The Use of Central Tendency Measures from an Operational Short Lead-time Hydrologic Ensemble Forecast System for Real-time Forecasts

Thomas E Adams, III

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Civil Engineering

Randel L. Dymond, Chair
Kevin J. McGuire
Andrew W. Ellis
Mark A. Widdowson

May 8, 2018
Blacksburg, Virginia

Keywords: Hydrology, Forecasting, Precipitation, Uncertainty, Prediction, Modeling

The Use of Central Tendency Measures from an Operational Short Lead-time Hydrologic Ensemble Forecast System for Real-time Forecasts

Thomas E Adams, III

(ABSTRACT)

A principal factor contributing to hydrologic prediction uncertainty is modeling error introduced by the *measurement* and *prediction* of precipitation. The research presented demonstrates the necessity for using probabilistic methods to quantify hydrologic forecast uncertainty due to the magnitude of precipitation errors. Significant improvements have been made in precipitation estimation that have lead to greatly improved hydrologic simulations. However, advancements in the prediction of future precipitation have been marginal. This research shows that gains in forecasted precipitation accuracy have not significantly improved hydrologic forecasting accuracy. The use of forecasted precipitation, referred to as quantitative precipitation forecast (QPF), in hydrologic forecasting remains commonplace. Non-zero QPF is shown to improve hydrologic forecasts, but QPF duration should be limited to 6 to 12 hours for flood forecasting, particularly for fast responding watersheds. Probabilistic hydrologic forecasting captures hydrologic forecast error introduced by QPF for all forecast durations. However, public acceptance of probabilistic hydrologic forecasts is problematic. Central tendency measures from a probabilistic hydrologic forecast, such as the ensemble median or mean, have the *appearance* of a single-valued deterministic forecast. The research presented shows that hydrologic ensemble median and mean forecasts of river stage have smaller forecast errors than current operational methods with forecast lead-time beginning at 36-hours for fast response basins. Overall, hydrologic ensemble median and mean forecasts display smaller forecast error than current operational forecasts.

The Use of Central Tendency Measures from an Operational Short Lead-time Hydrologic Ensemble Forecast System for Real-time Forecasts

Thomas E Adams, III

(GENERAL AUDIENCE ABSTRACT)

Flood forecasting is uncertain, in part, because of errors in measuring precipitation and predicting the location and amount of precipitation accumulation in the future. Because of this, the public and other end-users of flood forecasts should understand the uncertainties inherent in forecasts. But, there is reluctance by many to accept forecasts that explicitly convey flood forecast uncertainty, such as, "there is a 67% chance your house will be flooded". Instead, most prefer "your house will not be flooded" or something like "flood levels will reach 0.5 feet in your house". We hope the latter does not happen, but due to forecast uncertainties, explicit statements such as "flood levels will reach 0.5 feet in your house" will be wrong. If by chance, flood levels do *exactly* reach 0.5 feet, that will have been a *lucky* forecast, very likely involving some skill, but the flood level could have reached 0.43 or 0.72 feet as well. This research presents a flood forecasting method that improves on traditional methods by directly incorporating uncertainty information into flood forecasts that still appear like forecasts people are familiar and comfortable with and understandable by them.

Acknowledgements

I must recognize the most influential men in my life: my father, Lt. Col. Thomas E. Adams, Jr. (U.S. Army, Retired), my father-in-law, CMDR, Dr. James F. Phelan, PhD (U.S. Navy, Retired), Dr. Joseph C. Pitt, PhD who all helped me to become a better person; any and all failings are my own. I am very indebted to my PhD Advisor, Dr. Randel L. Dymond, PhD for helping me through the past several years to finally bring my PhD to a conclusion. I am grateful to my other PhD committee members for their time and feedback.

I am, of course, very grateful to my wife, Dr. Anne L. Phelan-Adams, MD for sticking with me through life.

It is with honor that I dedicate this work to Dr. G.V. Loganathan, PhD – friend, colleague, and teacher, whose life was cut short horrifically...

From Homer *Il.* XIV, 200 [101]:

> For I am going to see the limits of fertile earth, Okeanos begetter of gods and mother Tethys...

> Data do not give up their secrets easily. They must be tortured to confess.

Jeff Hopper, Bell Labs...

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Nature of the Problem

All forecasts are uncertain. Hydrologic forecasts are uncertain largely because of errors in the measurement and prediction of hydrologic model forcings, such as temperature and precipitation. Advances have been made over recent decades in the measurement and prediction of precipitation that have been adopted operationally for flood prediction and water resources forecasting by U.S. *National Oceanic and Atmospheric Administration (NOAA)*, *National Weather Service (NWS)* River Forecast Centers (RFCs). But hydrologic forecast accuracy gains have not been fully quantified with the adoption of these scientific advances. This dissertation quantifies error reduction in hydrologic forecasting derived from advancements in radar based precipitation measurement and precipitation forecasting at the NOAA/NWS *Ohio River Forecast Center (OHRFC)*. However, despite improvements in the measurement and prediction of hydrometeorological variables, considerable forecast uncertainty remains, yet *official* NWS river stage forecasts do not currently convey forecast uncertainty. Consequently, this study further explores the use of hydrologic ensemble median and mean river

stage forecasts as a possible alternative to current single-valued deterministic river stage forecasts to both reduce forecast uncertainty and to further motivate adoption of the use of probabilistic hydrologic forecasts by the public and decision makers.

1.2 Dissertation Objectives

The central research objective of this dissertation is to demonstrate the necessity for using probabilistic hydrologic forecasting, using hydrologic ensembles, in place of current methods that rely on the use of single-valued deterministic forecast of precipitation, known as *Quantitative Precipitation Forecast (QPF)*. Due to resistance by the general public and many decision-makers to accept probabilistic hydrologic forecasts, the use of hydrologic ensemble median and mean forecasts is explored as a mechanism to reduce hydrologic forecast uncertainty and to explicitly include the notion of uncertainty in hydrologic forecasts in terms of *expectation* or "best estimate". The manuscript is structured to:

- quantify, in relative terms, gains achieved in the reduction of hydrologic prediction error due to advancements in the measurement of precipitation, referred to as *Quantitative Precipitation Estimate (QPE)* and prediction of precipitation, QPF, over the past ∼50 years;

- quantify hydrologic prediction error using deterministic single-valued QPF, to determine (1) if the use of *non-zero* QPF is warranted in hydrologic forecasting, because of reduced forecast error relative to *zero-QPF* forecasts, and (2) if the answer is that *non-zero* QPF does produce hydrologic forecasts with smaller error, what hydrologic prediction error structures are incurred with the use of longer QPF periods (6-, 12-, 24-,. . . , 72-hours,. . .) and;

- quantify hydrologic ensemble median and mean forecasts error relative to current operational methods that use single-valued deterministic QPF.

These issues will be addressed in 3 chapters:

1. Hydrometeorological Forcing Errors for a Real-time Flood Forecast System in the Ohio River Valley, USA;

2. The Effect of QPF on Real-time Deterministic Hydrologic Forecast Uncertainty;

3. The Use of Central Tendency Measures from an Operational Short Lead-time Hydrologic Ensemble Forecast System for Real-time Forecasts.

1.3 Research Approach

The first research objective is achieved by quantifying radar derived precipitation estimation errors relative to a widely accepted historical precipitation database over an approximate 20-year period. A hydrologic simulation experiment is run to quantify the hydrologic impact of precipitation estimation improvements. QPF verification results from two sources are reported to show forecast improvements since the 1970s. A hydrologic *monte carlo* simulation experiment is conducted to assess the impact of the improvements on hydrologic forecast error.

The second research objective is attained by using two real-time hydrologic forecast experiments, the first to assess the magnitude of hydrologic prediction error with *zero-QPF* and *non-zero QPF*. The second experiment investigates hydrologic prediction error incurred due to varying ranges of QPF duration.

The third research objective is met by utilizing hydrologic ensemble forecasts from the *Meteorological Model-based Ensemble Forecast System (MMEFS)* [1] methodology to estimate hydrologic prediction uncertainty at numerous forecast point locations in the NOAA/NWS OHRFC area of responsibility. The research methodology makes use of the U.S. NOAA/NWS, National Centers for Environmental Prediction (NCEP) *North American Ensemble Forecast System (NAEFS)* numerical weather prediction (NWP) model output of precipitation and temperature as gridded hydrometeorological field forcings to a physically-based conceptual hydrologic model which generates, as output, ensemble hydrological time series. The resulting hydrologic time series ensembles are analyzed to provide probabilistic forecasts of peak flow and stage for short lead-time events out to 168 hours. Probabilistic verification measures are used to evaluate the reasonableness of the MMEFS hydrologic ensemble forecasts.

Chapter 2

Literature Review

Real-time, operational hydrologic forecasting is needed throughout the world for flood prediction and is necessary in many water resources applications. A key requirement, especially for flood forecasting, is the delivery of accurate and timely flood warnings/alerts to the general public and decision makers, thus providing the opportunity to initiate preventative flood defense measures or for possible emergency response. Currently, hydrologic forecasts typically take the form of single-valued deterministic river stage predictions that are derived from observed and forecasted temperature and precipitation as input to a hydrologic modeling system, as depicted in Figure 2.1. However, operational experience and significant research (Welles et al. [181]; Demargne et al. [56]; and Demargne et al. [57]) have demonstrated that errors resulting from *observational measurement* and *prediction* of air temperature and the magnitude and location of precipitation (or other hydrometeorological variables, such as relative humidity, wind speed and direction, etc.) can produce significant hydrologic prediction/forecasting errors, which can lead to erroneous alerts and warnings (*False Alarms*) or the failure to issue alerts and warnings. The *World Meteorological Organization (WMO)* statement on the *Scientific Basis for and Limitations of River Discharge and Stage Fore-*

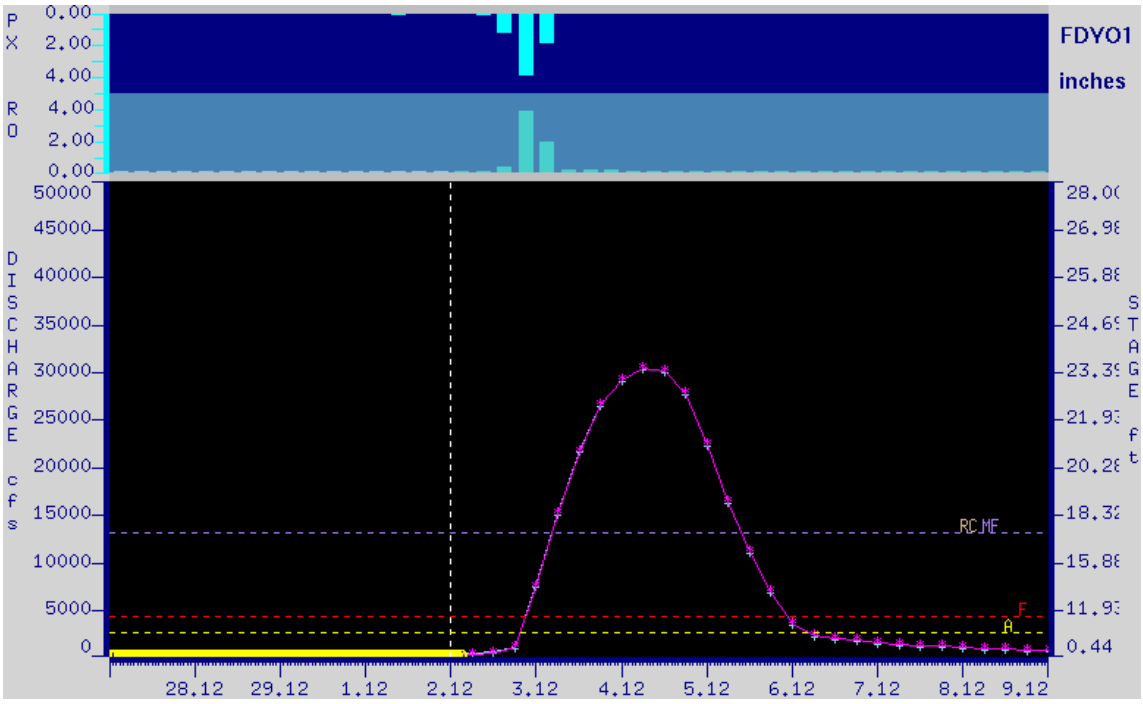

Figure 2.1: Example Ohio River Forecast Center (OHRFC) hydrologic forecast hydrograph for Findlay, OH (with the location identifier, FDYO1) showing Quantitative Precipitation Forecast (QPF) as downward directed cyan colored bars to the right of the current time (vertical white dashed line). The graphic was generated by the *NWS River Forecast System (NWSRFS) Interactive Forecast Program (IFP)* for the period 28 February 2008 to 9 March 2008. The forecast exceeds the *Major Flood* level (dashed purple line) and top of the forecast point *rating curve* by over 5 Feet.

casting [191] underscores the factors controlling hydrologic forecast uncertainty. Figure 2.2 shows *Root Mean Square Error (RMSE)* (see Equation 5.3), comparing above flood to below flood forecasts, for 13 National Weather Service (NWS) River Forecast Centers (RFCs), shown in Figure 2.3. These forecasts include the use of Quantitative Precipitation Forecast (QPF) – that is, forecasted precipitation, for the period 2002–2015. Two conclusions can be drawn:

1. Forecast errors increase with longer lead times;

2. Forecast errors are significantly greater for *above flood* forecasts than for *below flood*

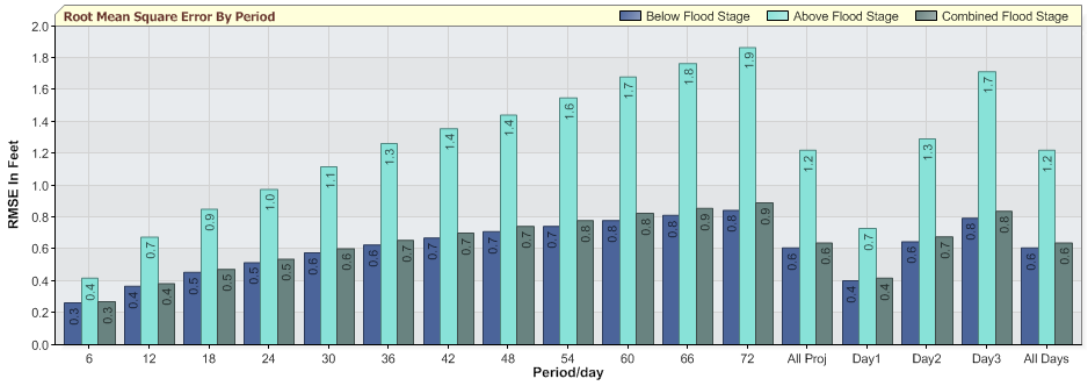

Figure 2.2: NWS forecast verification for 13 River Forecast Centers (RFCs), showing *Root Mean Square Error (RMSE)* by lead time, 2002 – 2015, and comparing above flood forecasts to below flood forecasts.

forecasts.

Item number 2 is particularly important since the NWS mission for hydrologic forecasting is focused on the "protection of lives and property" [134]. It should be noted that other verification measures, such as *mean error* (Equation 5.1) and *mean absolute error* (Equation 5.2) show the same trend as Figure 2.2 shows for *root mean square error*.

Forecast errors also arise from model parameter estimation uncertainty, uncertain model state initial conditions, the inability of deterministic hydrologic models to accurately capture the physical processes of the hydrologic cycle, and unwitting error introduced from manual modeling adjustments during forecaster intervention, etc. All sources of model and input error are either additive or multiplicative (McMillan et al. [121]; Salamon and Feyen [155]) and lead to uncertainty in hydrologic forecasts to varying degrees. There are, additionally, model structural uncertainties [109]. Unfortunately, the relative importance of the sources of hydrologic forecast error has not been studied in enough detail and has not been quantified sufficiently to rank sources of hydrologic forecast error in order of their relative contributions to overall forecast uncertainty. However, widespread agreement exists

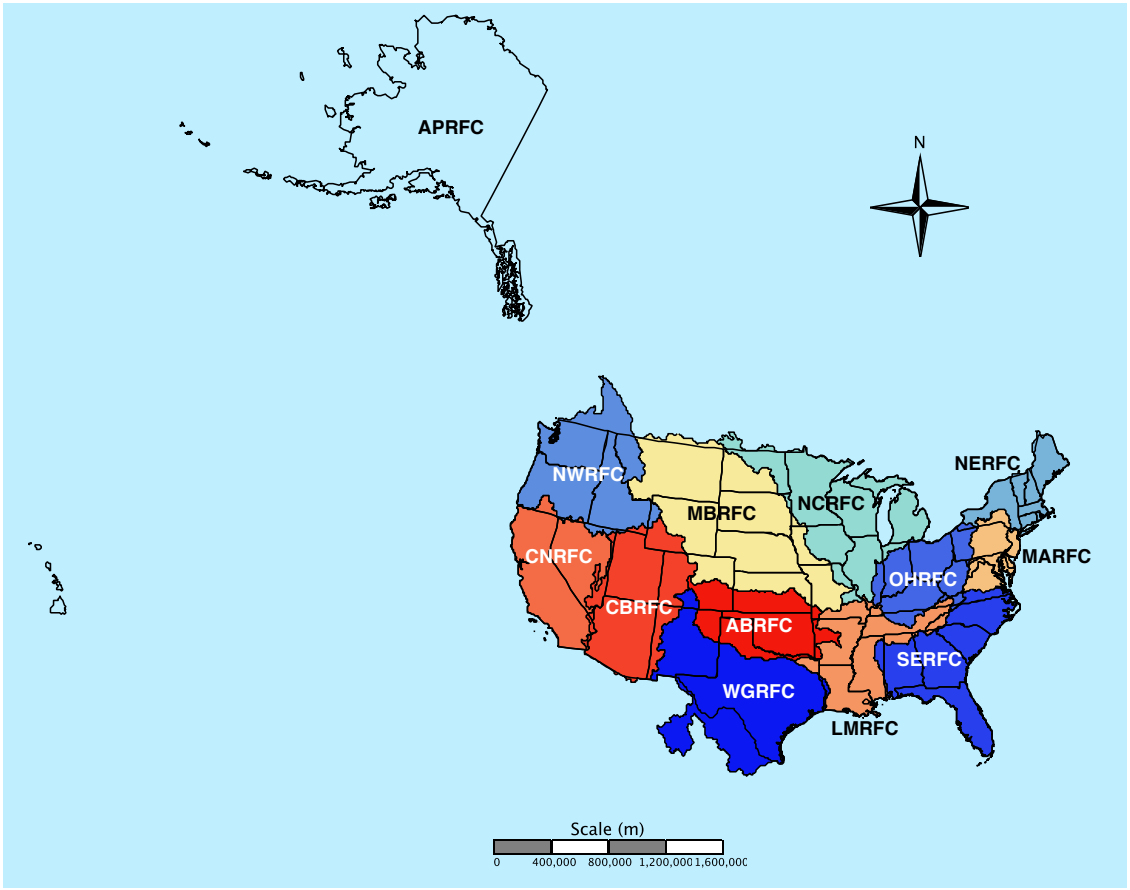

Figure 2.3: The NWS 13 River Forecast Centers (RFCs) – Alaska/Pacific RFC (APRFC), Arkansas-Red RFC (ABRFC), Colorado Basin RFC (CBRFC), California-Nevada RFC(CNRFC), Lower-Mississippi RFC (LMRFC)Middle Atlantic RFC (MARFC), Missouri Basin RFC (MBRFC), North Central RFC (NCRFC), Northwest RFC (NWRFC), Ohio RFC (OHRFC), Southeast RFC (SERFC), and West Gulf RFC (WGRFC). Please note that several RFC boundaries extend beyond the U.S. national boundary into Canada and Mexico.

that hydrologic forecast uncertainty must be quantified and that the magnitude of hydrologic forecast uncertainty should be passed on to decision makers and end-users in clear, understandable ways (Pappenberger et al. [144]; Wetterhall et al. [184]).

