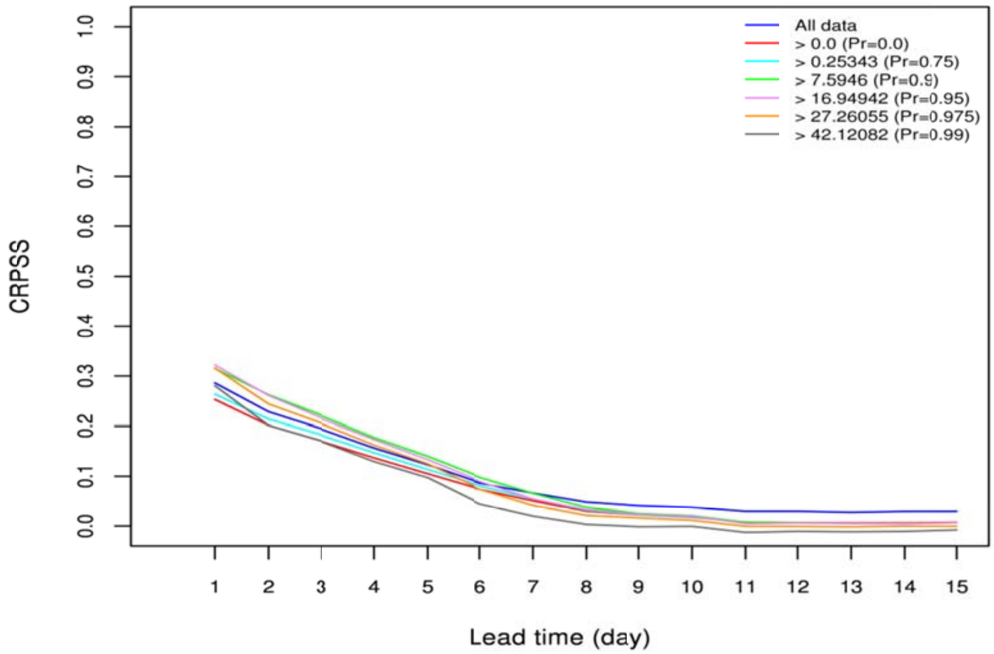


Figure A. 15 CRPSS of daily precipitation ensemble hindcasts for the SGET2 basin

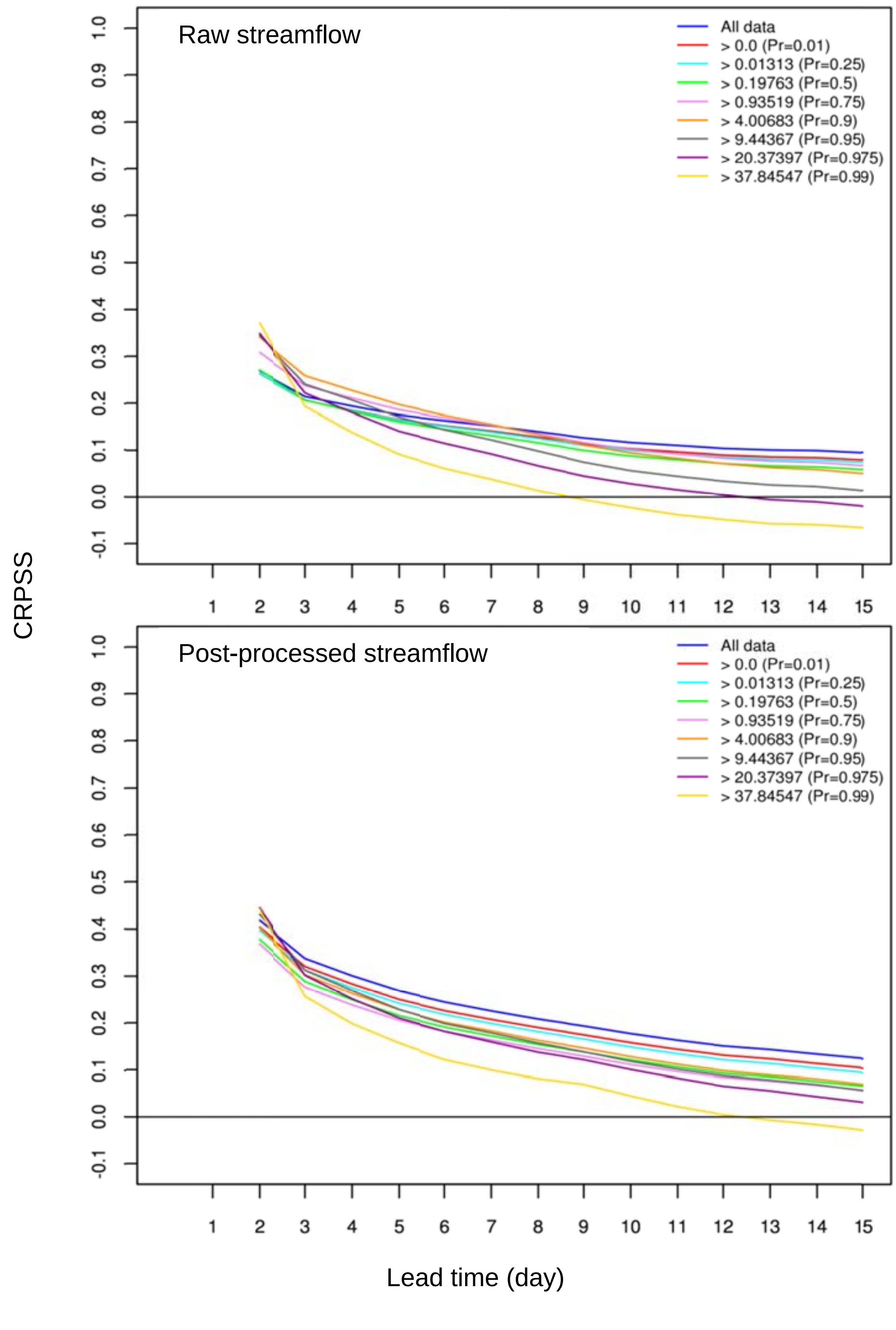


Figure A. 16 CRPSS of daily streamflow ensemble hindcasts for the BRPT2 basin

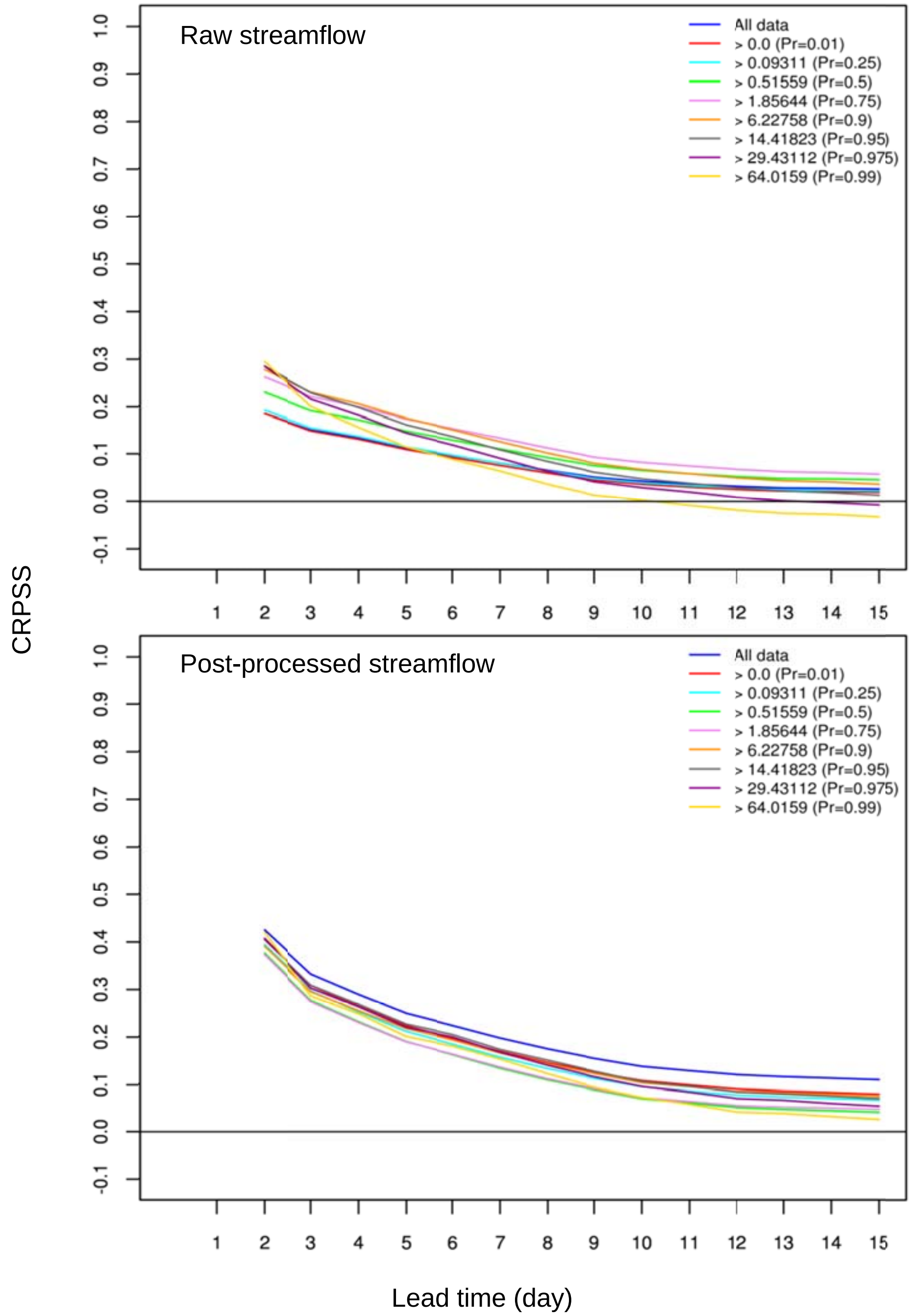


Figure A. 17 CRPSS of daily streamflow ensemble hindcasts for the DCJT2 basin

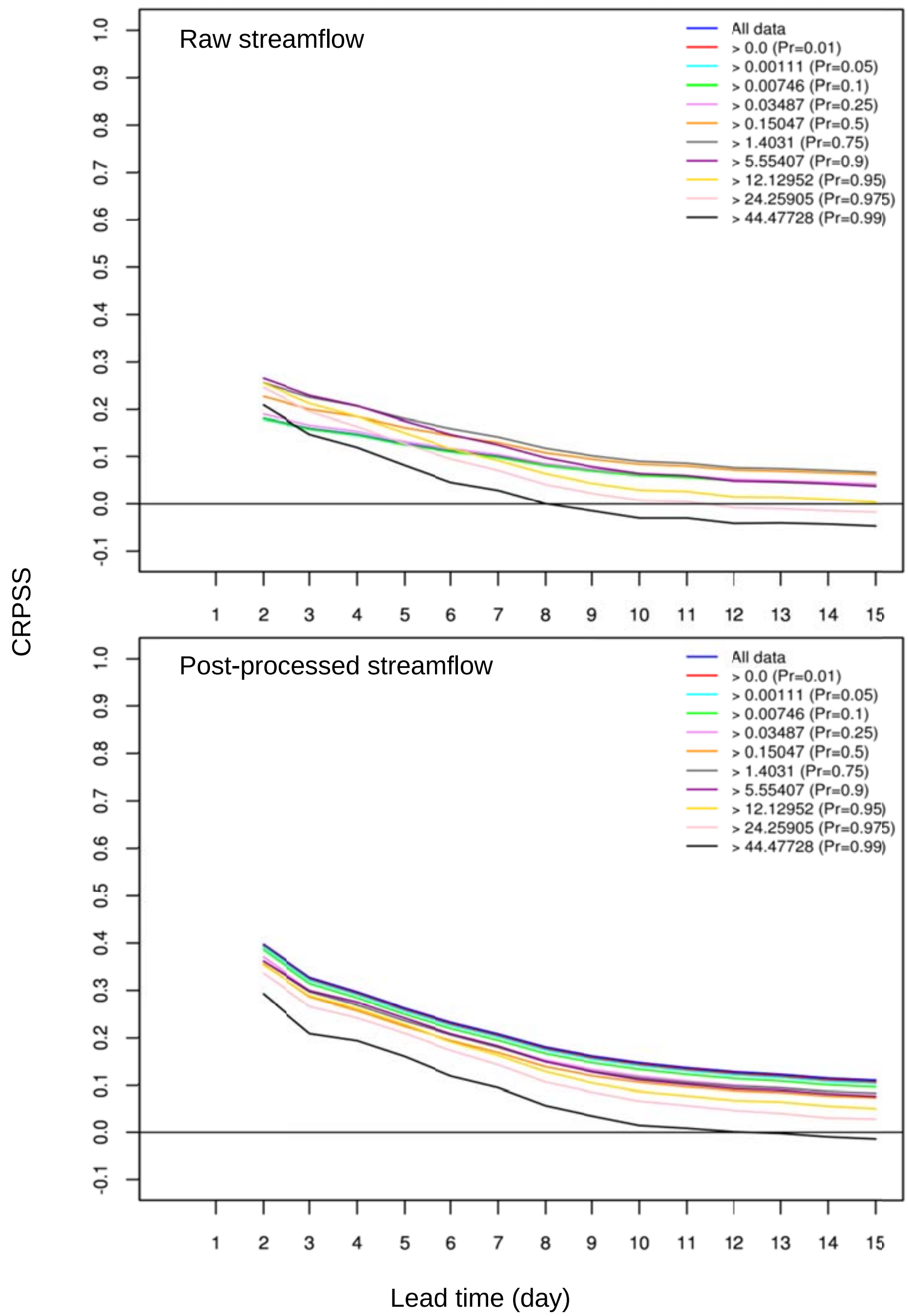


Figure A. 18 CRPSS of daily streamflow ensemble hindcasts for the GLLT2 basin

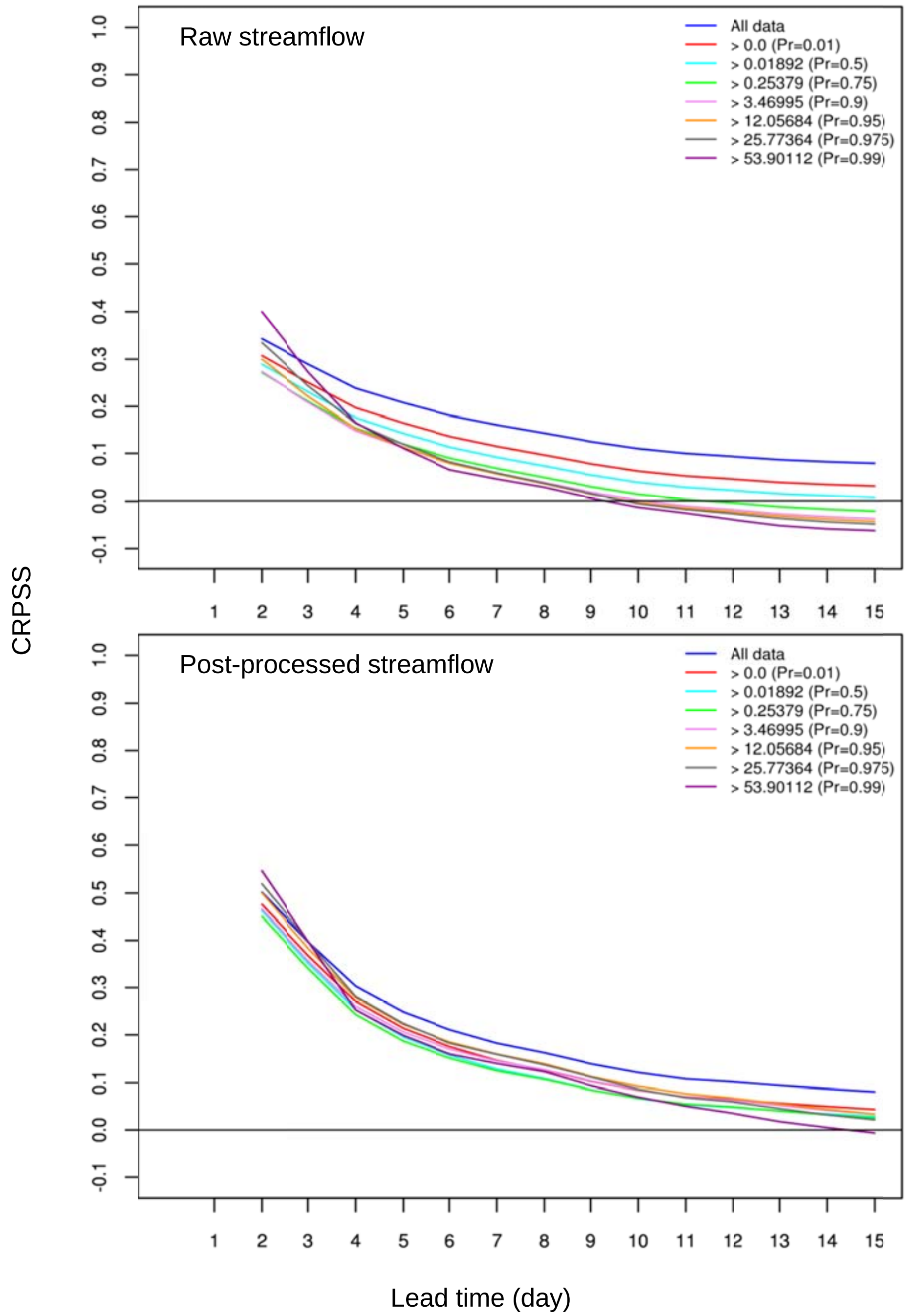


Figure A. 19 CRPSS of daily streamflow ensemble hindcasts for the JAKT2 basin