To illustrate the basis of hydrologic forecast uncertainty, this research will draw on data from U.S. National Oceanic and Atmospheric Administration (NOAA) NWS RFCs, specifically, the OHRFC, which is shown in Figure 2.3. The focus of the research is to establish the effect of model forcing (focusing on observed and forecast precipitation[1]) error on hydrologic forecast uncertainty. The main points to be made are that the most significant inputs to hydrologic models for rainfall-driven events used in forecasting are *QPE*, that is, observed precipitation and *QPF*, namely predicted/forecasted precipitation and that there are significant errors associated with their measurement and prediction, respectively.

Explicit quantification of hydrologic forecast uncertainty is one of the central themes of the NWS Hydrologic Services Program Advanced Hydrologic Prediction Services (AHPS) initiative [120]. The estimation of hydrologic forecast uncertainty for short lead-time (days 1 to 5) events is a area of active research within the NWS and elsewhere. Krzysztofowicz [112] outlines the need for probabilistic hydrologic forecasting, stating that probabilistic forecasts:

1. are scientifically more honest by providing prediction uncertainty
2. enable risk-based warnings for floods
3. allow rational decision making under the knowledge of prediction uncertainty
4. offer additional economic benefits due to improved decision making

[1]While significant in many regions of the world due to the influence of snow accumulation and melt processes, temperature estimation and prediction uncertainty will not be considered in order to limit the scope of the research task.

Probabilistic forecasts must include estimates of all the components of forecast uncertainty, including:

1. model input errors
2. inherent modeling errors (independent of the inputs)

Explicit quantification of hydrologic forecast uncertainty is one of the central themes of the NWS Hydrologic Services Program Advanced Hydrologic Prediction Services (AHPS) initiative, *National Research Council (NRC)* [141]. The estimation of hydrologic forecast uncertainty for short lead-time (days 1 to 5) events is a area of active research within the NWS. On the need to characterize the effects of input uncertainties for forecast precipitation and temperature, the *NRC* [133] states in *Completing the Forecast: Characterizing and Communicating Uncertainty for Better Decisions Using Weather and Climate Forecasts* (http://www.nap.edu/catalog/11699.html):

> *The NWS operational hydrology short-term forecast products carry uncertainty that is to a large degree due to forecasts of precipitation and temperature that serve as hydrologic model input and which are generated by objective or in some cases subjective procedures applied to the operational NCEP model forecasts.*

2.1 Precipitation variability

Principle data inputs for NOAA/NWS RFC hydrologic models are observed and forecasted precipitation and temperature. Observed precipitation is estimated through a multisensor estimation process using the Multisensor Precipitation Estimator (MPE) software [103] which utilizes rain gauges, NWS *Next Generation Radar (NEXRAD)* doppler radar, shown in

Figure 2.4, and, in some instances, satellite precipitation estimates to produce an un-biased optimal estimate of hourly precipitation fields. Forecasted precipitation is derived from numerical weather prediction (NWP) models, but meteorological forecaster adjustments are made at both the NWS Weather Prediction Center (WPC) and at local RFCs. However, the greatest sources of hydrologic prediction error derives from uncertainties in precipitation forecasts (Ebert and McBride, 2000 and Ebert et al, 2003), also known as quantitative precipitation forecast (QPF) and errors with the estimation of observed precipitation, or quantitative precipitation estimates (QPE) (see Anagnostou et al. [11], Seo et al. [160], and Krajewski and Ciach [108]).

2.1.1 Observed precipitation variability

One way to decrease hydrologic modeling uncertainty is to apply hydrologic models (and other models - snow model, for instance) at smaller subbasins scales with the hope of capturing the finer structure of precipitation and other hydrometeorological variability and spatial heterogeneities of basin characteristics. Finnerty et al. [73] and Smith et al. [165] with the Hydrologic Research Laboratory (HRL) of the NWS Office of Hydrology (OH) experimented with various approaches of applying the SAC-SMA model in a distributed modeling approach. Namely, they calibrated the SAC-SMA at a gaged location and applied the parameters to nested subbasins of varying sizes. These experiments demonstrated increased hydrograph peaks and runoff volumes with smaller basins and decreased hydrograph peaks and runoff volumes with larger basins. Attempts to identify consistent scaling relationships for parameter values between basins of differing sizes have been unsuccessful. It does not seem possible, as yet, to rationally adjust calibrated SAC-SMA parameters to be suitable for the differing characteristics of ungaged subbasins and maintain consistent hydrograph response.

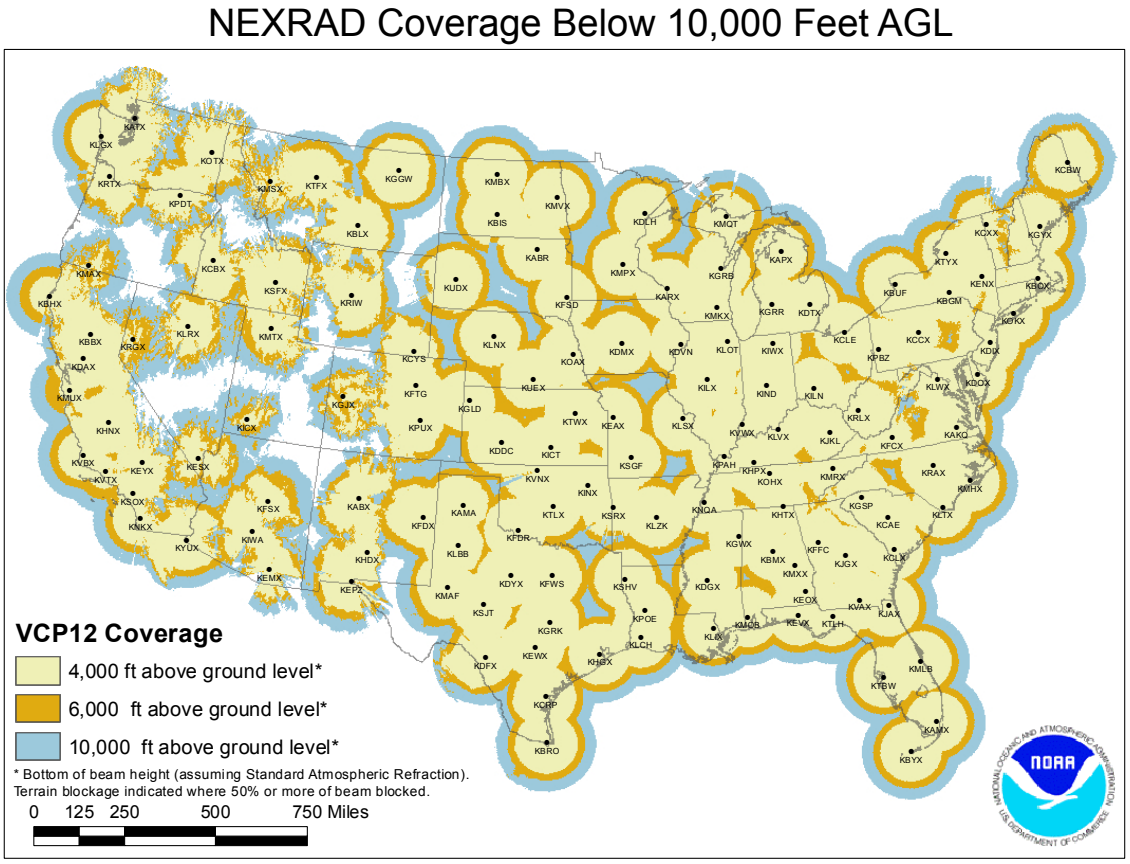

Figure 2.4: Location of NWS NEXRAD radar sites and radar coverage below 10,000 Feet above ground level (AGL). Note the areas in the western U.S. where there is no NEXRAD radar coverage (white).

Rainfall variability over watersheds is the dominant factor influencing runoff variability observed in hydrographs and provides the chief motivation for the adoption of distributed hydrologic modeling by the NWS in RFCs. Dawdy and Bergmann [50] showed that spatial rainfall variations significantly altered parameter calibration of the Stanford Watershed Model. Also using a hydrologic model, Wilson et al. [186] found large differences in the time-to-peak, peak discharge, and runoff volume depending on whether numerically generated rainfall input came from a single raingage, implying a spatially uniform distribution of rainfall, or from 20 covering the simulated drainage basin, where the rainfall was spatially variable. Based on raingage and flood data, Reich [152] found that there was no consistent relationship between point rainfall maxima and peak runoff maxima for 24 basins in Pennsylvania. However, subsequently, Larson and Reich [113] found that while there was high variability for individual years with Reich's 1970 data, the rank and recurrence interval of storm rainfall and peak runoff do have a central tendency of equality. This confirms, in other words, the accepted notion that the largest rain-induced floods tend to be produced by the greatest rainfalls.

In a somewhat different approach, Fogel [74] produced multiple regression equations for predicting runoff volumes from three small catchments ranging in area from 0.47 to 7.77 mi^2 (Table 2.1). He found that the spatially averaged storm rainfall and other factors accounted for appreciably less explained variance with increased drainage area, where Q: storm runoff (inches)*;* R: mean storm rainfall over the basin (inches)*;* i_{15}: maximum 15-minute rainfall intensity (inches)*;* t_{m}: time to the center of mass of rainfall (hours); and b_0, b_1, b_2: regression coefficients. Fogel's results indicate that the relationship between basin mean rainfall and peak storm runoff is consistent, that is, greater rainfalls produce larger flood peaks, but that considerable deviations occur about this tendency. Clearly, these deviations result from (1) the areal variability of rainfall over the individual basins, (2) the temporal distribution of

storm rainfall, and (3) antecedent watershed conditions. But for these Arizona watersheds, antecedent conditions are probably not significant since the time between rainfalls is large and considerable drying occurs during the intervening rain periods. By inspection, it appears the rainfall intensity factors explain less variance within basins than the differences in basin areas explain the variance between the different basins. This seems to confirm the idea that rainfall variability is the dominant factor in explaining runoff variability, which is especially evident with the simplest of the runoff prediction equations, $Q = b_0 + b_1 R$.

Studies of rainfall patterns of storms using dense raingage networks have shown that large spatial rainfall gradients exist within storms over short durations (<15 minutes). Huff [93], for example, obtained spatial correlations of 1, 5, and 10 minute rainfall rates, shown in Table 2.2, and total storm accumulations for a raingage network of 50 recording gages over a 100 sq. mi. area in east central Illinois from a 29 storm sample (see Figure 2.5). Additionally, an analysis by Jones and Wendland [98] of continuous recording raingage networks throughout the world reveals that 1-minute rainfall intensities for showery rains, that is, storms exhibiting thunderstorm or near-thunderstorm intensity rainfalls, were essentially uncorrelated at distances of 12 km from the reference raingage (Figure 2.6) for July and October storms. Osborn et al (1979) found that total rainfall accumulations of, primarily, air-mass thunderstorms, had correlation coefficients between 0.4 to 0.6 at 5 km, 0.1 to 0.3 at 15 km, and 0.0 to 0.1 at 25 km for raingage networks in Arizona and New Mexico. Since the deployment of NEXRAD systems by the NWS, routine observation of the spatial variability of rainfall is commonplace. NEXRAD Stage-3 Precipitation Processing rainfall estimates, which are made on a 4 km spatial grid, reveal detailed rainfall variations within storms that are evident nationwide.

Adams [3] studied the intra-storm spatial variability of flooding indicated by comparisons of interval estimates of the return periods of peak flows for basins in close proximity to

	Subwatershed		
	W-1B	W-2	W-3
Area, mi^2	7.77	4.49	0.47
Predictive Equation	r^2		
$Q = b_0 + b_1 R$	0.61	0.63	0.86
$Q = b_0 + b_1 R + b_2 i_{15}$	0.69	0.79	0.89
$Q = b_0 + b_1 R + b_2 i_{15} t_m^{1/3}$	0.75	0.87	0.94

Table 2.1: Table showing the coefficients of determination for runoff prediction equations Fogel [74].

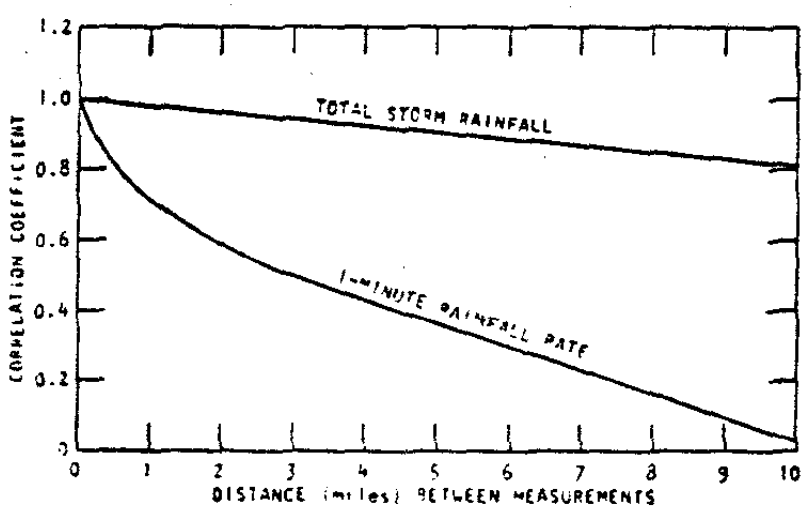

Figure 2.5: Mean correlation decay with distance between measurements for 1-minute rainfall rate & storm total rainfall for warm-season events (from Huff [93]).

Distance (miles)	Correlation Coefficient		
	1-min.	5-min.	10-min.
1	0.71	0.72	0.77
2	0.58	0.51	0.61
4	0.41	0.29	0.41
6	0.28	0.20	0.25
8	0.16	0.13	0.15

Table 2.2: Correlation coefficients of rainfall rates with respect to distance from a central raingage (from Huff [93]).

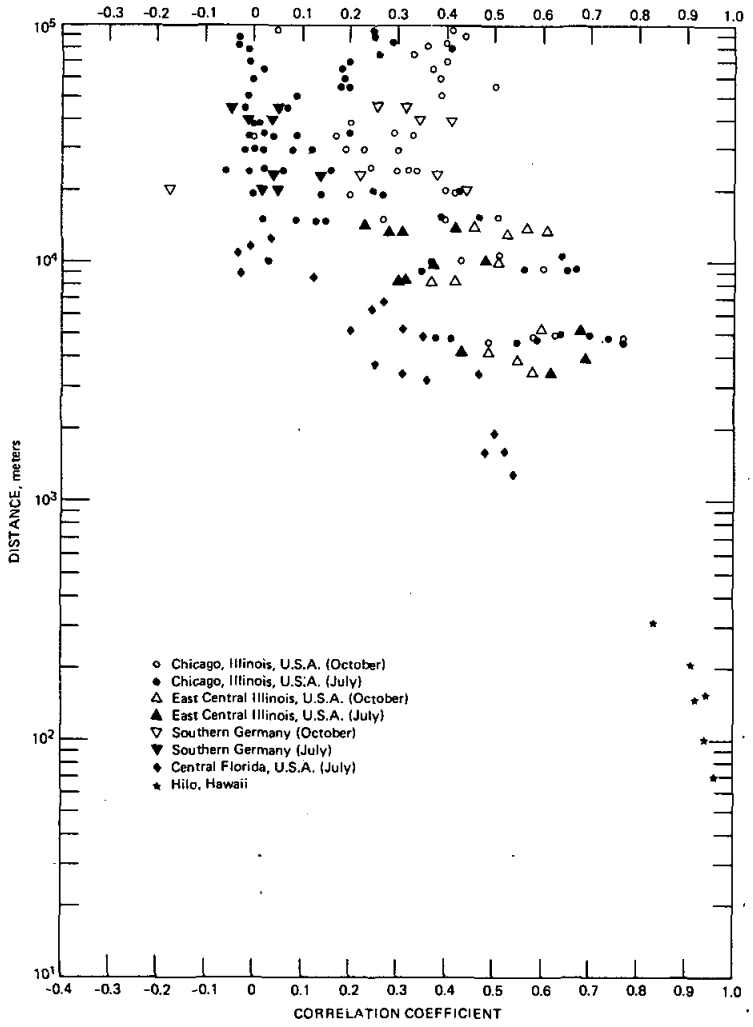

Figure 2.6: Correlation coefficients of concurrent rainfall intensity with distance from reference site (from Jones and Wendland [98]).

each other within the Piedmont region of Maryland. Results showed substantial differences between interval estimates on return periods between nearby basins within individual storms, which suggests substantial rainfall differences between basins.

Hamlin [89] emphasizes the importance of rainfall variability and the need for accurate rainfall estimates for variable source area models. He suggests that lumped-parameter models do not require as stringent rainfall estimates, since significant averaging is made within the lumped process modeling. Faurès et al. [72] and Goodrich et al. [84] found very significant rainfall measurement and runoff modeling errors for a small, 4.4 ha semiarid catchment, where the coefficient of variation for peak rate and runoff volume ranged from 9 to 76%, and from 2 to 65%, respectively, over eight storm events. They concluded that the assumption of spatial uniformity of rainfall at the 5 ha scale in convective environments appeared to be invalid.