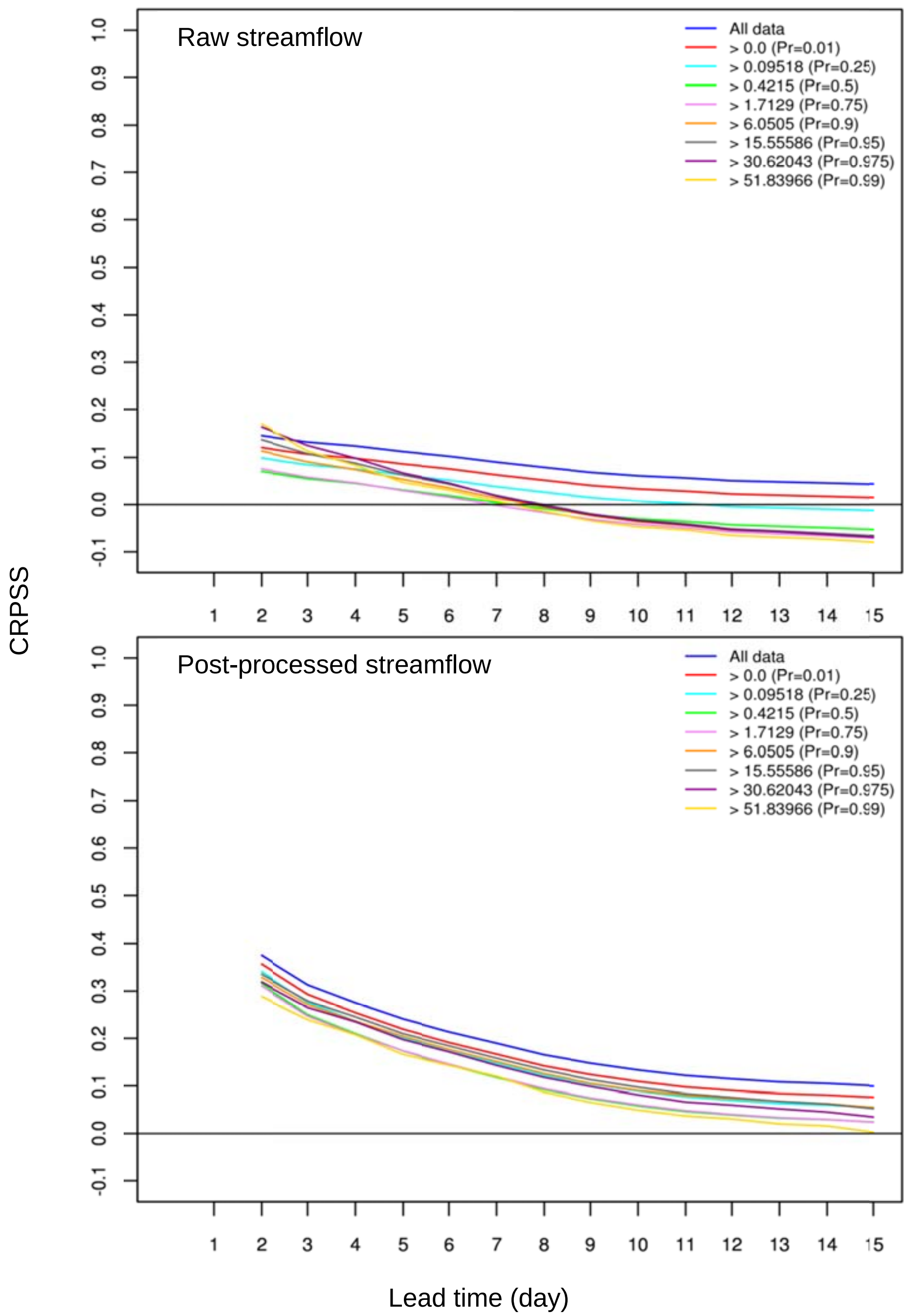


Figure A. 20 CRPSS of daily streamflow ensemble hindcasts for the SGET2 basin

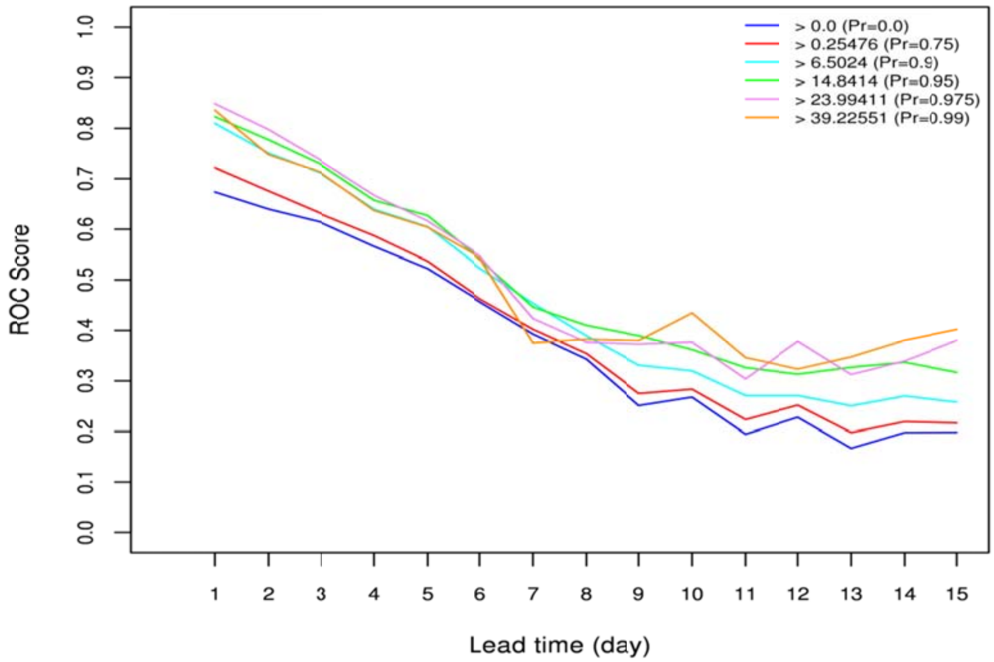


Figure A. 21 ROC Score of daily precipitation ensemble hindcasts for the BRPT2 basin

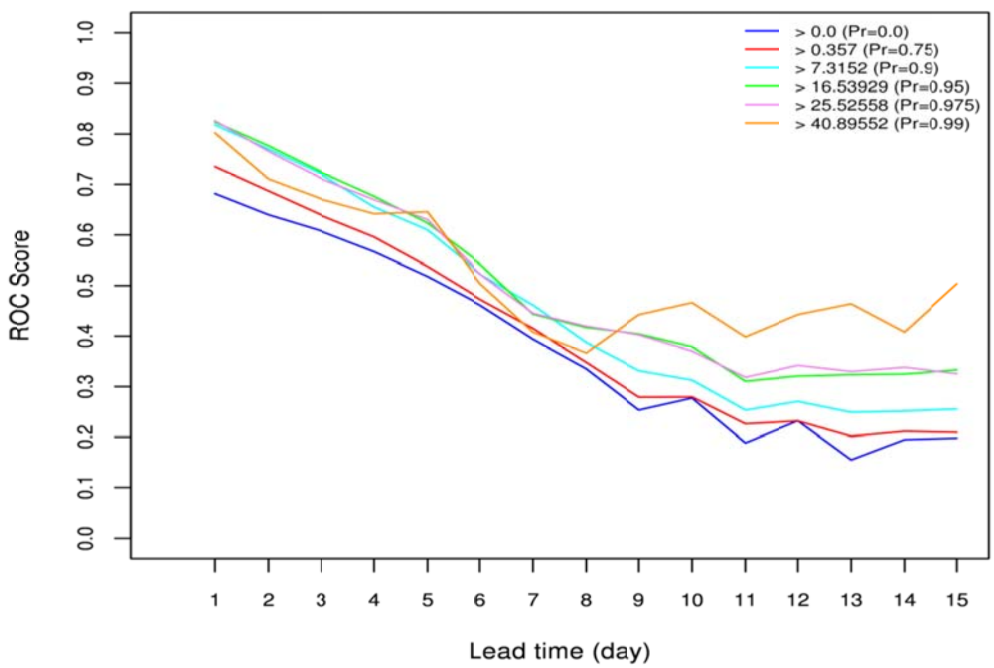