Analyses of precipitation estimates from radar can be problematic. Reed and Maidment [150] have identified coordinate transformation errors in the NEXRAD HRAP coordinate system which is the basis of the digital precipitation data used by the NWS for input to NWSRFS hydrologic models. The magnitude of these shape distortion errors depend on latitude and range between 0.3% to 0.6% in the conterminous U.S. These errors translate into HRAP grid size variations, causing the true area of HRAP grid cells to range from 13 km^2 in Miami to 19 km^2 in Minneapolis.

The use of gridded historical datasets, such as the *Parameter-elevation Regressions on Independent Slopes Model (PRISM)* at the Spatial Climate Analysis Service Oregon State University (http://www.prism.oregonstate.edu), desribed by Taylor et al. [171], Daly et al. [46], Taylor et al. [172], Daly et al. [47], and Daly et al. [48] are useful analyzing the spatial bias patterns of radar-derived precipitation estimates. PRISM is an expert system that uses point data and a digital elevation model (DEM) to generate gridded estimates

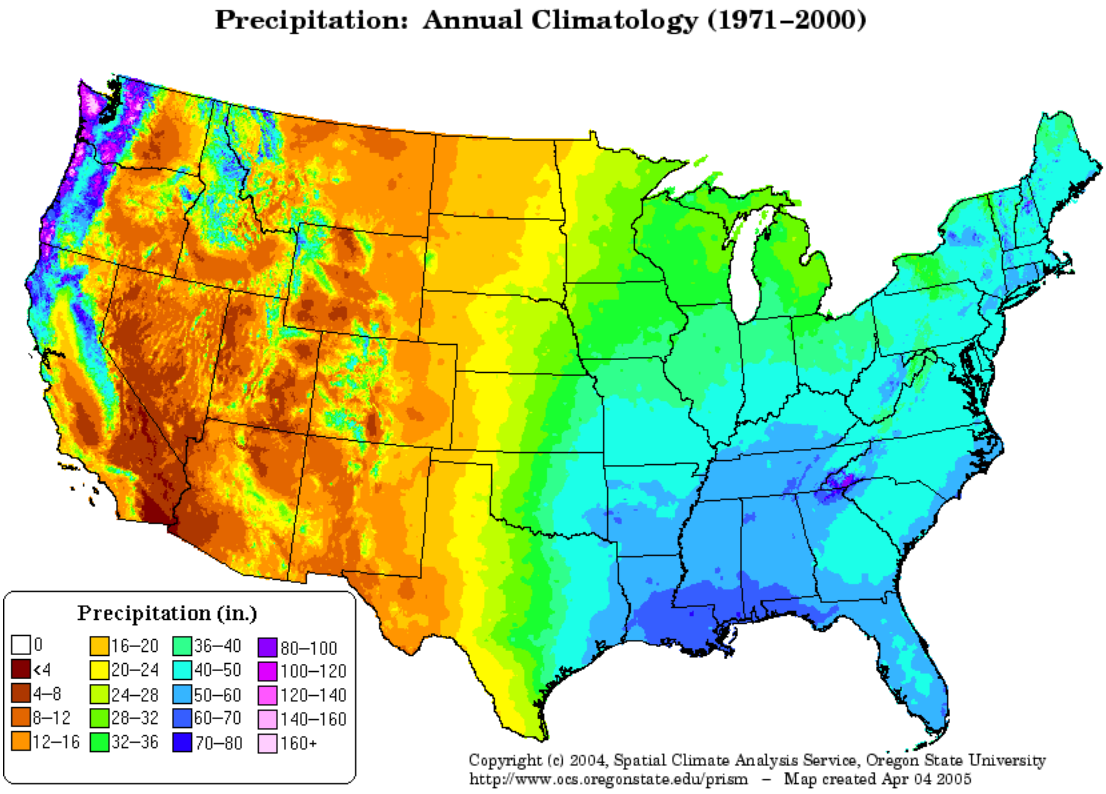

Figure 2.7: PRISM precipitation climatology for the period 1971-2000.

of climate parameters estimates of annual, monthly, and event-based climatic elements including precipitation and temperature. Figure 2.7 shows an example of the PRISM analysis for precipitation climatology for the conterminous U.S for the period 1971-2000. Figure 2.8 shows significant underestimation by NEXRAD MPE estimated mean areal precipitation relative to PRISM gage only estimates for 2010. Systematic biases such as this are common.

2.1.2 Forecast precipitation variability

The focus of this discussion is quantitative precipitation forecast (QPF) generation methods for the purposes of hydrologic prediction. The importance of QPF in hydrologic forecasting is a long standing issue. Georgakakos and Hudlow [83] discussed the urgency to develop

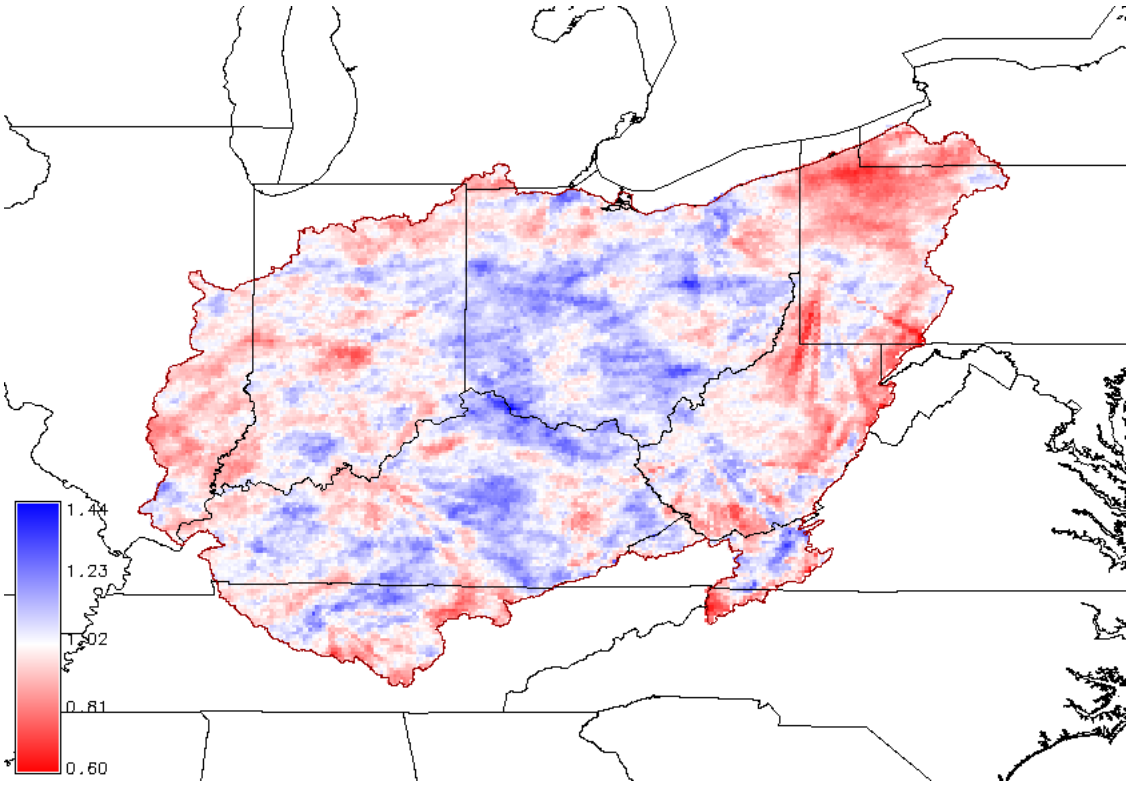

Figure 2.8: MPE bias with respect to PRISM for 2010 for the OHRFC forecast region; MPE over-estimation is indicated by blue colors and under-estimation are shades of red. Bias values equal to 1.0 are unbiased (white areas).

QPF methods to meet the needs of hydrologic prediction. Unfortunately, as recently as 1998 Gaudet and Cotton [81] report that "precipitation is notorious for being difficult to predict accurately".

Recently Ebert et al. [65] studied QPF performance of *General Circulation Models* (GCMs) from the National Centers for Environmental Prediction (NCEP) in the United States, the Deutscher Wetterdienst (DWD) in Germany, and the Bureau of Meteorology Research Centre (BMRC) in Australia for the period 1997 through 2000. This work was done within auspices of the *Working Group on Numerical Experimentation (WGNE)*, established under the World Meteorological Organisation's World Climate Research Programme (WCRP) and Commission for Atmospheric Sciences (CAS). They report that QPFs produced by NWPs easily outperformed *persistence* and provided useful routine guidance, but the forecasts, also, were far from perfect. Ebert et al. [65] also found that the predicted rainfall of models is highly sensitive to the predicted atmospheric and surface conditions, imply that a good rainfall forecast points to a good forecast of other atmospheric variables. On the other hand, a bad rainfall forecast may have little to do with the model parameterization for precipitation, but yet me be much more a function of how a NWP is tuned to optimize model performance of other variables. They state:

> "The process of improving model numerics and physics is a complicated juggling act. Unless the accurate prediction of rainfall is made a top priority then improvements in NWP model QPF will continue to be realized slowly."

Buizza [31] performed an experiment to test the magnitude of QPF errors resulting from initial conditions alone with forecasts of rainfall over Australia during January and July 1998 from the *European Centre for Medium-Range Forecasts (ECMWF) Ensemble Prediction System (EPS)* for 24- and 48-hr forecasts. Results showed that most of the difference in

performance between what is currently achieved in skill and perfect QPF skill could be eliminated with a perfect model. This suggests that, by far, errors in the model initial conditions were far less important than the errors induced by current model numerics and physics in QPF skill. Ebert et al. [65] draw some important conclusions, stating:

> "... one of the most promising and practical ways to improve quantitative precipitation forecasting using existing NWP models is the use of ensembles to generate multiple rain scenarios and probabilistic forecasts."

and continues by saying:

> "While improvements in our understanding of rainfall process, numerical models, and data assimilation are important steps toward improving quantitative precipitation forecasting, ensemble prediction may offer the most effective means of making best use of the imperfect QPFs available to us at present."

Work by Stensrud et al. [168], Wandishin et al. [178], and Ebert [64] have shown the utility of NWP model ensembles of QPF.

2.2 Ensemble Hydrologic Forecasting

There has been considerable research into probabilistic methods to quantify hydrologic forecast uncertainty (see for example, Buizza [31], Wandishin et al. [178], Franz et al. [77], National Research Council [133], Schaake et al. [157], and Adams and Ostrowski [1]).

Probabilistic hydrologic forecasting addresses the inherent uncertainties found in deterministic forecasting discussed in previous sections, ranging from short lead-time (1-7 days) to long

lead-time (monthly, seasonal, and annual) temporal scales. For short lead-time probabilistic forecasting, Krzysztofowicz [111] proposed a *Bayesian* approach while others have employed *monte carlo* methods utilizing variations of *ensemble* methodologies, such as Adams and Ostrowski [1] with the MMEFS, Demargne et al. [58] with the *Hydrologic Ensemble Forecast Service (HEFS)*, as part of the Advanced Hydrologic Prediction Service, and Werner et al. [183] with medium-range meteorological ensemble inputs of temperature and precipitation derived from the NCEP Medium-Range Forecast (MRF) model. Example output from such an ensemble hydrologic forecast system is shown in Figure 2.9 for the OHRFC MMEFS NAEFS, for the Greenbrier River at Alderson, WV, for the period March 1-7, 2015. Hydrologic model inputs for the MMEFS are forecasted mean areal precipitation and temperature time-series derived from output grids from numerical weather prediction (NWP) models comprising the NOAA/NWS *National Centers for Environmental Prediction (NCEP) North American Ensemble Forecast System (NAEFS)* [35] and *Short Range Ensemble Forecast System (SREF)* [62]. A recent review by Cloke and Pappenberger [40] describes features of many recently implemented medium-range lead-time ensemble hydrologic forecast systems. Siddique and Mejia [162] and Alfieri et al. [10], further illustrate regional and global systems, respectively, for ensemble hydrologic forecasting. These forecasting systems have been implemented for the issuance of routine flood alerts and warnings and broader water resources applications, important in reservoir and drought management (Hamlet et al. [88]; Raff et al. [146]; Anghileri et al. [14]; Turner et al. [175]).

International efforts in ensemble hyrometeorological modeling include *The Observing System Research and Predictability Experiment (THORPEX) Interactive Grand Global Ensemble (TIGGE)* project, which includes as one of its primary goals "facilitate exploring the concept and benefits of multimodel probabilistic weather forecasts, with a particular focus on high-impact weather prediction" [22]. *Hydrological Ensemble Prediction Experiment*

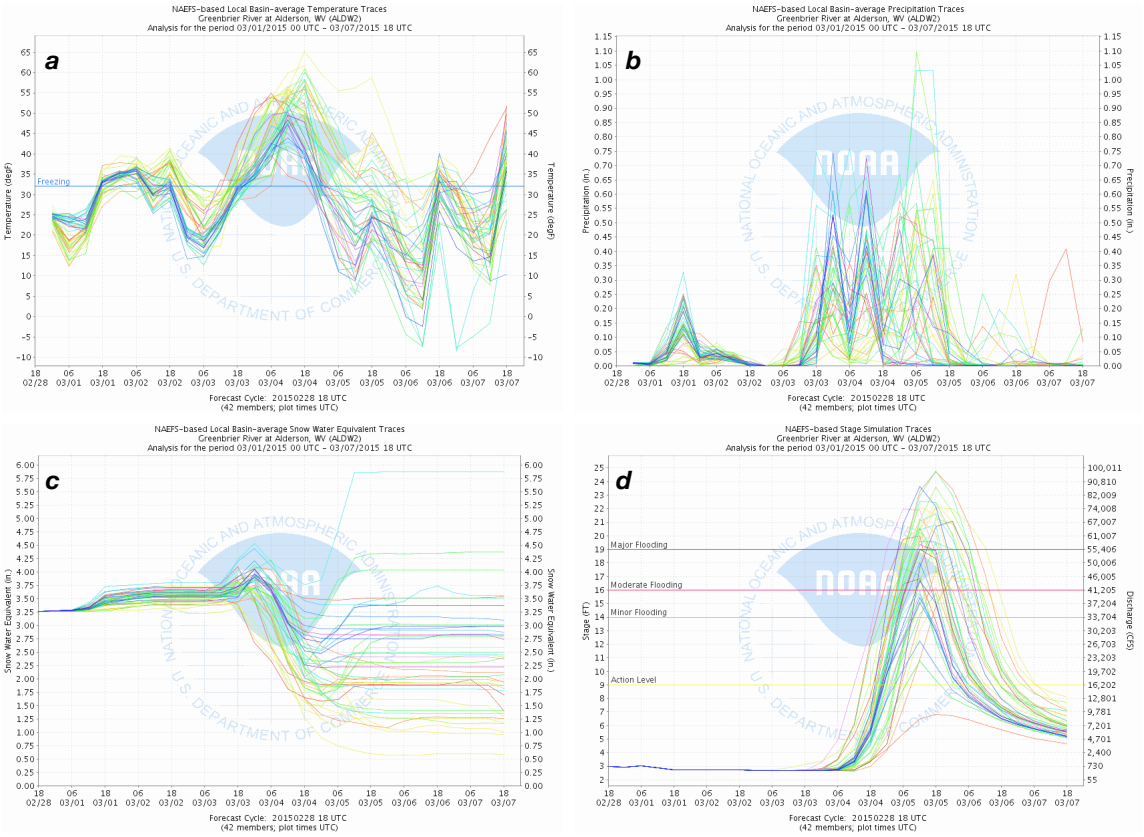

Figure 2.9: Example ensemble hydrologic forecast from the NOAA/NWS MMEFS using *NAEFS* ensemble (a) temperature and (b) precipitation inputs, producing (c) snow water equivalent (SWE) from the NOAA/NWS SNOW-17 model and (d) hydrologic stage/discharge forecasts from the SAC-SMA rainfall-runoff model within the CHPS-FEWS forecast system at the OHRFC for the Greenbrier River at Alderson, WV, for the period March 1-7, 2015.

(HEPEX; www.hepex.org/), launched in 2004, has facilitated communication and collaboration among the atmospheric and hydrologic communities, including involvement from forecast users with goals of improving ensemble forecasts and demonstrating their utility in decision making in water management.

Pioneering development of ensemble hydrologic forecasting methodologies for water resources is described by Twedt et al. [176] and Day [51] with, what was called at that time, *Extended Streamflow Prediction (ESP)* within the NWS *River Forecast System (NWSRFS)*. The initial development and application of ESP methodology is chronicled in "Tracing The Origins of ESP"[2]. For long lead-time predictions at NWS RFCs, ESP utilizes basin averaged historical temperature and precipitation time-series as surrogates for *possible* future hydrologic model forcings for the generation of ensemble monthly, seasonal, and annual streamflow forecasts. An example of ESP output for *exceedance probability* in the OHRFC area for the Ohio River at Golconda, IL is shown in Figure 2.10. A few points of interest are:

1. that the two conditional simulations (CS), conditional because of their dependence on *initial basin conditions*, show high-exceedance probability values beginning at a stage of 38.2 feet, reflecting that initial flow conditions for the Ohio River are at that level;

2. both conditional simulations are shifted to the right of the historical simulation (HS)[3], which implies that there is a lower probability of attaining a given stage/flow level. This, in turn, implies that the basin conditions are drier than normal[4], relative to the historical simulation;

3. the CS (black) utilizing NOAA/NWS *Climate Prediction Center (CPC)* climate ad-

[2]April 26, 2016 by Andy Wood, https://hepex.irstea.fr/tracing-the-origins-of-esp/.

[3]A historical simulation is made as a single, continuous model simulation at the beginning of the historical record for precipitation and temperature time-series through the most recent available data, at the 1- or 6-hour model time step, utilizing all model components in the forecast system.

[4]Observed antecedent rainfall that had occurred weeks and months prior to the ESP run also showed less-than-normal rainfall.

justments is shifted to the right of the CS (green), which does not utilize CPC climate adjustments, thus implying drier future conditions.

Item (3) points to the need that, for longer lead-time forecasts, the influence of predictable, climate-scale meteorological features should be included in hydrological forecasts[5], such as El Niño–Southern Oscillation (ENSO) and, the cooling phase, La Niña effects (Werner et al. [183]; Wood and Lettenmaier [188], Moradkhani and Meier [127]; Bastola et al. [16]; Forzieri et al. [75]; Bradley et al. [25]; Beckers et al. [17]; Mendoza et al. [122]; Crochemore et al. [42]). For OHRFC AHPS monthly and seasonal streamflow forecasts, CPC near term climate adjustments of historical precipitation and temperature time-series are made prior to use in ESP simulations to reflect wetter/dryer or warmer/cooler future conditions that are associated with climate influences. For short lead-time hydrologic forecasts climate influences are minimal, but other factors are important, such as initial basin conditions or model structure that lead to systematic biases. Considerable research can be found on these topics (Ebtehaj et al. [66]; DeChant and Moradkhani [52], Zalachori et al. [198], DeChant and Moradkhani [53]).

Franz et al. [77], Demargne et al. [56], Demargne et al. [57], DeChant and Moradkhani [54], and others have identified the need for bias correction and correction of ensemble spread of hydrologic ensemble forecasts. Wood and Schaake [190] and Bogner and Pappenberger [20] discuss methods for correcting hydrologic ensemble forecast bias and reliability errors. Consequently there has been considerable effort to develop methodologies to address hydrologic ensemble biases and spread using pre- and post-processing techniques. Zhao et al. [203] evaluated the performance of a statistical post-processor for imperfect hydrologic model forecasts and show that a proposed General Linear Model (GLM) Post-Processor (GLMPP),

[5]Chapman Conference (2013) on Seasonal to Interannual Hydroclimate Forecasts, http://chapman.agu.org/watermanagement/files/2013/07/Final-Program1.pdf.

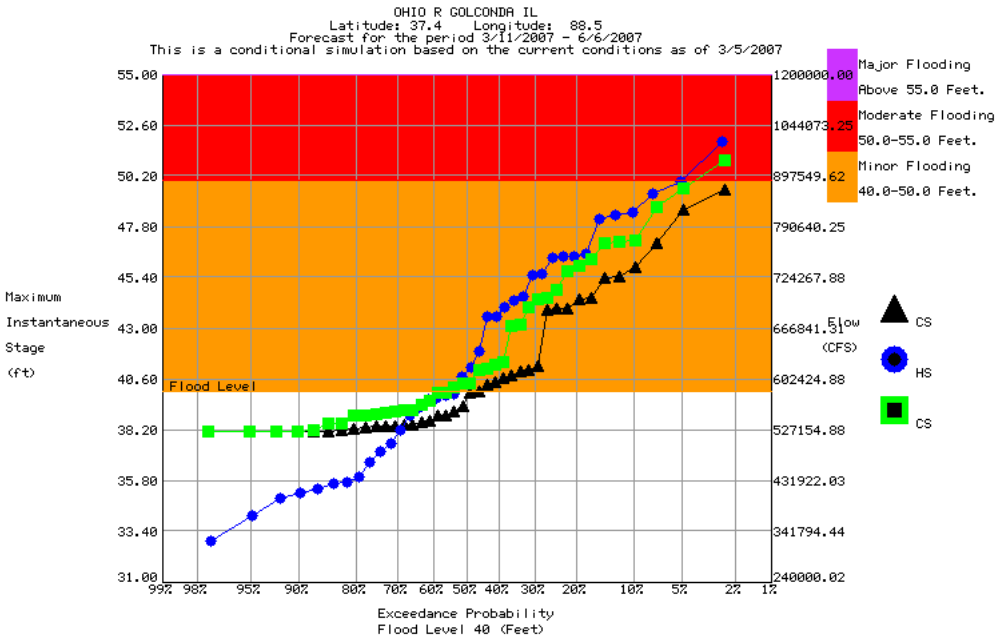

Figure 2.10: Probability of exceedance for OHRFC AHPS/ESP ensemble hydrologic forecast for the Ohio River at Golconda, IL, March 11 – June 6, 2007, showing historical simulation (HS, blue), conditional simulation without CPC climate adjustments (CS, green), and conditional simulation with CPC climate adjustments (CS, black). The orange region designates above *Minor Flood* level and red above *Moderate Flood* level.

built using data from the calibration period, removes the mean bias when applied to hydrologic model simulations from both the calibration and verification periods. Li et al. [116] present a comprehensive review of commonly used statistical post-processing methods for meteorological and hydrological forecasts. Sharma et al. [161] propose a method for preprocessing ensemble precipitation forecasts for hydrologic forecasting, finding greater skill than the raw forecasts.

Ensemble pre-processing methods and hydrologic hindcast experiments proposed by Demargne et al. [58] are specifically aimed at bias correction of forecast meteorological inputs and quantification of hydrologic model error, respectively. For the purposes of the proposed research, no pre- or post-processing or bias-correction techniques will be utilized. The reason for this is that applying such techniques could obfuscate the underlying goal of the research, which is to assess the whether or not ensemble hydrologic mean or mean forecasts are superior to current deterministic forecasts. The literature shows that the application of various methodologies will improve ensemble forecasts; this is known. Making use of such techniques in the proposed research could cause confusion as to whether the outcomes resulted from the underlying hypothesis or the use of bias correction or some other pre- or post-processing methodology.

Chapter 3

Hydrometeorological Forcing Errors for a Real-time Flood Forecast System in the Ohio River Valley, USA

3.1 Abstract

Errors in hydrometeorological forcings for hydrologic modeling lead to considerable prediction uncertainty of hydrologic variables. Analyses of *Quantitative Precipitation Estimate (QPE)* and *Quantitative Precipitation Forecast (QPF)* errors over the *Ohio River Valley* were made to quantify QPE and QPF errors and identify hydrologic impacts of forcing errors and possible improvements resulting from advancements in precipitation estimation and forecasting. Monthly, seasonal, and annual bias analyses of Ohio River Forecast Center (OHRFC) *NEXt-generation RADar (NEXRAD)* based *Stage III* and *Multisensor Precipitation Estimator (MPE)* precipitation estimates, for the period 1997-2016, were made with respect to *Parameter-elevation Regressions on Independent Slopes Model (PRISM)* precipita-

tion estimates. Verification of QPF from NWS River Forecast Centers from the NOAA/NWS *National Precipitation Verification Unit (NPVU)* was compared to QPF verification measures from several numerical weather prediction models and the NOAA/NWS *Weather Prediction Center (WPC)*. Improvements in NEXRAD based QPE over the OHRFC area have been dramatic from 1997 to present. However, from the perspective of meeting hydrologic forecasting needs, QPF shows marginal improvement. A hydrologic simulation experiment illustrates the sensitivity of hydrologic forecasts to QPF errors based on *Threat Score (TS)*. Experiments show there is considerable hydrologic forecast error associated with QPF at expected WPC TS levels and, importantly, that higher TS values do not necessarily translate into improved hydrologic simulation results.

3.2 Introduction

Hydrologic forecast accuracy is largely dependent on the magnitude of measurement and prediction errors of hydrometeorological forcings used as model inputs (Maurer and Lettenmaier [119]; Tetzlaff and Uhlenbrook [173]; Benke et al. [18]; Wood and Lettenmaier [189]; Newman et al. [137]). As early as 1969, research by Fogel [74] quantified differences in watershed runoff due to rainfall variability, using a dense raingauge network for the Atterbury experimental watershed in Arizona. More recently, using distributed precipitation inputs, Wilson et al. [186] and Faurès et al. [72] demonstrated that large variations in modeled watershed runoff can result from spatially variable rainfall, on the order of 9 to 76% for peak runoff rates and 2 to 65% for runoff volume, for a 4.4 ha semiarid catchment [72]. Also utilizing dense raingauge networks, Jones and Wendland [98], Goodrich et al. [84], and Zhang et al. [201] report the occurrence of significant rainfall variability over short distances (100-1000 m) which, with gridded precipitation fields, would be considered the subgrid scale.

In an operational setting, hydrologic model forcings usually take the form of quantitative estimates of observed and forecasted precipitation and temperature. Past studies have shown that accurate quantification of observed precipitation, known as *quantitative precipitation estimate (QPE)*, is problematic for both raingauge-based and radar-derived estimates. Raingauge based estimation errors arise, largely, due to insufficient gauge density of raingauge networks (Huff [93]; Sungmin et al. [170]; Cecinati et al. [36]) and low-*catch* biases of individual raingauges (Humphrey et al. [94]; Ciach and Krajewski [39]; Ciach [38]). It has been shown that the highest quality radar based estimates of precipitation depend on raingauge based bias adjustments and other corrections within precipitation processing algorithms (Anagnostou et al. [11]; Young et al. [196]), to account for systematic detection and measurement errors associated with beam attenuation, full/partial beam blockage, ground clutter, beam overshooting, curvature of the Earth, anomalous propagation, brightband contamination, conversion from reflectivity to rainfall rates, i.e., non-unique *Z-R* relationships, beam attenuation, and range effects, including sampling and averaging errors.

Significant research and development related to radar precipitation processing algorithms in the U.S. is evident since the 1980s. These include development efforts by the *National Oceanic and Atmospheric Administration (NOAA), National Weather Service (NWS)* with *Stage II & Stage III* (Fulton et al. [79]; Young et al. [196]), *Muti-sensor Precipitation Estimator (MPE)* (Seo [158]; Seo et al. [159]; Breidenbach et al. [27]; Breidenbach and Bradberry [26]; Kitzmiller et al. [103]; Eldardiry et al. [67]), *Stage IV* [117], which is a nationwide mosaick of Stage III or MPE products from RFCs. A recent study by Nelson et al. [136] made an assessment of NCEP Stage IV QPE. RFCs have utilized *Multi-Radar/Multi-Sensor (MRMS)* [200] precipitation processing systems (PPSs) estimates within their QPE workflows since the data became available in 2011. The MRMS PPS, originally called the *National Mosaic and QPE (NMQ)* algorithm package, was developed at the *National Severe Storms Labo-*

ratory (NSSL) and subsequently moved to the NOAA *National Centers for Environmental Prediction (NCEP)* for operational support of NWS River Forecast Centers (RFCs), shown in Figure 3.1, and Weather Forecast Offices (WFOs).

In western regions of the U.S., where radar beam blockage is problematic in mountainous areas, NWS estimation methods rely on data from raingauge and *Natural Resources Conservation Service (NRCS)*, *Snow Telemetry (SNOTEL)* networks for precipitation estimation. Gauge data are processed at RFCs, using spatial interpolation algorithms and historical data, such as *Parameter-elevation Relationships on Independent Slopes Model (PRISM)* (Taylor et al. [171]; Taylor et al. [172]; Daly et al. [47]), within the *Advanced Weather Interactive Processing System (AWIPS)*, to generate gridded estimates of precipitation utilizing *Mountain Mapper* and *Data QC* [156].

Difficulties with the prediction of future precipitation, referred to as *quantitative precipitation forecast (QPF)*, are compounded by the need to accurately predict occurrences of heavy precipitation accumulations spatially. That is, the location of flood producing rainfall matters significantly, which is demonstrably evident with flash flood scale events, where the occurrence of excessive rainfall, accompanied by flooding, can be *hit-or-miss* over very short distances with devastating outcomes (Smith et al. [163]; Baeck and Smith [15]; Smith et al. [164]; Borga et al. [21]; Alfieri et al. [9]; Broxton et al. [30]).

With winter time storms, errors in temperature estimation and prediction can incorrectly identify the physical state of hydrometeors, suggesting the occurrence of rainfall rather than snowfall, or the reverse [180]. Mizukami et al. [125] and Hunter and Holroyd [95] discuss the implications of mis-typing the physical state of precipitation (rain, snow, ice, hail, etc.) and how such errors lead directly to hydrologic forecast error. Moine et al. [126], Rössler et al. [154], and Wayand [179] show that errors in the estimation of snow accumulation and *snow water equivalent (SWE)* become especially problematic during rain-on-snow and significant

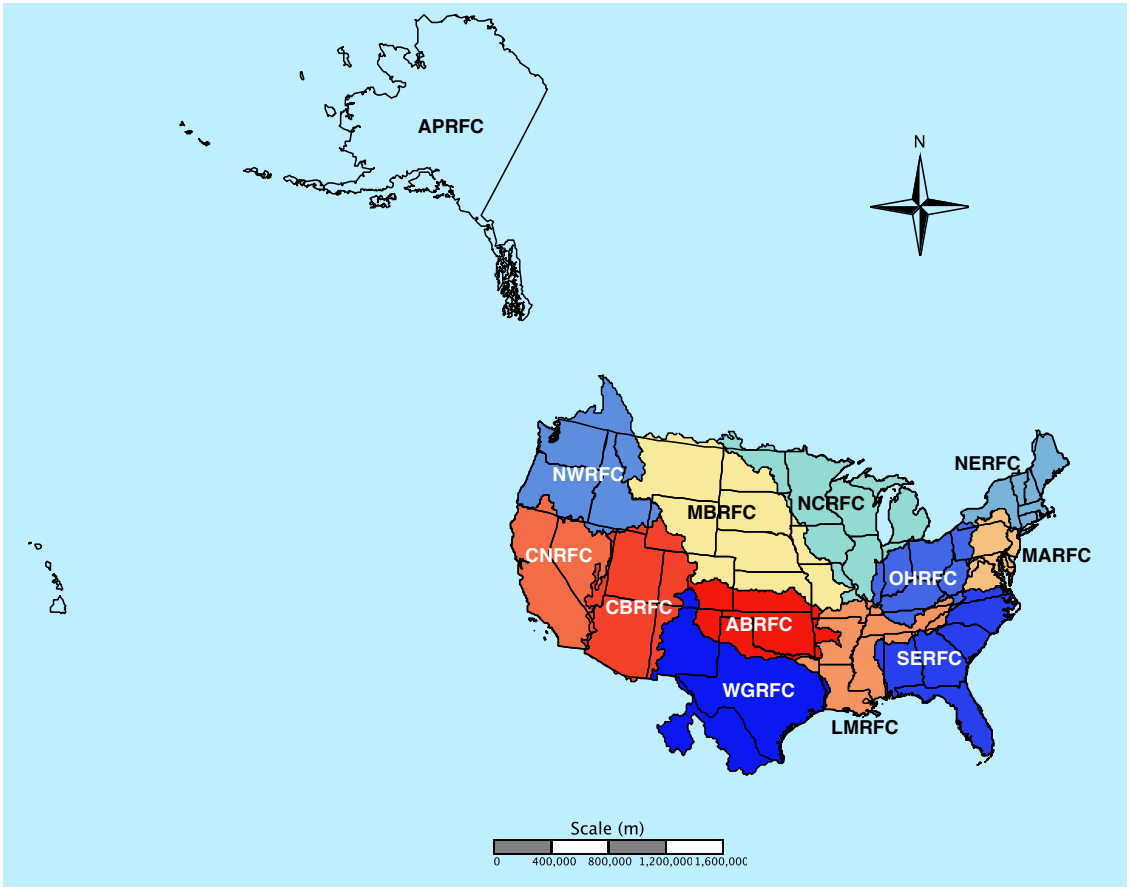

Figure 3.1: The NWS 13 River Forecast Centers (RFCs) – Alaska/Pacific RFC (APRFC), Arkansas-Red RFC (ABRFC), Colorado Basin RFC (CBRFC), California-Nevada RFC(CNRFC), Lower-Mississippi RFC (LMRFC)Middle Atlantic RFC (MARFC), Missouri Basin RFC (MBRFC), North Central RFC (NCRFC), Northwest RFC (NWRFC), Ohio RFC (OHRFC), Southeast RFC (SERFC), and West Gulf RFC (WGRFC). Please note that several RFC boundaries extend beyond the U.S. national boundary into Canada and Mexico.

temperature-driven snowmelt events. With temperature-index based snow models, such as the NWS *Snow Accumulation and Ablation* model, *SNOW-17* [12], used by NWS RFCs, erroneous temperature estimates can lead to inaccurate snowmelt rates. The effects of wind, terrain, and vegetation on snow estimation (Winstral et al. [187]; Essery and Pomeroy [69]) and modeling (Essery et al. [70]; Xiao et al. [195]; Bowling et al. [23]; Liston and Elder [118]) are significant as well.

Thiboult et al. [174] identify three broad sources of of total hydrologic forecast uncertainty, namely (1) *model structure*, (2) *model state initial conditions* (such as initial soil moisture, snow water equivalent, streamflow, etc.), and (3) *forcing* uncertainties. The focus of this study are the latter two sources of hydrologic forecast uncertainty. Specifically, first, we analyze QPE and QPF errors over the NOAA/NWS *Ohio River Forecast Center (OHRFC)* area of responsibility, shown in Figure 3.1, and, second, the hydrologic modeling and forecast error produced in response to QPE and QPF forcing errors. The OHRFC region was selected for this study because of the availability of data. Section 3.3 examines the spatial bias patterns of Stage III/MPE precipitation estimates and changes over time for the OHRFC area. A historical simulation using Stage III/MPE precipitation estimates as the principal model forcing, demonstrates improvements in hydrologic model simulation resulting from changes in precipitation estimation. Section 3.4 presents results of WPC and NOAA/NWS *National Precipitation Verification Unit (NPVU)* QPF verification. Additionally, the implications to hydrologic forecast uncertainty are examined with respect to past and current expected levels of QPF accuracy, using a hydrologic *monte carlo* simulation experiment. A summary and discussion of the limitations of the work, as well as implications to hydrologic forecasting, and final conclusions are presented in section 3.5.

3.3 QPE biases

The OHRFC, shown in Figure 3.1, has produced radar-based precipitation estimates derived from the NEXRAD network of *Weather Surveillance Radar-1988 Doppler (WSR-88D)* radars [43] since 1996. NEXRAD radars utilized by the OHRFC are shown in Figure 3.2 and are listed in Table 3.1. However, use of NEXRAD data as model forcings for operational hydrologic forecasting did not begin immediately due to significant changes in the OHRFC operational hydrological environment, which included operational implementation of the *Advanced Weather Interactive Processing System (AWIPS)* [140] and the *NWS River Forecast System (NWSRFS)* (U.S. Department of Commerce [177]; Adams [5]).

3.3.1 History

Although Stage II/III products were generated beginning in 1996 at the OHRFC, these data were not used in hydrologic forecast operations until 1998. A timeline of the period of use of *Stage II & III* and *MPE* by the OHRFC is shown in Figure 3.3. Stage II & III and MPE estimates are generated from the individual radars to produce mosaicked, gridded precipitation fields within the NWS *Hydrologic Rainfall Analysis Project (HRAP) polar stereographic* projection grid, which is described by Fulton [78] and Reed and Maidment [151]. Significant changes to the NEXRAD PPS at RFCs followed Stage II & III implementation. With the development of MRMS [200] at NSSL, the OHRFC and other RFCs began use of MRMS precipitation estimates in late 2012 as the initial step in MPE precipitation processing, which includes manual, interactive quality control by hydrometeorological forecasters using AWIPS MPE software.