Figure A. 22 ROC Score of daily precipitation ensemble hindcasts for the DCJT2 basin

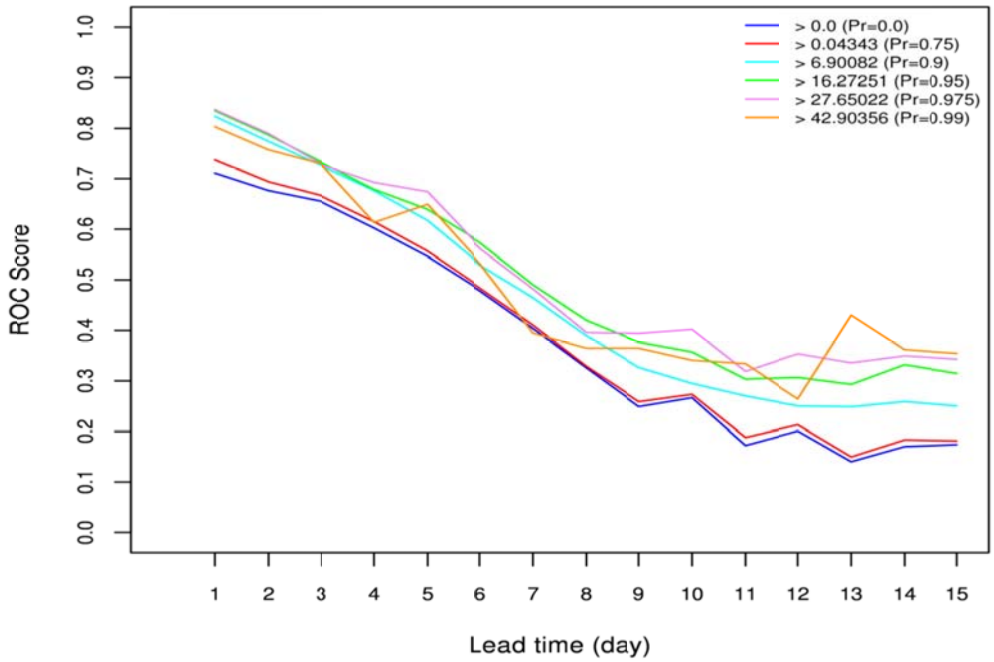


Figure A. 23 ROC Score of daily precipitation ensemble hindcasts for the GLLT2 basin

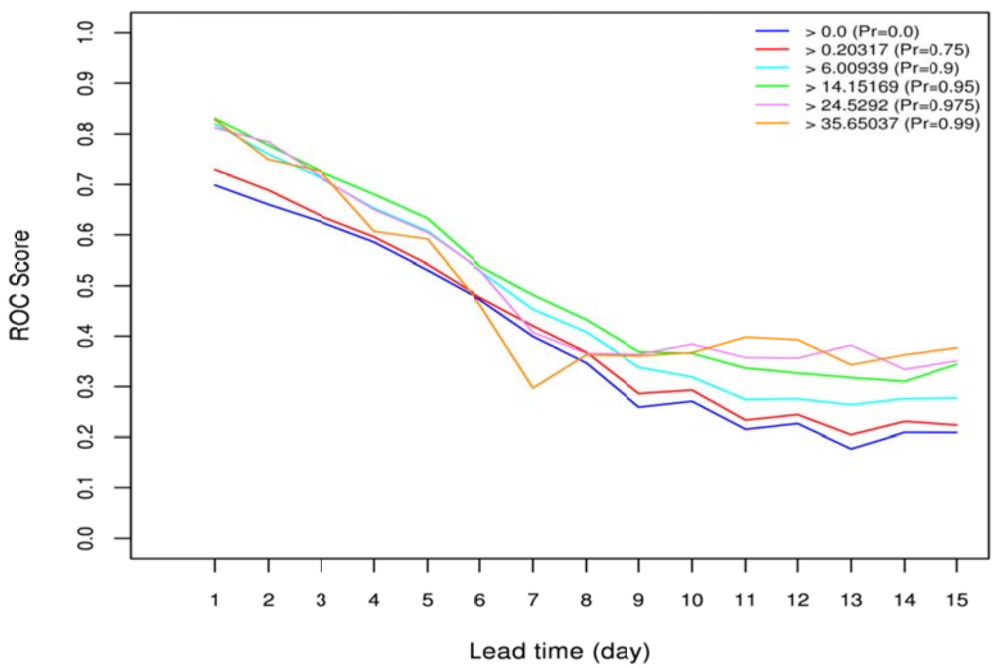


Figure A. 24 ROC Score of daily precipitation ensemble hindcasts for the JAKT2 basin