An early significant MPE improvement, identified in Figure 3.3, was the correction of an algorithmic error that produced truncated rain-rate values, leading to precipitation under-

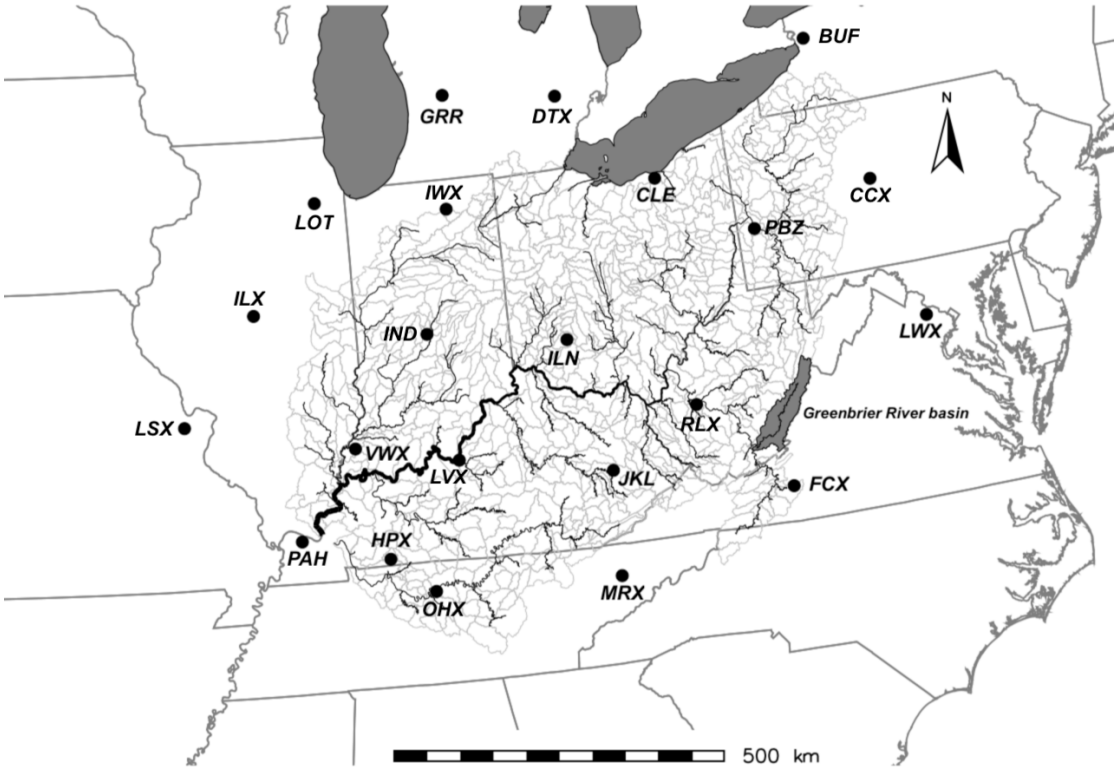

Figure 3.2: NEXRAD WSR-88D radar locations (black circles) in the NOAA/NWS OHRFC area of forecast responsibility. Refer to Table 3.1 for details. Also shown are 796 OHRFC modeling subbasins (light gray outlined areas) modeled operationally within the *CHPS-FEWS* hydrologic forecasting system and, for reference, the Ohio River and major tributaries (black lines). The Greenbrier River basin, WV, discussed below, is shaded gray.

estimation [80]. This error was corrected by the development and implementation of the *Open Radar Product Generator (ORPG)*, with software Build 1 which was deployed primarily over the period April-July 2002, Build 3, which was deployed during the months April-July 2003, and Build 4, which was delivered during the October-December 2003 period [80]. ORPG Build 1 contained the most significant improvements to precipitation estimation of the three OPRG software builds. Additional enhancements to the NEXRAD PPS have followed (Kitzmiller et al. [102]; Kitzmiller et al. [103]), including the deployment of NEXRAD *dual polarization* in 2011, which was completed for the OHRFC region before June 2013.

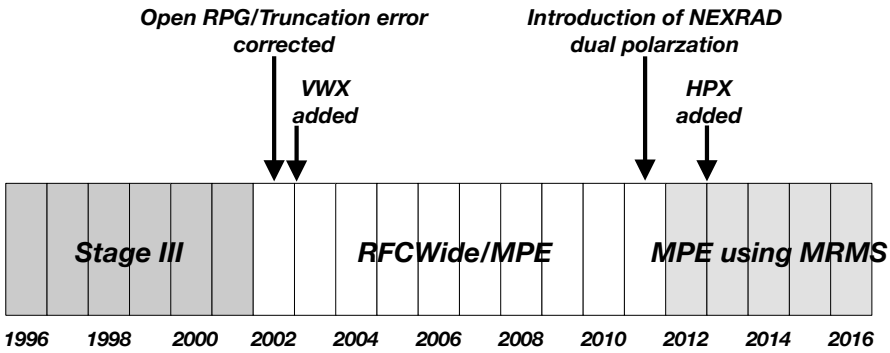

Figure 3.3: Timeline for OHRFC implementation of Stage III and MPE with changes to the NEXRAD network, with the addition of VWX and HPX radars (see Figure 3.2), and PPS changes.

Figure 3.3 also shows the addition of the VWX and HPX NEXRAD radars to those used by the OHRFC.

3.3.2 Data Analysis

The HRAP grid is nominally an ~4.7-km resolution, polar stereographic grid. PRISM estimates are developed at a 30-arcsec resolution in geographic (latitude-longitude) coordinates. Consequently, re-projection [71] and spatial interpolation of the PRISM grids to the HRAP coordinate system is necessary for PRISM-MPE/Stage III comparisons and analysis. Re-projection of the PRISM grids, employing a *bi-linear interpolation* algorithm, and bias analyses of the gridded fields use the *Geographic Resource Analysis Support System (GRASS)* Geographic Information System (GIS) [86]. The magnitude and spatial patterns of estimation error from *Stage III* and *MPE* precipitation estimates are calculated on a *pixel-by-pixel* basis with respect to gridded PRISM estimates over the OHRFC area. MPE/Stage III bias with respect to PRISM for the *masked* OHRFC region, using Equation 3.1:

Table 3.1: NEXRAD WSR-88D locations used by the OHRFC in the Stage III and MPE PPS, with the radar commissioning date, ground elevation, and tower height, listed in order of the commissioning date.

Location	**Longitude**	**Latitude**	**Elevation (m)**	**Height (m)**	**Date**	**ID**
Sterling, VA	-77.4781	38.9753	88.54	30	06/15/1994	LWX
St. Louis, MO	-90.6828	38.6989	197.00	30	07/15/1994	LSX
Louisville, KY	-85.9439	37.9753	219.15	30	11/29/1994	LVX
Chicago, IL	-88.0842	41.6044	202.08	25	12/16/1994	LOT
Pittsburgh, PA	-80.2183	40.5317	361.19	20	01/19/1995	PBZ
Cleveland, OH	-81.8600	41.4131	323.56	25	02/09/1995	CLE
Detroit, MI	-83.4719	42.7000	326.75	30	03/23/1995	DTX
State College, PA	-78.0047	40.9231	733.04	20	04/06/1995	CCX
Indianapolis, IN	-86.2800	39.7080	240.79	25	05/23/1995	IND
Wilmington, OH	-83.8214	39.4200	321.87	30	06/01/1995	ILN
Morristown, TN	-83.4022	36.1681	407.52	25	06/22/1995	MRX
Nashville, TN	-86.5622	36.2469	176.48	25	07/06/1995	OHX
Blacksburg, VA	-80.2742	37.0239	874.17	25	08/03/1995	FCX
Charleston, WV	-81.7233	38.3111	329.18	30	08/24/1995	RLX
Paducah, KY	-88.7719	37.0683	119.48	30	09/13/1995	PAH
Lincoln, IL	-89.3367	40.1503	177.39	30	01/03/1996	ILX
Grand Rapids, MI	-85.5450	42.8940	237.13	25	02/01/1996	GRR
Buffalo, NY	-78.7370	42.9490	211.23	30	04/04/1996	BUF
Jackson, KY	-83.3131	37.5908	415.75	25	10/25/1996	JKL
North Webster, IN	-85.7000	41.3600	292.30	25	03/17/1998	IWX
Ft. Campbell, KY	-87.2856	36.7370	172.00	10	05/21/1998	HPX
Evansville, IN	-87.7245	38.2603	155.75	30	12/16/2004	VWX

$$bias = \frac{Stage\ III\ or\ MPE}{PRISM} \tag{3.1}$$

where $bias = 1$ is perfect agreement.

Gridded HRAP monthly, seasonal, and annual precipitation totals are derived from OHRFC hourly *xmrg* (refer to Fulton [78]) format files.

Table 3.2: Annual OHRFC Stage III/MPE PPS bias statistics by year.

Year	Minimum	Maximum	Mean	Standard Deviation	Variance
1997	0	1.4334	0.6752	0.1134	0.0129
1998	0	1.3750	0.8176	0.1088	0.0118
1999	0	2.2144	0.7328	0.1319	0.0174
2000	0	1.2112	0.8800	0.0932	0.0087
2001	0	1.3814	0.9008	0.1102	0.0122
2002	0	1.3685	0.9382	0.1049	0.0110
2003	0.1913	2.0946	0.9150	0.1019	0.0104
2004	0.1912	1.6529	0.9266	0.1587	0.0252
2005	0.4562	1.7219	1.0100	0.1028	0.0106
2006	0.4018	1.5189	0.9961	0.1112	0.0124
2007	0.5480	1.3898	0.9790	0.1031	0.0106
2008	0.4411	1.5465	0.9884	0.1039	0.0108
2009	0.5606	1.3600	0.9863	0.0971	0.0094
2010	0.6004	1.4369	1.0005	0.1007	0.0101
2011	0.6494	1.4369	1.0421	0.0863	0.0074
2012	0.6757	1.9550	1.0539	0.0987	0.0097
2013	0.5886	1.6463	1.0965	0.0909	0.0083
2014	0.7285	1.4086	1.0936	0.0712	0.0051
2015	0.6361	1.2657	1.0319	0.0569	0.0032
2016	0.6718	1.3090	1.0216	0.0581	0.0034

3.3.3 Statistical methods

Analyses utilize *R* [145] verification measures and statistical analyses from the *verification* [135] and *hydroGOF* [199] contributed packages. For *R boxplots* we have,

$$IQR = Q_3 - Q_1 \tag{3.2}$$

$$UpperWisker = min(max(x), Q_3 + 1.5IQR) \tag{3.3}$$

$$LowerWisker = max(min(x), Q_1 + 1.5IQR) \tag{3.4}$$

where IQR is the interquartile range (box length in R boxplots) of a set of values, x, and Q_1 and Q_3 are, respectively, the 25th and 75th percentiles (lower and upper quartiles).

$$ME = \frac{1}{n}\sum_{k=1}^{n}(y_k - o_k) \tag{3.5}$$

$$PBIAS = 100 \cdot \frac{\sum_{k=1}^{n}(y_k - o_k)}{\sum_{k=1}^{n}(o_k)} \tag{3.6}$$

$$MAE = \frac{1}{n}\sum_{k=1}^{n}(|y_k - o_k|) \tag{3.7}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(y_k - o_k)^2} \tag{3.8}$$

$$NRMSE = 100 \cdot \frac{\sqrt{\frac{1}{n}\sum_{k=1}^{n}(y_k - o_k)^2}}{range} \tag{3.9}$$

With paired data, namely, predicted, y_i, and observed, x_i, we have $(x_1, y_1), \cdots, (x_n, y_n)$ and model,

$$Y_j = \beta_0 + \beta_1 x_j + \epsilon$$

where ϵ is random noise with $\mathbb{E}\epsilon = 0$ and $\mathbf{Var}\epsilon = \sigma^2$. No distribution for ϵ is assumed other than its mean is zero. It is noted that $\mathbb{E}Y_j = \beta_0 + \beta_1 x_j$ and $\mathbb{V}Y_j = \sigma^2$.

Least squares estimation. β_0 and β_1 are estimated by minimizing the sum of the squared errors:

$$\sum_{j=1}^{n}(y_j - \beta_0 - \beta_1 x_j)^2.$$

Consequently, we get:

$$SSE = \sum_{j=1}^{n} y_j^2 - \hat{\beta}_0 \sum_{j=1}^{n} y_j - \beta_1 \sum_{j=1}^{n} x_j y_j \tag{3.10}$$

$$SST = \sum_{j=1}^{n}(y_j - \bar{y})^2 \tag{3.11}$$

$$R^2 = 1 - \frac{SSE}{SST} \tag{3.12}$$

$$NSE = 1 - \frac{\sum_{t=1}^{T}(Q_m^t - Q_o^t)^2}{\sum_{t=1}^{T}(Q_o^t - \overline{Q_o})^2} \tag{3.13}$$

where we have the *Mean Error (ME)*, *Percent Bias (PBIAS)*, *Mean Absolute Error (MAE)*, *Root Mean Square Error (RMSE)*, *Normalized Root Mean Square Error (NRMSE)*, and *Coefficient of Determination (R^2)*, with quantities y_k and o_k the predicted and observed kth values, respectively, for n total paired values; $range = max(o_k : k = 1, \ldots, n) - min(o_k : k = 1, \ldots, n)$.

The *Nash-Sutcliffe Efficiency (NSE)* [131], for T periods, where Q_o is the observed discharge, Q_m is the modeled discharge, and Q_0^t is the observed discharge at time t, can range from $-\infty$ to 1. An efficiency of 1 (NSE = 1) corresponds to a perfect match of modeled discharge to the observed data. An efficiency of 0 (NSE = 0) indicates that the model predictions are as accurate as the mean of the observed data; values of NSE less than zero (NSE < 0) occurs when the observed mean is a better predictor than the model. Units of measure for *river flow* are $m^3 s^{-1}$, unless reported otherwise. Values for *ME*, *MAE*, and *RMSE* = 0 implies perfect agreement, i.e., no error.

Annual variability

Making use of Equation 3.1, we obtain Figure 3.4, which shows the spatial bias pattern of OHRFC Stage III and MPE precipitation estimates on an annual basis from January 1, 1997 through December 31, 2016. Two features should be evident, namely that (1) Stage III and MPE precipitation estimates are significantly under-estimated with respect to the PRISM estimates beginning in 1997, but improve significantly by 2002; and (2) the character of the spatial bias pattern changes from an *apparent* random variation (1997-2001) to one that exhibits distinct polygonal artifacts (2002-2011), to a pattern showing more of a random character (2012-2016). The changes to the bias patterns can be directly attributed to changes in the method used for bias correction initially in Stage III (1997-2001), then MPE (2002-2011), and finally with MPE utilizing initial MRMS estimates (2012-2016). There are also clear indications of persistent beam blockage in the MPE estimates (2002-2011), which are greatly reduced when MRMS is introduced to the OHRFC PPS late in 2012. Bias variations viewed as an annual series in Figure 3.5(a) using boxplots (see Equations 3.2 to 3.4) show marked bias reduction from 1997-2005. With the introduction of the use of MRMS in late 2012, the variance in bias is substantially reduced over previous years and by 2015-2016

median biases are close to 1. Results presented in Figure 3.5(b) for the Greenbrier River basin, WV, identified in Figure 3.2, are discussed below.

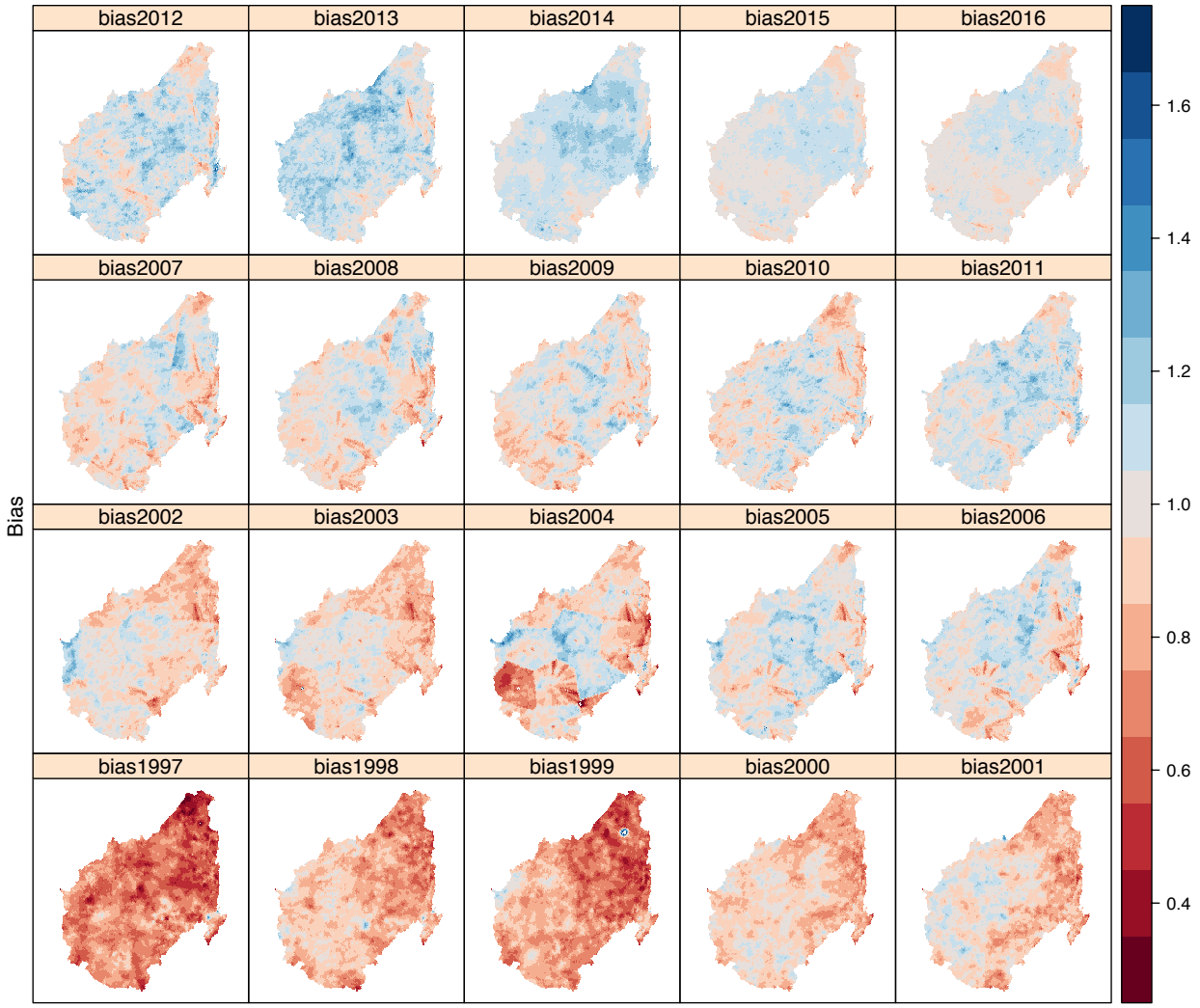

Figure 3.4: Spatial pattern of *Stage III/MPE* precipitation estimate biases with respect to PRISM over the OHRFC forecast area of responsibility, 1997-2016

The change in annual Stage III/MPE bias is further illustrated in Figure 3.6 with bias density plots, which shows a shift in mean bias, with significant underestimation in 1997, to nearly unbiased estimates by 2015-2016, and a consistent reduction in bias spread from 1997 to 2016.