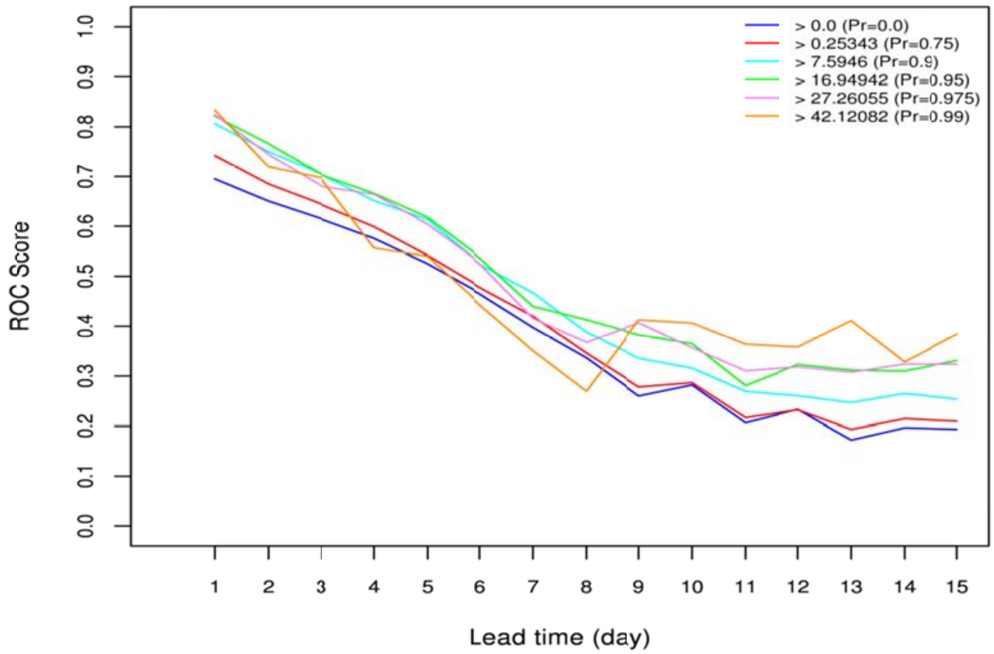


Figure A. 25 ROC Score of daily precipitation ensemble hindcasts for the SGET2 basin

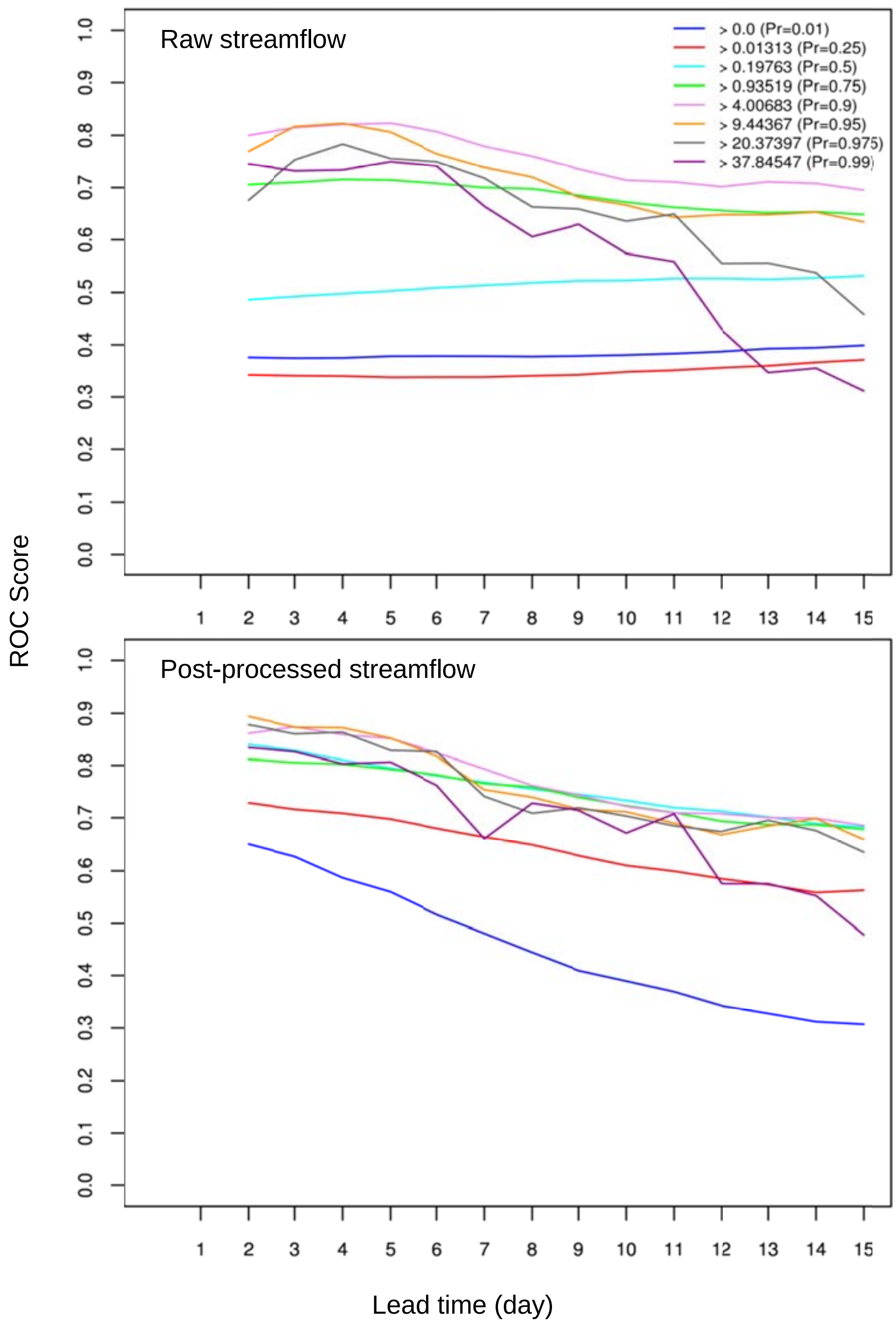


Figure A. 26 ROC Score of daily streamflow ensemble hindcasts for the BRPT2 basin

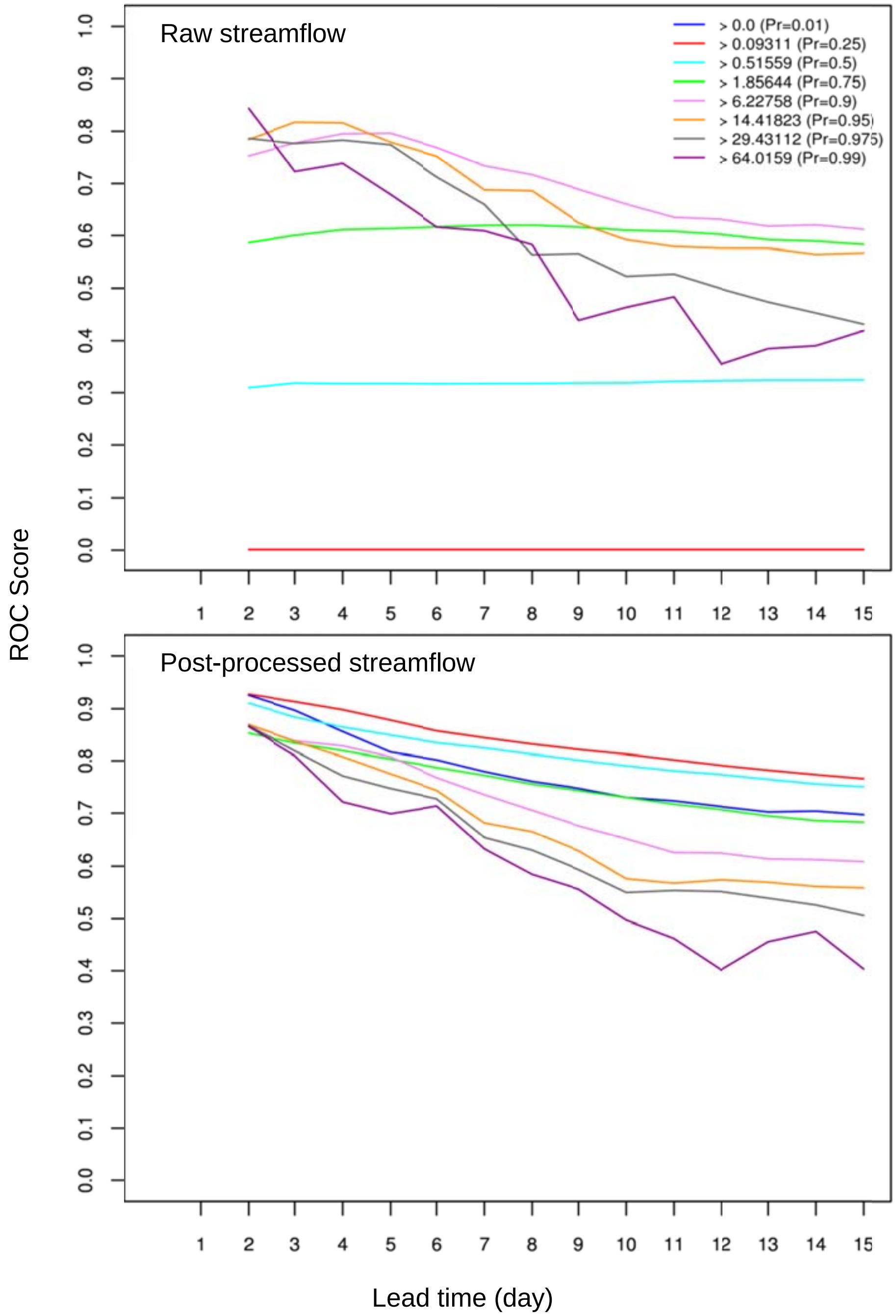


Figure A. 27 ROC Score of daily streamflow ensemble hindcasts for the DCJT2 basin

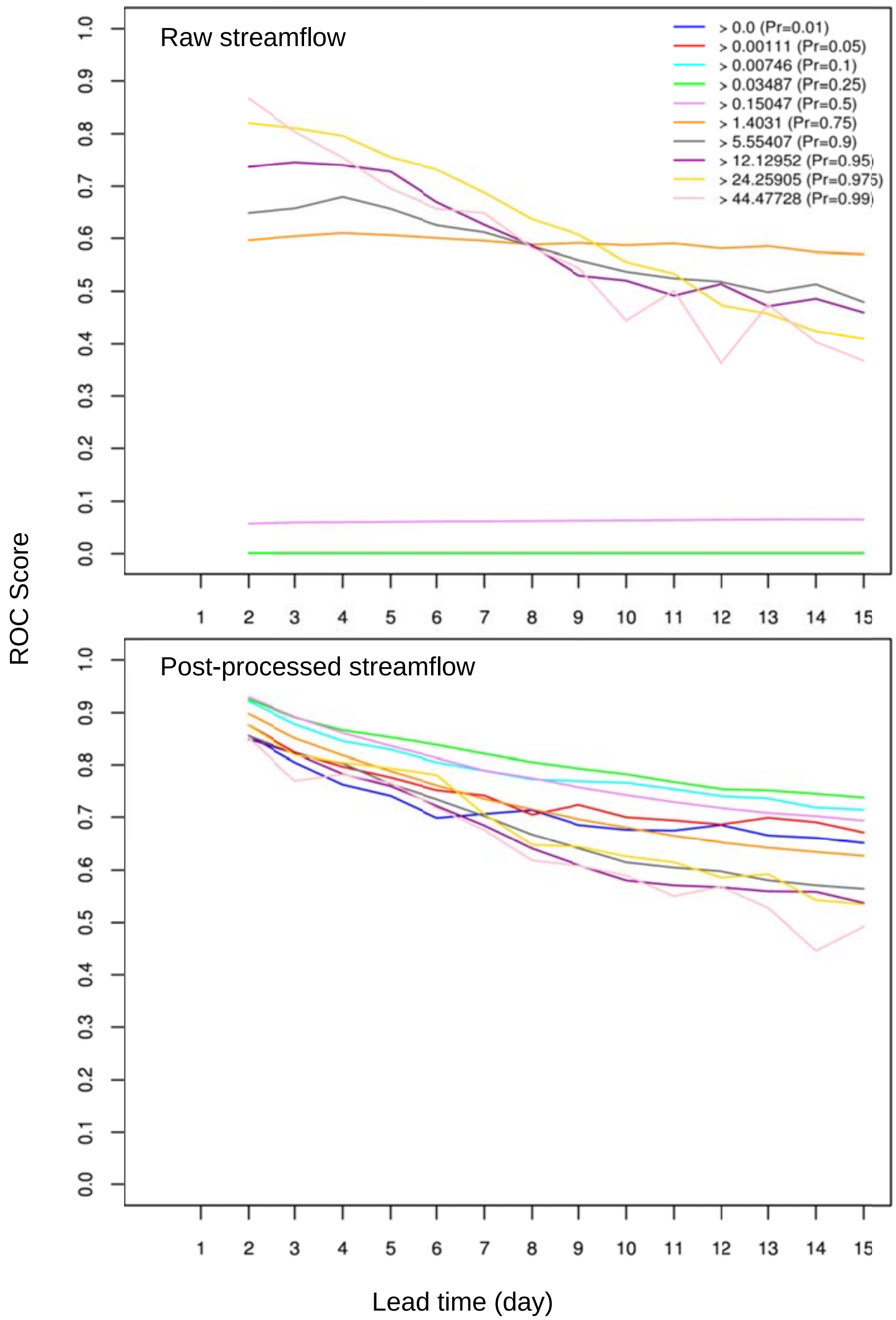


Figure A. 28 ROC Score of daily streamflow ensemble hindcasts for the GLLT2 basin