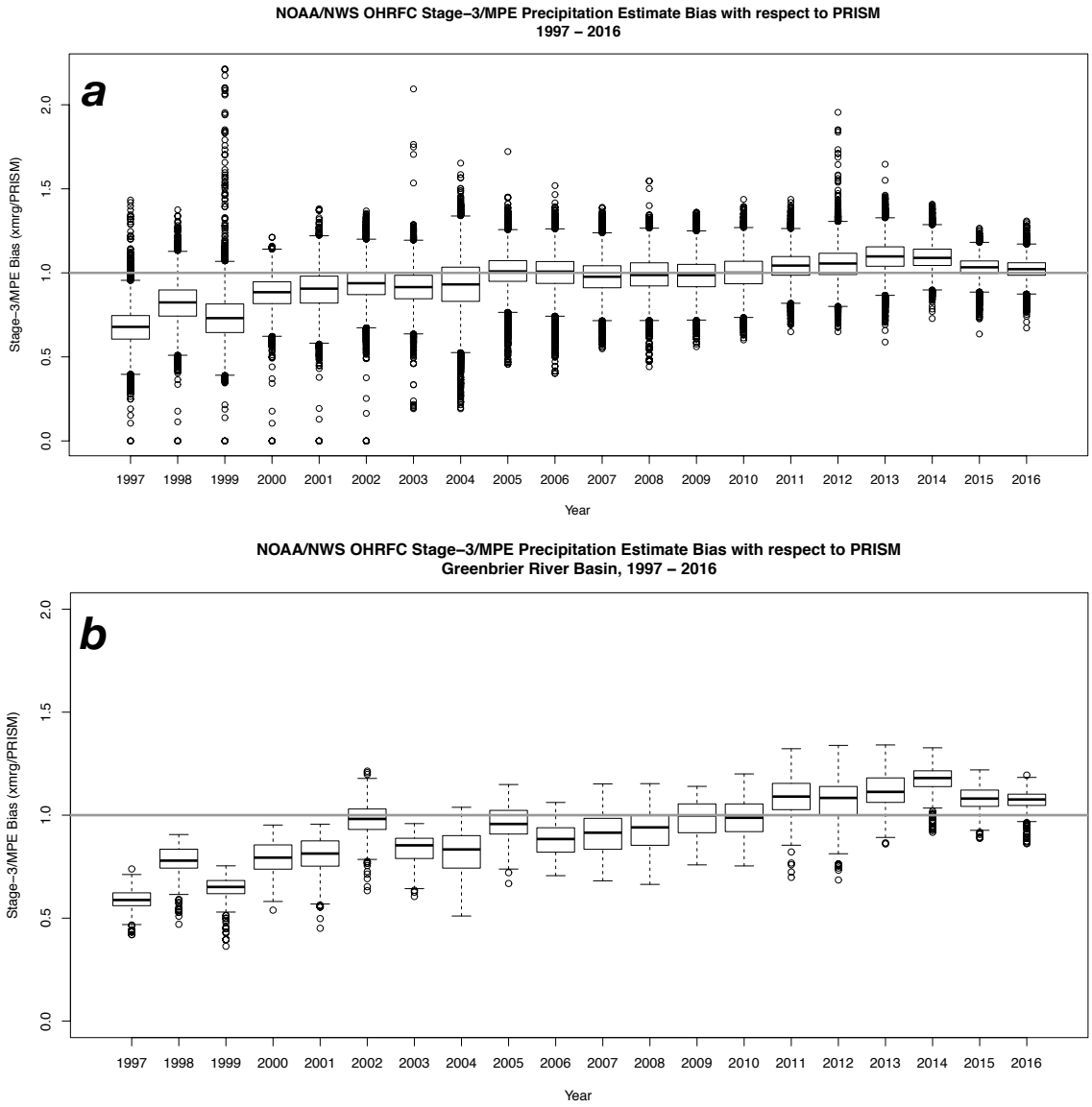

Figure 3.5: Annual time-series of *Stage III/MPE* precipitation estimate biases with respect to PRISM over the (a) OHRFC forecast area of responsibility and (b) the Greenbrier River basin at Alderson, West Virginia (see Figure 3.2), 1997-2016. The horizontal gray line is used for reference with $bias = 1$.

Seasonal variability

Seasonal variability of MPE biases is presented for the years 2015 and 2016 only because these years reflect current *best estimates* and PPS performance. Summer months are defined as June-July-August (JJA) and winter months as December-January-February (DJF). Seasonal bias values are calculated as monthly JJA and DJF averages using Equation 3.1. Figure 3.7 shows the summer and winter spatial bias patterns for 2015 and 2016. Bias patterns during the summer months exhibit more of a random pattern compared to the winter months for 2015-2016. These differences are expected due to the prevalence of more isolated, convective rainfall during the summer months and wide-spread, stratiform and synoptic-scale precipitation during the winter months. Also evident during the 2015 and 2016 winter seasons is widespread MPE over-estimation in central and east-central regions of the OHRFC forecast area, quite possibly due to *brightband* influences (Gourley and Calvert [85]; Cunha et al. [44]).

Clear differences are seen between winter and summer season biases for the 2015-2016 period in Figure 3.8 with bias density plots. Summer biases are very close to 1 and, while winter season biases are not much different from 1, with slight over-estimation, the spread in bias values is much greater during the winter season compared to summer.

Monthly variability

The analysis of monthly MPE biases, using Equation 3.1, is based on monthly accumulations of hourly *xmrg* HRAP gridded fields. The bias analysis is restricted to the years 2015 and 2016. Figure 3.9 shows significant month-to-month bias variability. Monthly median biases are very close to unity, *interquartile* (25 to 75 percentile) differences are generally small, but large outliers are evident. This points to the complex nature of both precipitation pro-

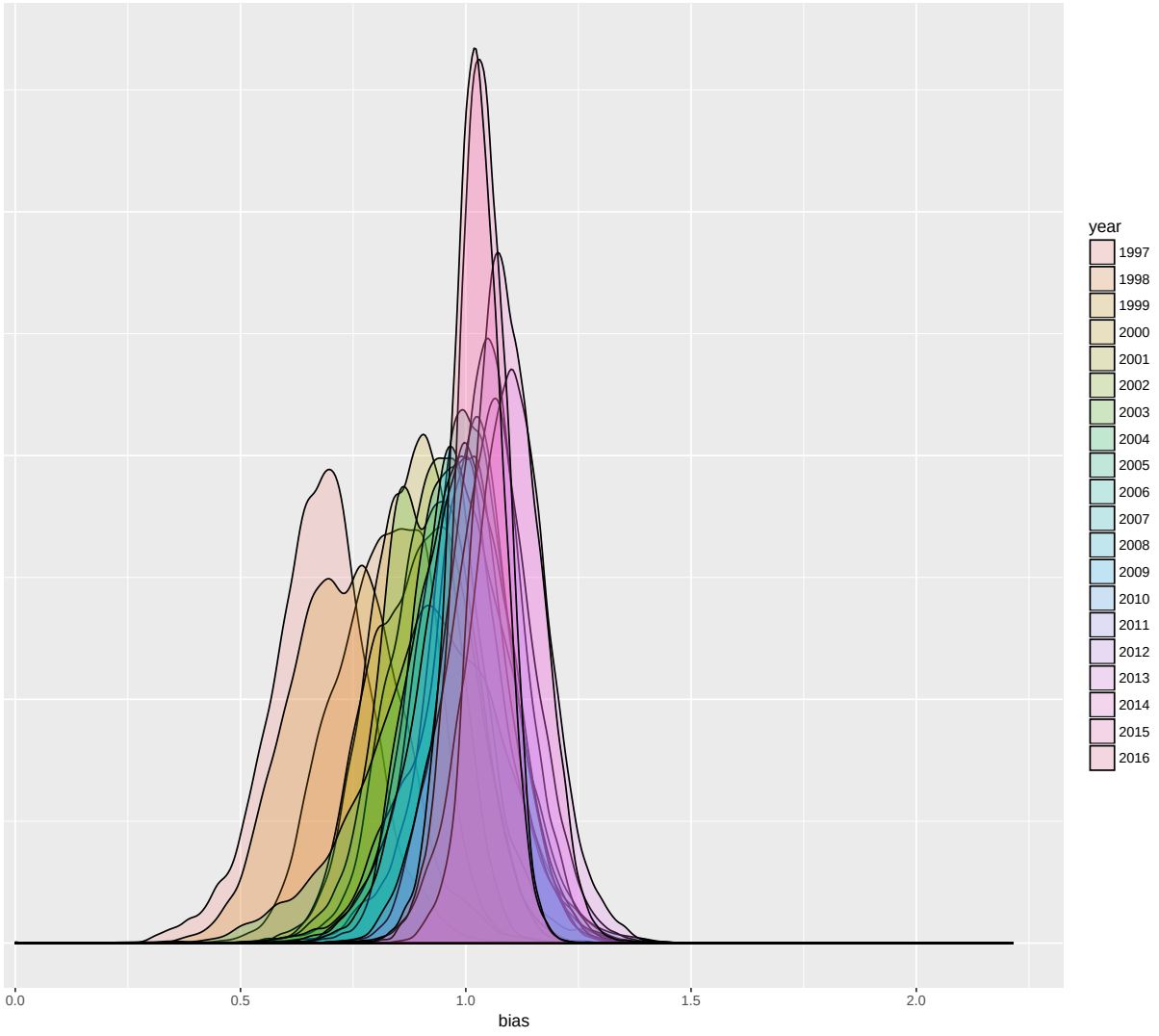

Figure 3.6: OHRFC *Stage III/MPE* precipitation estimation bias density by year with respect to *PRISM* for 1997-2016.

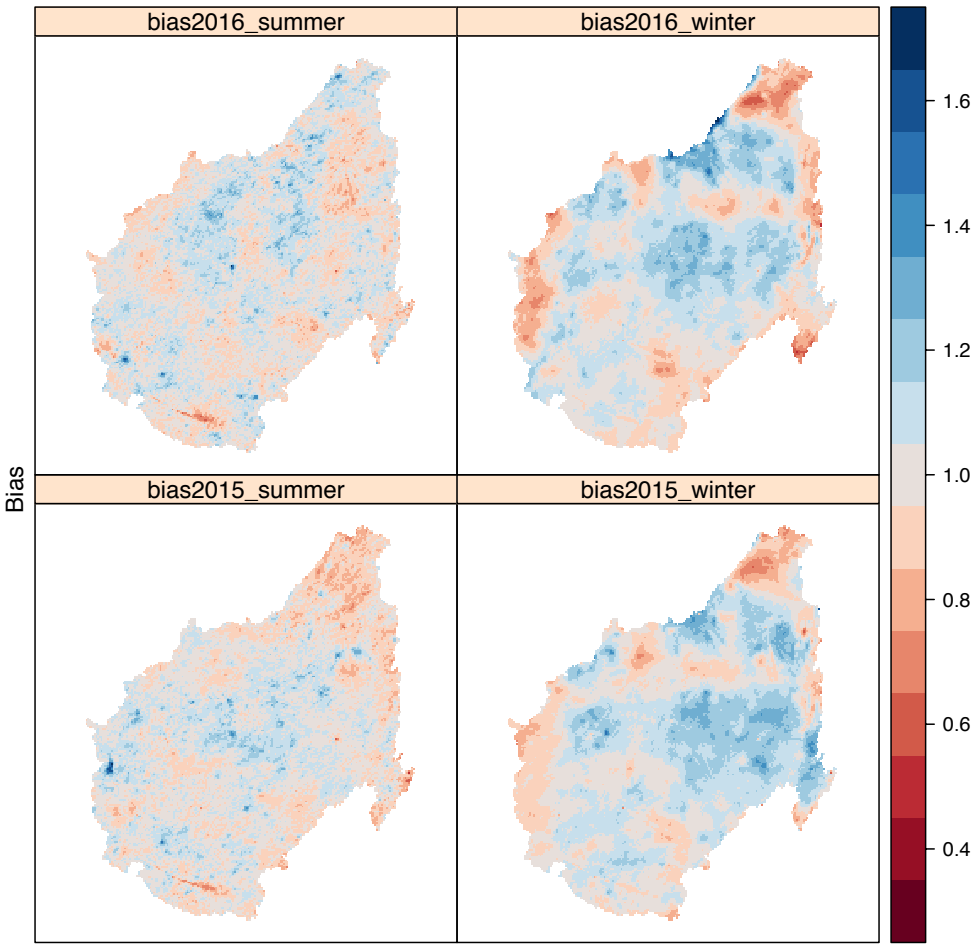

Figure 3.7: OHRFC spatial pattern of *MPE* precipitation estimation bias by season, summer (JJA) and winter (DJF), with respect to *PRISM* for 2015-2016.

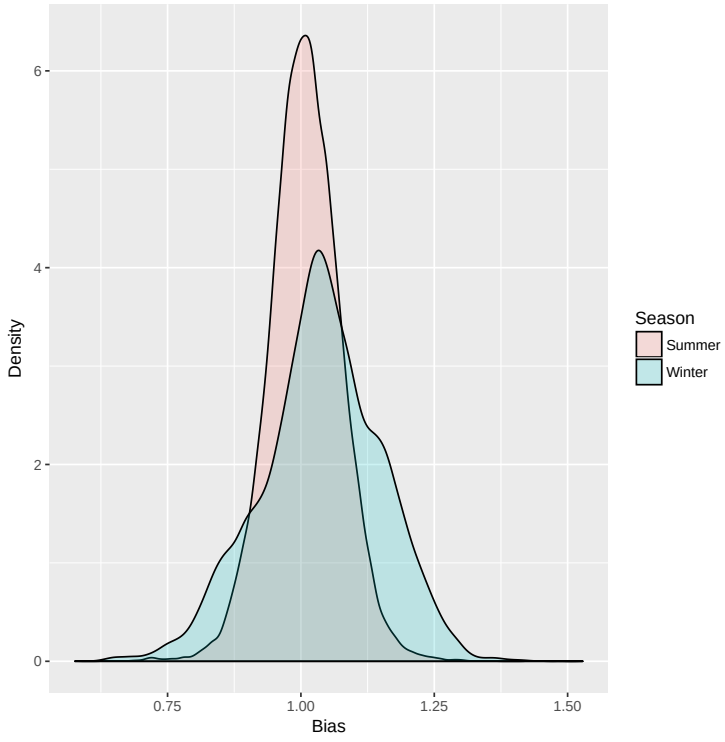

Figure 3.8: OHRFC *MPE* precipitation estimation bias density by season, summer (JJA) and winter (DJF), with respect to *PRISM* for 2015-2016.

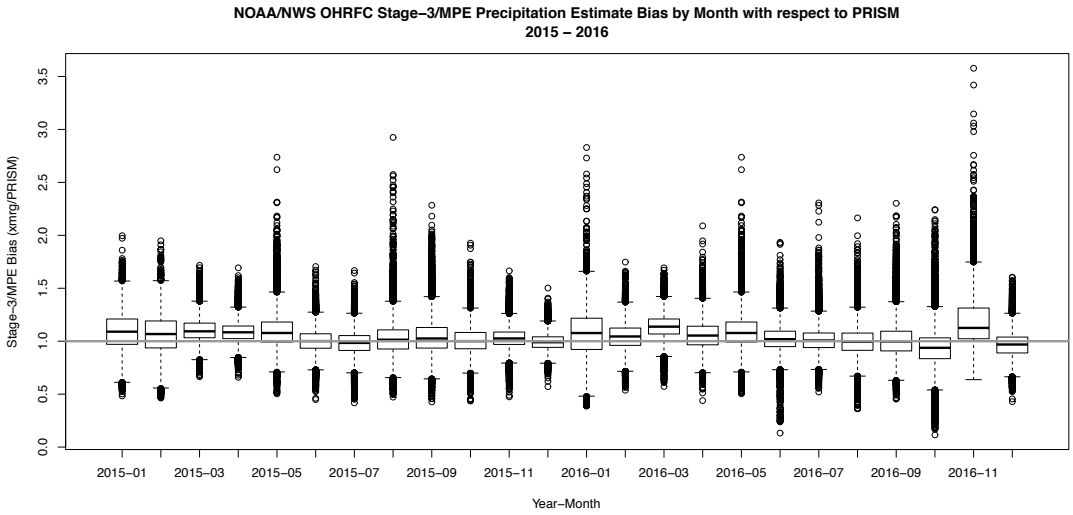

Figure 3.9: Monthly time-series of *MPE* biases with respect to PRISM over the OHRFC forecast area of responsibility, 2015-2016. The horizontal gray line is used for reference with $bias = 1$.

cesses and difficulty in removing estimation biases at short time scales, suggesting continued challenges with correcting estimation biases in real-time.

Hydrologic modeling impacts

We desire knowing what benefits, if any, to hydrologic modeling and forecasting is evident from improvements in Stage III and MPE precipitation estimation. We address this issue by using the Stage III/MPE precipitation estimates and hourly temperature data obtained from the North American Land Data Assimilation System (NLDAS-2) (Xia et al. [193]; Xia et al. [194] as the primary hydrologic model forcings for simulations spanning the historical period of record, 1997-2016.