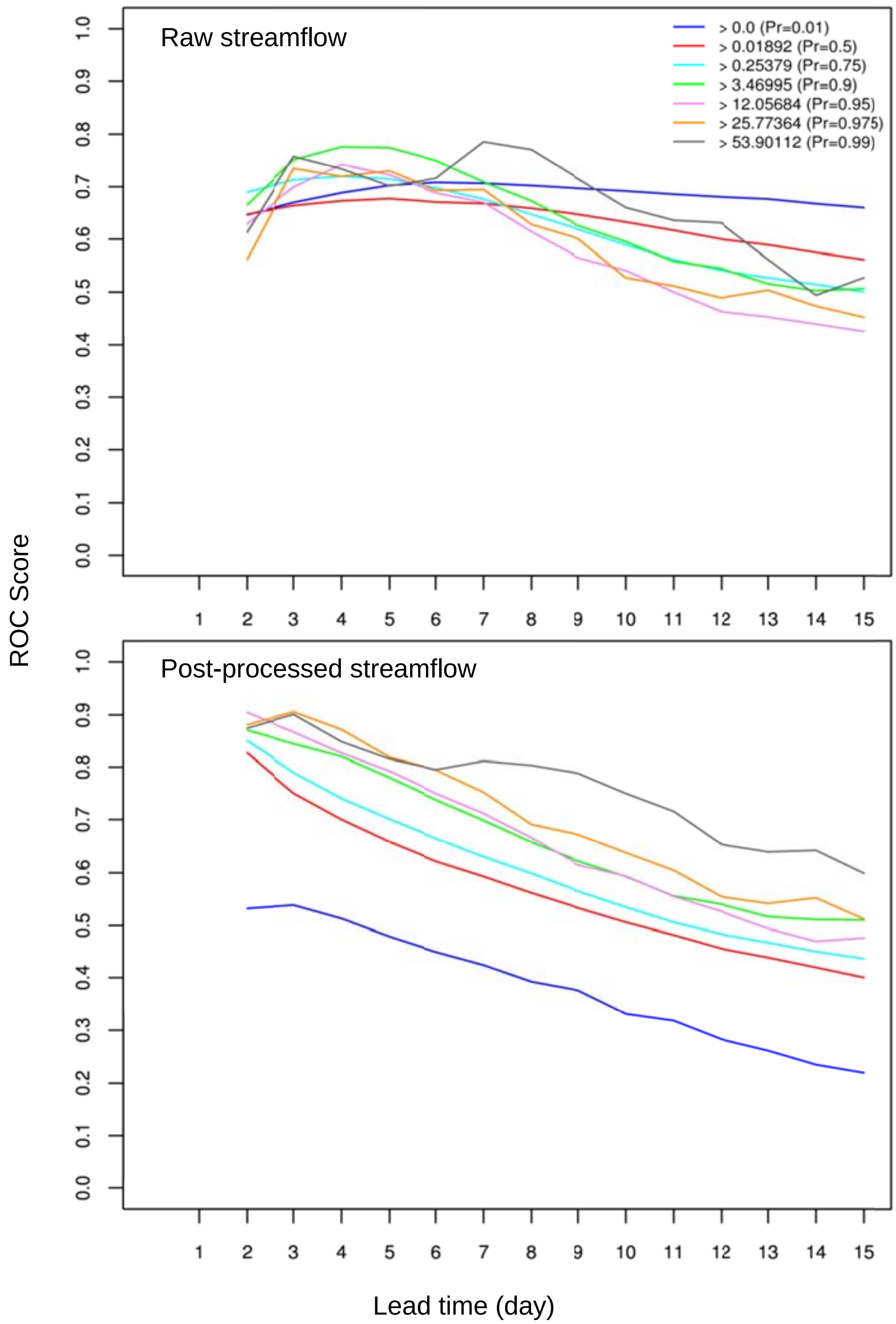


Figure A. 29 ROC Score of daily streamflow ensemble hindcasts for the JAKT2 basin

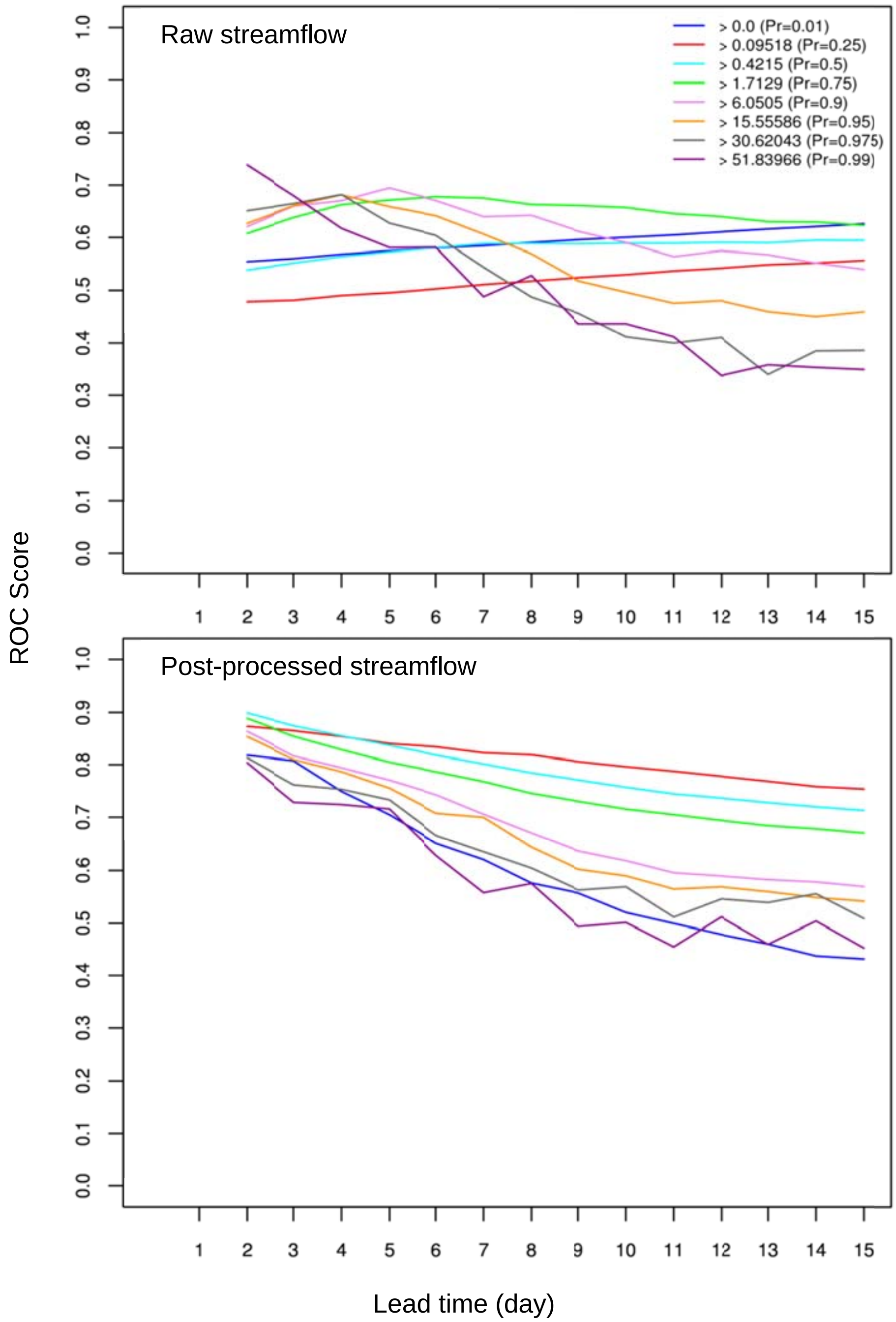


Figure A. 30 ROC Score of daily streamflow ensemble hindcasts for the SGET2 basin

References

Adams, T., & Ostrowski, J. (2010, May). Short lead-time hydrologic ensemble forecasts from numerical weather prediction model ensembles. In *Proc. World Environmental and Water Resources Congress* (pp. 2294-2304).

Bogner, K., Pappenberger, F., & Cloke, H. L. (2012). Technical Note: The normal quantile transformation and its application in a flood forecasting system. *Hydrology and Earth System Sciences*, *16*(4), 1085-1094.

Brown, J. D. (2015). Ensemble Verification System (EVS) user's manual. Retrieved from http://www.nws.noaa.gov/oh/hrl/general/indexdoc.htm.

Brown, J. D., & Heuvelink, G. (2005). Assessing uncertainty propagation through physically based models of soil water flow and solute transport. *Encyclopedia of hydrological sciences*.

Brown, J. D., Demargne, J., Seo, D. J., & Liu, Y. (2010). The Ensemble Verification System (EVS): A software tool for verifying ensemble forecasts of hydrometeorological and hydrologic variables at discrete locations. *Environmental Modelling & Software*, *25*(7), 854-872.

Brown, J. D., He, M., Regonda, S., Wu, L., Lee, H., & Seo, D. J. (2014a). Verification of temperature, precipitation, and streamflow forecasts from the NOAA/NWS Hydrologic Ensemble Forecast Service (HEFS): 2. Streamflow verification. *Journal of Hydrology*, *519*, 2847-2868.

Brown, J. D., Wu, L., He, M., Regonda, S., Lee, H., & Seo, D. J. (2014b). Verification of temperature, precipitation, and streamflow forecasts from the NOAA/NWS Hydrologic Ensemble Forecast Service (HEFS): 1. Experimental design and forcing verification. *Journal of Hydrology*, *519*, 2869-2889.

Burnash, R.J.C. (1995). The NWS River Forecast System - Catchment Model. In V. P. Singh (Ed.), *Computer models of watershed hydrology* (pp. 311-366). Water Resources Publications.

Clark, M., Gangopadhyay, S., Hay, L., Rajagopalan, B., & Wilby, R. (2004). The Schaake shuffle: A method for reconstructing space-time variability in forecasted precipitation and temperature fields. *Journal of Hydrometeorology*, *5*(1), 243-262.

Cloke, H. L., & Pappenberger, F. (2009). Ensemble flood forecasting: a review. *Journal of Hydrology*, *375*(3), 613-626.

Demargne, J., Mullusky, M., Werner, K., Adams, T., Lindsey, S., Schwein, N., ... & Welles, E. (2009). Application of forecast verification science to operational river forecasting in the US National Weather Service. *Bulletin of the American Meteorological Society*, *90*(6), 779-784.

Demargne, J., Wu, L., Regonda, S. K., Brown, J. D., Lee, H., He, M., ... & Zhu, Y. (2014). The science of NOAA's operational hydrologic ensemble forecast service. *Bulletin of the American Meteorological Society*, *95*(1), 79-98.

Deutsch, C. V., & Journel, A. G. (1992). Geostatistical software library and user’s guide. *New York*.

Gneiting, T., Balabdaoui, F., & Raftery, A. E. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *69*(2), 243-268.

Hamill, T. M., Bates, G. T., Whitaker, J. S., Murray, D. R., Fiorino, M., Galarneau Jr, T. J., ... & Lapenta, W. (2013). NOAA's second-generation global medium-range ensemble reforecast dataset. *Bulletin of the American Meteorological Society*, *94*(10), 1553-1565.

Hamill, T. M., Hagedorn, R., & Whitaker, J. S. (2008). Probabilistic forecast calibration using ECMWF and GFS ensemble reforecasts. Part II: Precipitation. *Monthly weather review,* 136(7), 2620-2632.

Hartman, R., Seo, D.-J., Lawrence, B., Shumate, S., Ostrowski, J., Halquist, J., Dietz, C., & Mullusky, M. (2007). The Experimental Ensemble Forecast System (XEFS) Design and Gap Analysis. NOAA/NWS, Silver Spring, MD.

Krzysztofowicz, R. (1999). Bayesian theory of probabilistic forecasting via deterministic hydrologic model. *Water Resources Research,* 35(9), 2739-2750.

MetEd (2015). Introduction to verification of hydrologic forecast. Retrieved from https://www.meted.ucar.edu/

National Oceanic and Atmospheric Administration. (2015, June 2). Flood disaster in Texas and Oklahoma. Retrieved from https://www.climate.gov/news-features/event-tracker/flood-disaster-texas-and-oklahoma

National Weather Service Office of Hydrologic Development (2015a). Ensemble Postprocessor (EnsPost) User’s Manual. Retrieved from http://www.nws.noaa.gov/oh/hrl/general/indexdoc.htm

National Weather Service Office of Hydrologic Development (2015b). Meteorological Ensemble Forecast Processor (MEFP) User’s Manual. Retrieved from http://www.nws.noaa.gov/oh/hrl/general/indexdoc.htm

Nester, T., Komma, J., Viglione, A., & Blöschl, G. (2012). Flood forecast errors and ensemble spread—A case study. *Water Resources Research,* 48(10).

Regonda, S. K., Seo, D. J., Lawrence, B., Brown, J. D., & Demargne, J. (2013). Short-term ensemble streamflow forecasting using operationally-produced single-valued streamflow forecasts–A Hydrologic Model Output Statistics (HMOS) approach. *Journal of Hydrology,* 497, 80-96.

Saharia, M. (2013). Ensemble Streamflow Forecasting For The Upper Trinity River Basin In Texas.

Schaake, J., Demargne, J., Hartman, R., Mullusky, M., Welles, E., Wu, L., ... & Seo, D. J. (2007). Precipitation and temperature ensemble forecasts from single-value forecasts. *Hydrology and Earth System Sciences Discussions Discussions,* 4(2), 655-717.

Seo, D. J., Demargne, J., Wu, L., Liu, Y., Brown, J. D., Regonda, S., & Lee, H. (2010, June). Hydrologic ensemble prediction for risk-based water resources management and hazard mitigation. In *Joint Federal Interagency Conference on Sedimentation and Hydrologic Modelling (JFIC2010)* (Vol. 27).

Seo, D. J., Herr, H. D., & Schaake, J. C. (2006). A statistical post-processor for accounting of hydrologic uncertainty in short-range ensemble streamflow prediction. *Hydrology and Earth System Sciences Discussions,* 3(4), 1987-2035.

Texas Water Development Board (TWDB) (http://waterdatafortexas.org), Accessed 2015

WGRFC, QPF http://www.srh.noaa.gov/wgrfc/?n=wgrfcqpfpage, Accessed in 2015

Wu, L., Seo, D. J., Demargne, J., Brown, J. D., Cong, S., & Schaake, J. (2011). Generation of ensemble precipitation forecast from single-valued quantitative precipitation forecast for hydrologic ensemble prediction. *Journal of hydrology,* 399(3), 281-298.

Yuan, X., Wood, E. F., & Liang, M. (2014). Integrating weather and climate prediction: Toward seamless hydrologic forecasting*. Geophysical Research Letters,* 41(16), 5891-5896.

Biographical Information

Hossein Sadeghi was born in Shiraz, Fars, Iran, in 1976. He received his Bachelor's degree in Civil Engineering from Azad University, Fars, Estahban, Iran, in 1999. He received his Master's degree in Civil Engineering-Hydraulic Structure from Azad University, Kerman, Iran, in 2003. He worked as a hydraulic engineering expert and resident inspector for the Taleghan Dam and Power Plant Project at Lar Consulting Engineers Company from late 2003 to early 2005. He has experience in hydraulic design and also in construction for which he worked as a site manager in residential steel and concreate building projects. He was a full-time faculty in Civil Engineering Department, Azad University, Sepidan Branch, Iran, from 2007 to early 2013, where he was also the head of the Civil Engineering Department from 2012 to early 2013.