Historical simulation To illustrate the benefits gained from precipitation estimation improvements, a retrospective hydrologic simulation using the NOAA/NWS Hydrology Laboratory *Research Distributed Hydrologic Model (HL-RDHM)* (Koren et al. [106]; Koren [104];

Koren et al. [107]) is made. The historical simulation spans the full period (1997-2016) of available Stage III/MPE, NEXRAD radar derived precipitation estimates. RDHM simulations are made for the Greenbrier River basin in West Virginia, shown in Figure 3.2, with the model defined at the HRAP grid resolution, using an hourly time-step. The Greenbrier River basin defined at Alderson, WV (USGS 05050003), located in the *valley and ridge* physiographic province of the Appalachian Mountains, has an area of 3533 km^2, and ranges in elevation from 466 to 1433 m. The basin was selected for study largely because of recent major flooding in June 2016 and data availability. Retrospective RDHM simulations are made without prior calibration to avoid biasing model performance to any model period. In this way, model results will best reflect PPS improvements implemented at the OHRFC without confounding influences from other factors. Initial, *a priori*, estimation of RDHM *Sacramento Soil Moisture Accounting (SAC-SMA)* (Burnash et al. [34]; Burnash [33]) model parameters is described by Koren et al. [105]. Parameter estimation includes 12 SAC-SMA parameters and several channel routing parameters. RDHM simulations also include use of the SNOW-17 snow accumulation and ablation model to account for wintertime precipitation and snowmelt. SNOW-17 model parameters are also estimated and the model is used uncalibrated.

RDHM model simulations begin June 1, 1996 from a *cold state*, that is, without prior model state initialization that reflect existing basin conditions, using hourly NLDAS-2 precipitation and temperature model inputs. OHRFC Stage III and MPE are used for the full period January 1, 1997 to December 31, 2016. Verification of RDHM simulations is restricted to the January 1998 to December 2016 period, to allow sufficient RDHM SAC-SMA model warm-up, nearly 18 months.

Table 3.3: *Goodness-of-fit* statistics for the uncalibrated RDHM historical simulation and USGS observed flows for the Greenbrier River at Alderson, WV (USGS 05050003), by year, for the period 1998-2016. Unless indicated otherwise, units are m^3s^{-1}, except for *NSE* and *R2*, which are dimensionless.

Year	ME	MAE	RMSE	NRMSE (%)	PBIAS (%)	NSE	R2
1998	-30.02	32.32	76.97	75.10	-48.60	0.44	0.61
1999	-23.62	24.16	51.51	90.00	-67.50	0.19	0.58
2000	-17.69	19.29	52.35	76.10	-36.80	0.42	0.53
2001	-18.18	20.23	57.46	78.70	-43.90	0.38	0.52
2002	-13.05	24.92	61.07	68.50	-21.40	0.53	0.56
2003	-38.94	47.25	100.83	75.30	-36.60	0.43	0.53
2004	-31.66	33.63	68.87	76.80	-43.70	0.41	0.60
2005	-15.73	17.67	38.81	63.10	-33.20	0.60	0.70
2006	-11.68	16.10	44.35	65.80	-24.70	0.57	0.63
2007	-11.20	19.74	61.76	62.20	-21.40	0.61	0.65
2008	-1.53	19.90	54.20	64.20	-3.10	0.59	0.59
2009	-3.90	21.56	41.74	59.80	-6.50	0.64	0.65
2010	2.16	25.03	61.00	56.20	4.40	0.68	0.72
2011	0.67	23.37	58.55	60.30	1.10	0.64	0.64
2012	-3.97	13.57	38.16	55.50	-9.10	0.69	0.70
2013	-21.18	27.32	68.11	67.50	-33.00	0.54	0.61
2014	16.21	26.49	50.90	67.90	33.70	0.54	0.63
2015	9.69	28.58	60.21	51.90	13.80	0.73	0.75
2016	6.34	24.21	50.68	40.30	10.00	0.84	0.84

Simulation results Using USGS measured discharges and RDHM historical simulation, goodness-of-fit statistics, such as *Nash-Sutcliffe Efficiency (NSE)*, ME, RMSE, R^2, MAE, *Normalized Root Mean Square Error (NRMSE)*, and *Percent Bias (PBIAS)* are calculated, using Equations 5.1 through 3.13, to assess hydrologic modeling improvement based on MPE improvements. RDHM historical simulation results are shown as a hydrograph in Figure 3.10 for a representative period, October 2015 - June 2016. Analyses of the historical flows simulation compared to USGS measured flows are reported in Table 3.3. The statistical results presented in Table 3.3 for the 1998-2016 simulation period are summarized graphically in Figure 3.11.

With the exception of Mean Absolute Error (MAE), which is relatively unchanged at about 25 m^3s^{-1} and Root Mean Square Error (RMSE) at about 55 m^3s^{-1}, all other measures indicate improvement between observed and simulated flow values over the 1998-2016 retrospective simulation period. Improvements in values for Mean Error (ME) (from -30.02 to 6.34), Nash-Sutcliffe Efficiency (NSE) (from 0.44 to 0.84), Coefficient of Determination (R^2) (from 0.61 to 0.84), and Percent Bias (PBIAS) (from -48.60 to 10.00) are notable.

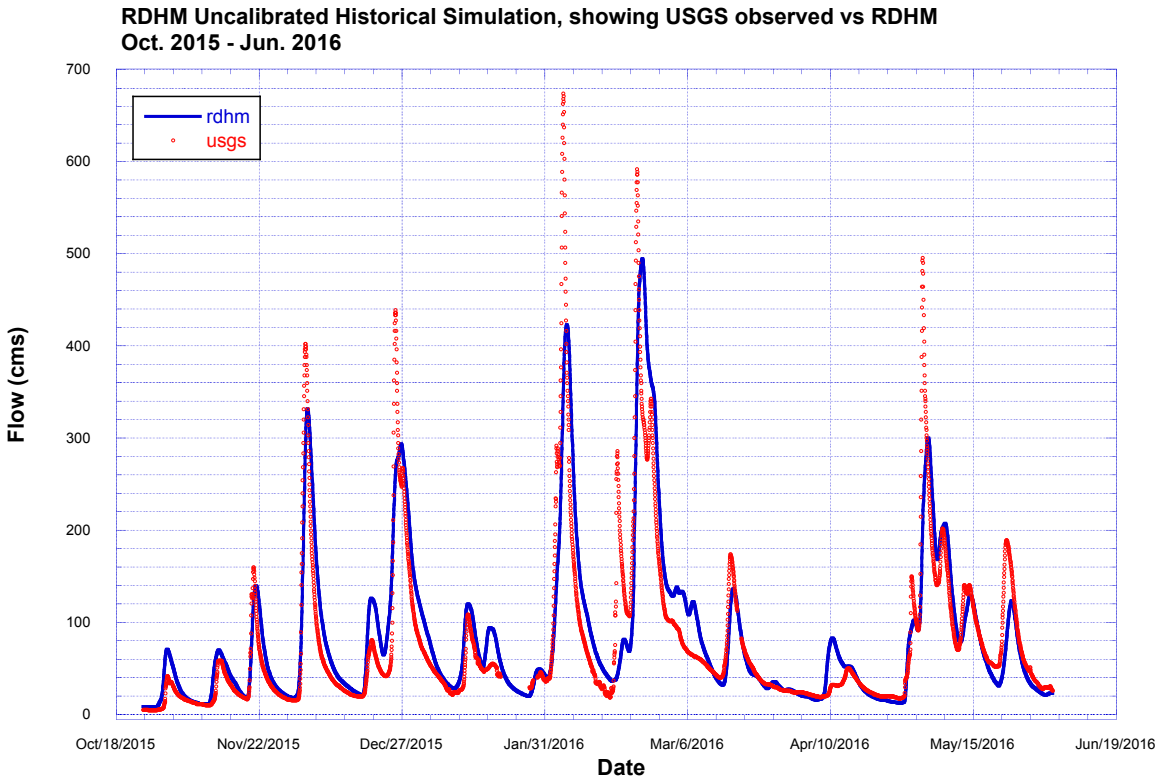

Figure 3.10: RDHM uncalibrated historical simulation for the Greenbrier River at Alderson, WV, 1997-2016, compared against USGS observed flows for the period October 18, 2015 to June 19, 2016.

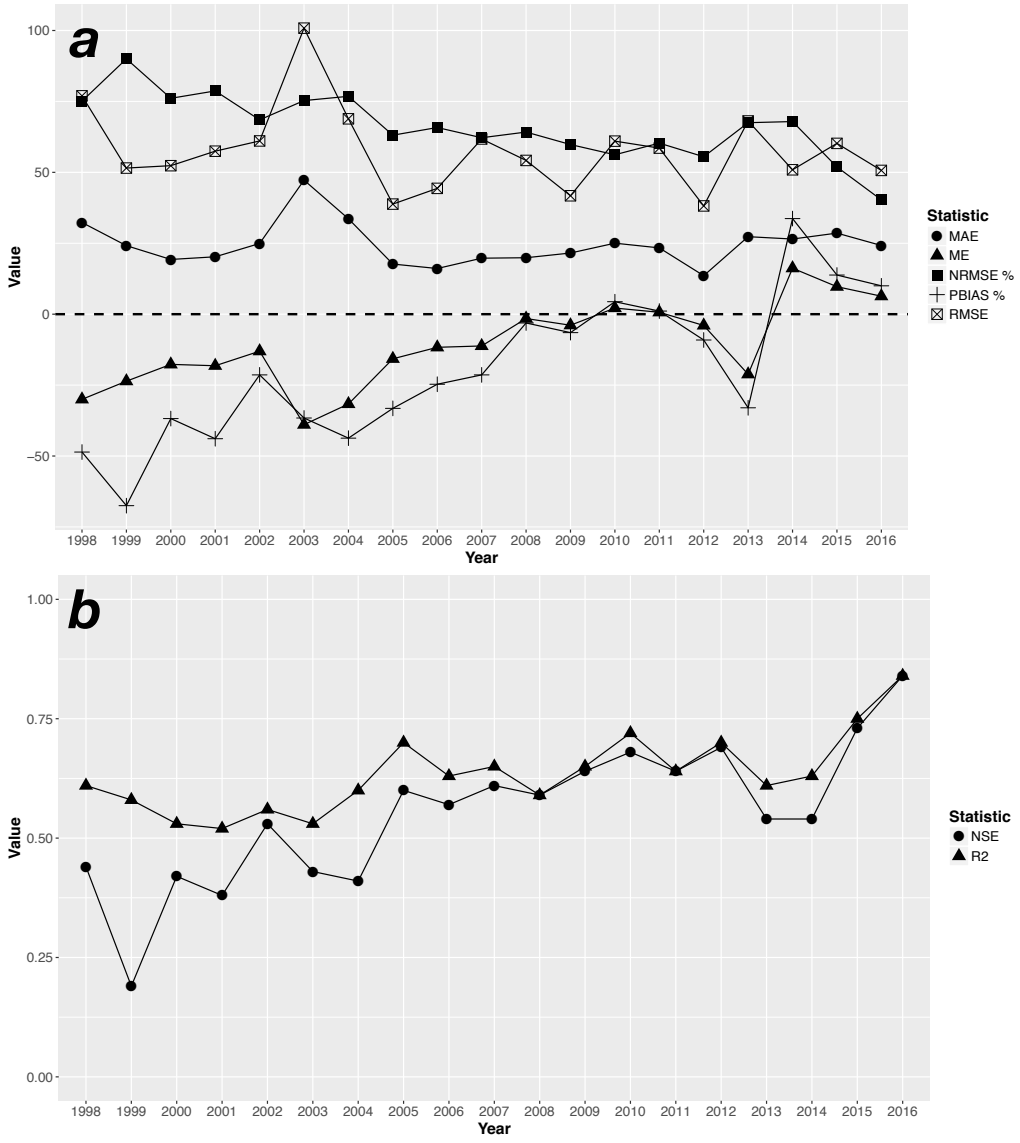

Figure 3.11: Mean annual goodness-of-fit statistics for the uncalibrated RDHM historical simulation and USGS observed flows, (a) MAE (m^3s^{-1}), ME (m^3s^{-1}), NRMSE (%), PBIAS (%), and RMSE (m^3s^{-1}) and (b) NSE and Coefficient of Determination (R2), for the Greenbrier River basin, for years 1998-2016.

Table 3.4: Contingency table for QPF *Threat Score* calculation.

	Observed	
Forecast	yes	no
yes	a	b
no	c	d

Consequently, the results demonstrate that improvements in QPE estimation has lead to significantly improved hydrologic simulations for the Greenbrier River basin in the OHRFC area of forecast responsibility. However, QPF is another major model forcing used in hydrologic forecasting. In section 3.4 we examine QPF error and the magnitude of hydrologic forecast error in response to QPF uncertainty.

3.4 QPF errors

Novak et al. [139] report improved NOAA/NWS *Weather Prediction Center* QPF performance from 1960 through 2012 for days 1, 2, and 3 lead-time, 24-h, 1 in (25.4 mm) forecasts. Methods used for forecaster generation of WPC QPF and comparisons relative to various numerical weather prediction (NWP) models are also presented. WPC QPF improvement is measured in terms of *threat score (TS)*, given by Wilks [185], using Table 5.5, is:

$$TS = \frac{a}{a+b+c} \tag{3.14}$$

A 24-h, 1 in (25.4 mm) precipitation threshold is, in most instances, too low to identify flood producing events. Consequently, an analysis of higher intensity, flood-producing events is warranted. WPC monthly QPF TS data for accumulations $\geq$2 in (50.8 mm) is likely to be more relevant to addressing QPF performance relative to meeting hydrologic flood

forecasting needs. Two important issues are, (1) the degree to which errors in operational QPF influence hydrologic prediction and (2) how these influences can be quantified. In an attempt to address these concerns, results from a hydrologic modeling study are presented, noting that:

1. We present QPF verification statistics taken from the NOAA/NWS *National Precipitation Verification Unit (NPVU)* covering the 12 *Conterminous U.S. (CONUS)* RFCs;

2. Results from *Monte Carlo* simulations using the RDHM for the Greenbrier River basin, WV are presented spanning the June 22-24, 2016 flooding episode to assess the range of hydrologic errors in response to expected WPC QPF accuracy;

3. *Threat Score (TS)*, a commonly used statistical measure of forecast accuracy in meteorology, will serve as the basis for evaluating QPF accuracy in the RDHM hydrologic experiments.

3.4.1 WPC

Monthly mean bias and TS data for precipitation accumulations $\geq$2 in (50.8 mm), obtained from NOAA/NWS WPC [148], are summarized in Figure 3.12 for the period 1970-2015 for Day-1 and 1991-2015 for Day-2 lead-times. Day-2 data were not collected before 1991 by WPC. Annual averages are also shown, based on monthly averaged values. The large variability of monthly bias and TS values is evident. We note that expected Day-1, $\geq$2 in (50.8 mm) TS values are about 0.06 in 1970 and 0.22 by 2015.

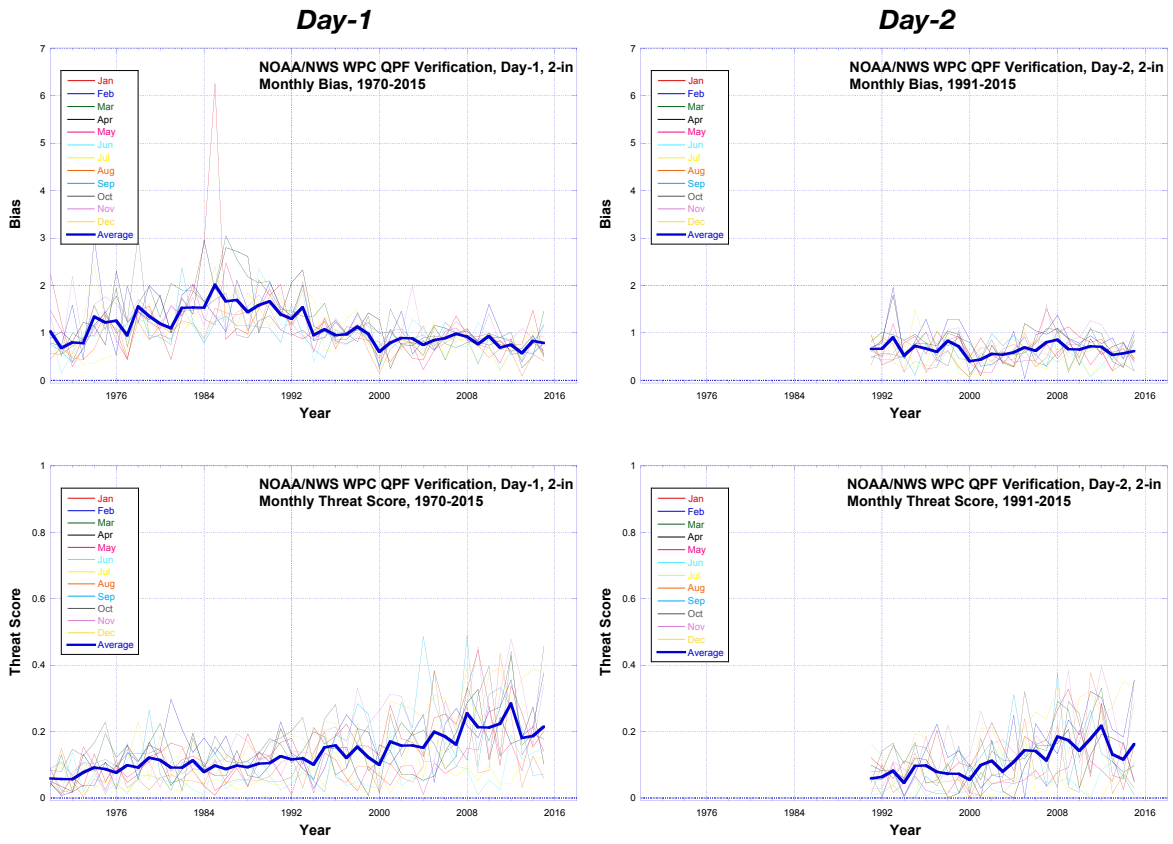

Figure 3.12: WPC monthly and annual average QPF Bias and Threat Score, by year, for Day-1 (1970-2015) and Day-2 (1991-2015), for accumulations≥2.00 in (50.8 mm).

3.4.2 NPVU

Charba et al. [37] describe a methodology for evaluating QPF forecast accuracy relative to RFC produced Stage III and MPE QPE at WPC, RFCs, and *Weather Forecast Offices (WFOs)* compared against national guidance produced by operational numerical weather prediction (NWP) models run at the National Centers for Environmental Prediction (NCEP). Also described is the implementation of the *National Precipitation Verification Unit (NPVU)* to carry-out national QPF verification from the various sources (WPC, WFOs, RFCs, NWP models). NPVU QPF verification data was produced for the period 2001-2012.

Figure 3.13 shows a time series of NPVU monthly (a) *Mean Absolute Error (MAE)* and (b) *Root Mean Square Error (RMSE)* for all NOAA/NWS CONUS River Forecast Centers (**rfc**), NOAA/NWS Nested Gridded Model (**ngm**), North American Model (**nam**), Hydrometeorological Prediction Center (**hpc**) – now Weather Prediction Center (WPC) – Global Forecast System (**gfs**), ETA Model (**eta**), and Aviation Model (**avn**) for the period June 2001 to December 2009. The gap in data from late 2003 to mid-2004 was due to an NPVU data processing failure. Several points are notable (1) seasonal variability for all the QPF sources is clear; (2) while there are differences between QPF sources, there is relatively little discernible MAE or RMSE improvement over the June 2001 to December 2009 analysis period displayed by any of the QPF sources; and (3) for all QPF sources there is less QPF error for the smaller precipitation intervals, $\geq$0.01 in (0.254 mm), than the larger intervals, $\geq$1.0 in (25.4 mm). The highest precipitation interval used by NPVU is $\geq$1.0 in (25.4 mm).

Figure 3.14 shows the *correlation coefficient, R*, of QPF versus QPE, aggregated across the June 2001 to December 2009 NPVU analysis period, grouped by QPF source and precipitation interval. It is evident, as reported by Charba et al. [37], that forecasters add value over NWP modeled QPF, based on HPC (WPC) and RFC results compared to NWP model

results. Also apparent is that for 24-h accumulations, especially for the larger precipitation intervals, ≥1.0 in (25.4 mm), QPF is poorly correlated with observed QPE and that there is considerable correlation spread within the precipitation intervals. Little difference is apparent between HPC and RFC QPF on the basis of R values and spread.

Implications

We can see that deterministic QPF can be quite erroneous, based on a range of statistical verification measures, demonstrated by both NWP model generated QPF and QPF produced with the aid of forecaster input. It has been reported [139], correctly, that by some statistical measures, specifically, *threat score*, QPF accuracy has improved since the beginning of systematic record keeping to the present. Nevertheless, the question must be asked, what benefits have accrued to hydrologic forecasting from QPF improvements? To answer this question a hydrologic simulation experiment is presented using the near-record Greenbrier River basin (defined at Alderson, WV, USGS 05050003) flooding event in West Virginia, June 22-24, 2016.

3.4.3 Hyrdrologic simulation experiments

A hydrologic simulation experiment is used to assess how QPF improvements have impacted hydrologic prediction. Simulations take the form of a *monte carlo* experiment, using a storm transposition methodology (Foufoula-Georgiou [76]; England et al. [68]; Wright et al. [192]). The aim of the experiment is to illustrate the inadequacy of current deterministic QPF for hydrologic forecasting purposes. The experiment first produces 1000 randomly located storm centers based on the maximum 24-h MPE grid cell precipitation accumulation for the period beginning June 23, 2016 0600 UTC within the region, shown in Figure 3.15 (a). An example

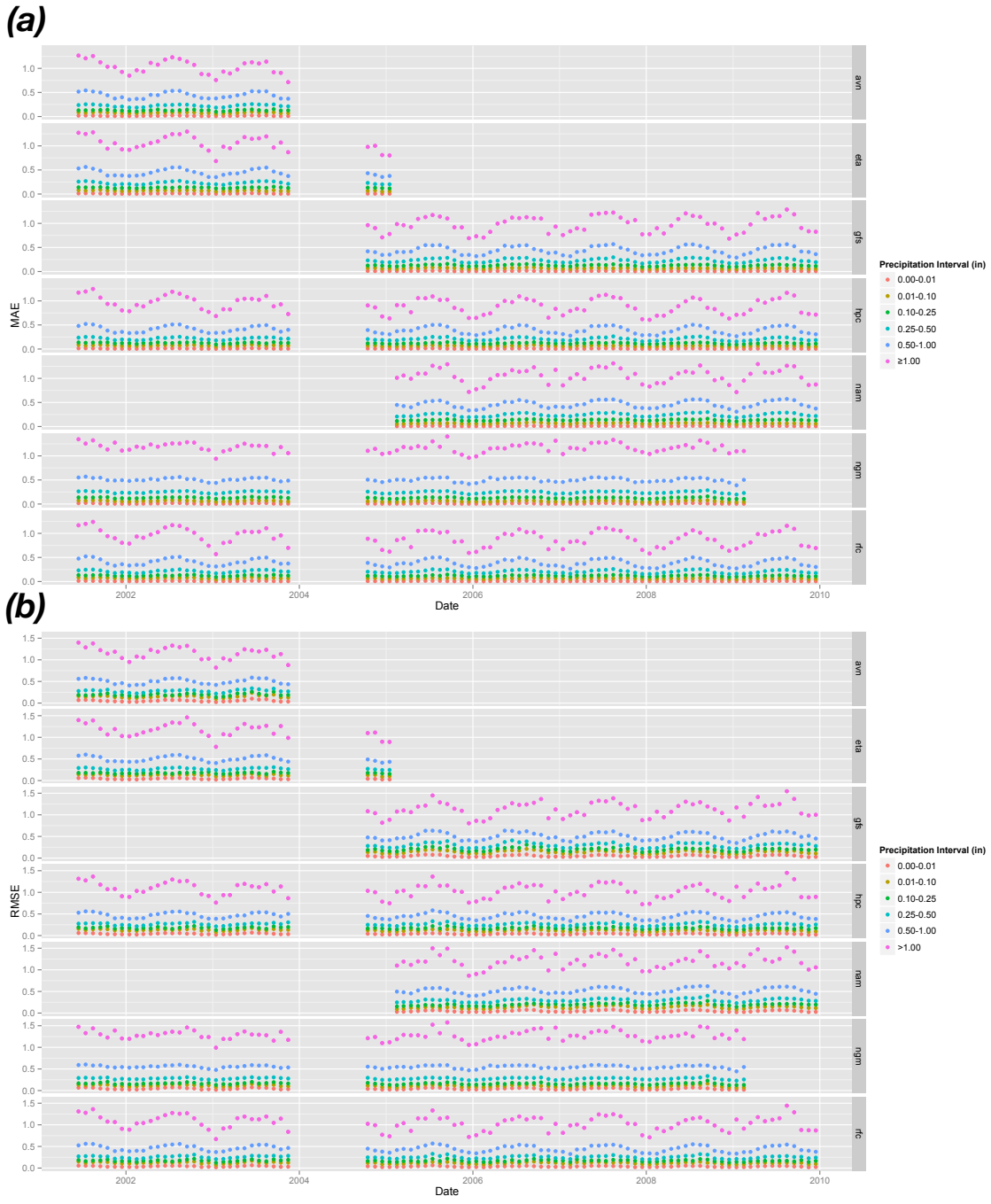

Figure 3.13: NPVU (a) *Mean Absolute Error (MAE)* and (b) *Root Mean Square Error (RMSE)*, by month, for QPF thresholds, ranging from ≤0.01 (0.254 mm) to ≥1.00 in (25.4 mm), for the period June 2001 to December 2009 for all NOAA/NWS CONUS River Forecast Centers (**rfc**), NOAA/NWS Nested Gridded Model (**ngm**), North American Model (**nam**), Hydrometeorological Prediction Center (**hpc**) – now Weather Prediction Center (WPC) – Global Forecast System (**gfs**), ETA Model (**eta**), and Aviation Model (**avn**).

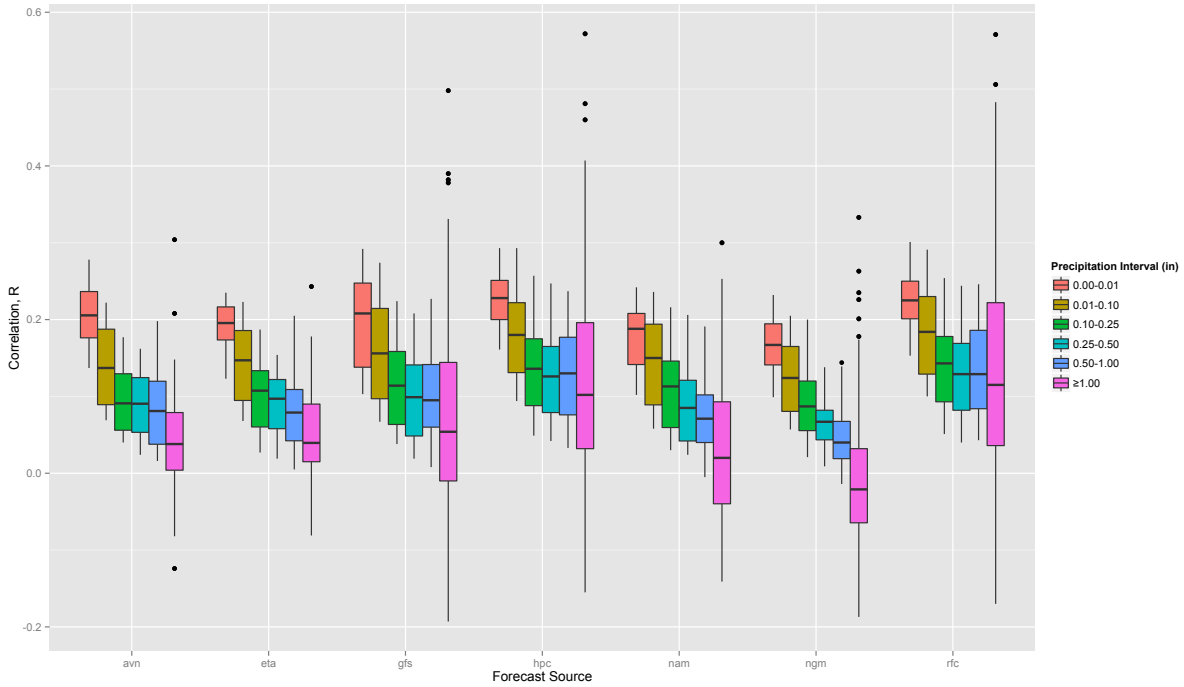

Figure 3.14: NPVU *Correlation Coefficient (R)* for QPF thresholds, ranging from ≤0.01 (0.254 mm) to ≥1.00 in (25.4 mm), for the period June 2001 to December 2009 for all NOAA/NWS CONUS River Forecast Centers (**rfc**), NOAA/NWS Nested Gridded Model (**ngm**), North American Model (**nam**), Hydrometeorological Prediction Center (**hpc**) – now Weather Prediction Center (WPC) – Global Forecast System (**gfs**), ETA Model (**eta**), and Aviation Model (**avn**).

transposed storm is shown in Figure 3.16, which shows that all details of the observed MPE field are retained. The June 23, 2016 0600 UTC to June 24, 2016 0600 UTC period was selected because this was the greatest MPE precipitation accumulation for a 24-h *synoptic* period during the flooding event. The transposed individual hourly grids of the 24-h MPE precipitation serve as QPF for the RDHM simulations. Figure 3.15 (b) shows the location of the maximum 24-h MPE grid cell precipitation accumulation, as well as some of the 1000 randomly transposed storm centers closest to the storm center maximum. Only precipitation grid cells with MPE accumulations ≥50.8 mm (2.0 in) are shown in Figure 3.15 (a) and (b).

TS values are calculated for each of the randomly transposed 24-h storms using Equation 3.14 and Table 5.5 based on precipitation amounts ≥50.8 mm (2.0 in). From these, only storms with TS≥0.06 are used in the RDHM *monte carlo* simulations. Consequently, 88 randomly transposed storms, shown in Figures 3.15 (a) and (b), were identified and are used in the RDHM *monte carlo* simulation experiment. The RDHM was initialized from model *warm states*, that is, model states generated from the 1996-2016 historical simulation discussed previously. However, for the purpose of the *monte carlo* simulation experiment only, the RDHM was calibrated for the June 22-24, 2016 event to produce good agreement between observed and simulated peak flows using the observed MPE precipitation. The hydrological context of the experimental results with respect to USGS observed flows is improved by using a calibrated model.

Monte Carlo experiment results

Figure 3.17 shows the USGS observed flows, simulated RDHM flow hydrograph from the calibration, and simulated flows from the transposed storms with TS values ranging 0.06-0.15 and 0.15-0.25. These TS ranges were used to identify reasonable TS value ranges, 0.06-0.15 and 0.15-0.25, that reflect WPC QPF skill for Day-1, ≥50.8 mm (2.0 in) precipitation

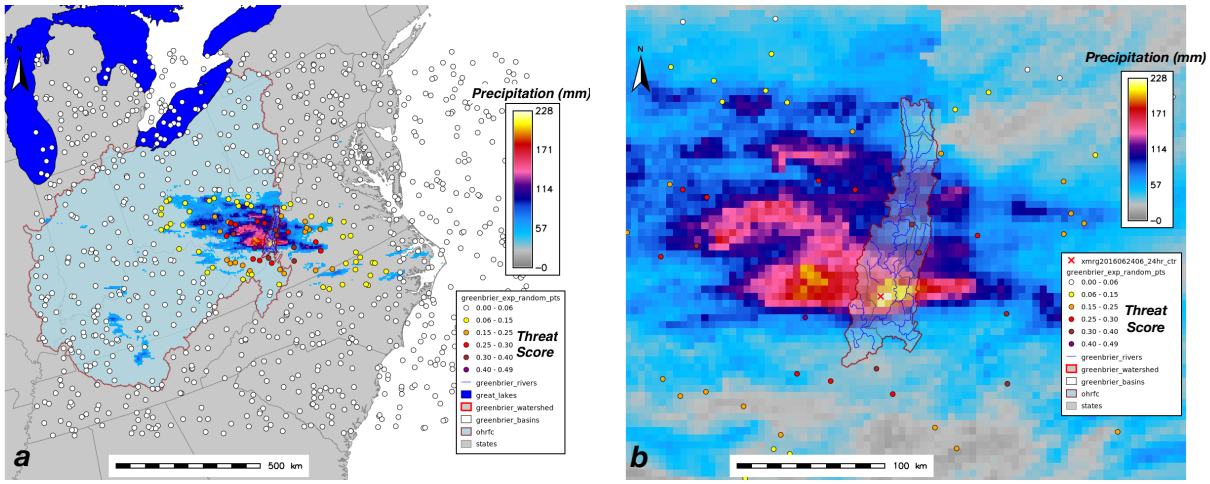

Figure 3.15: OHRFC forecast area of responsibility (a) (blue shading) showing 1000 randomly generated locations for QPF transposition of the 24-h precipitation accumulation for amounts $\geq$50.8 mm (2.0 in) from June 23, 2016 07 UTC to June 24, 2017 06 UTC. Points identifying transposition locations with Threat Scores $\geq$0.06 are colored yellow to purple; values $<$0.06 are filled white. A closer view (b) shows the reference location, used for storm transposition (identified with a red cross), which is the location of the maximum 24-h precipitation.

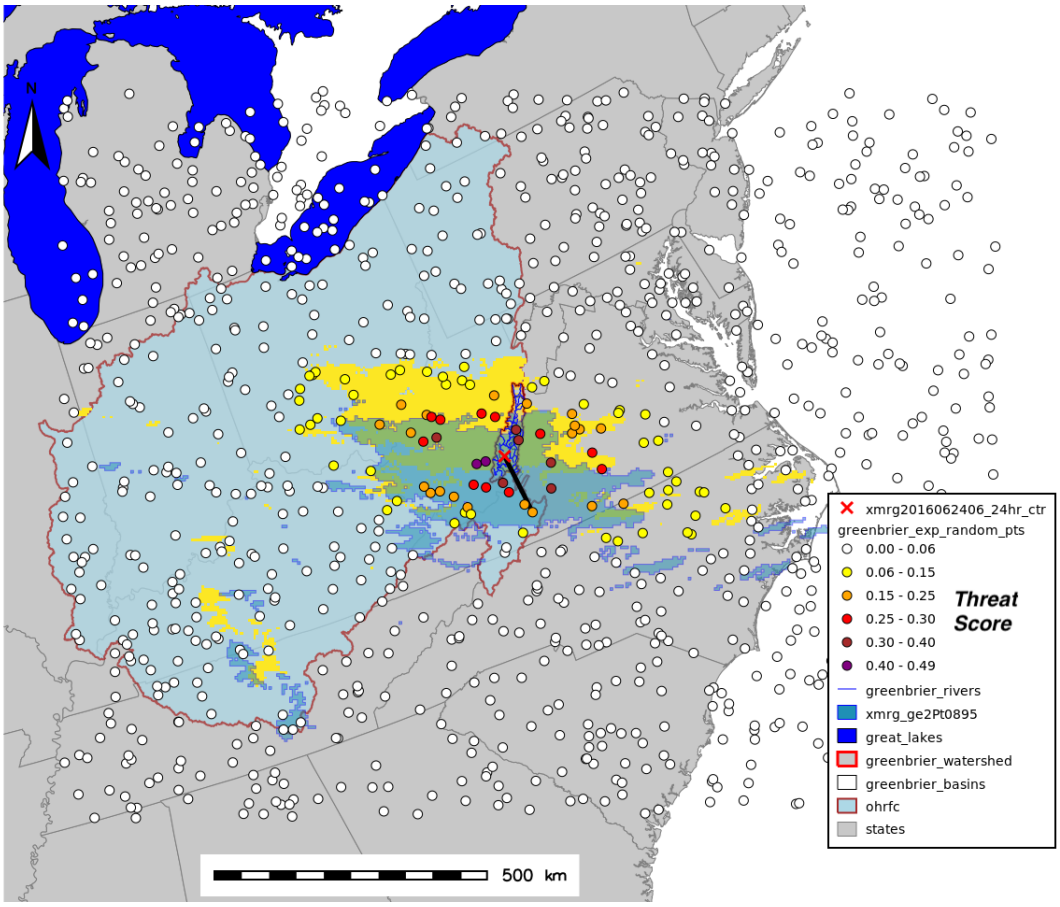

Figure 3.16: Example of a transposed storm (shaded blue) relative to the observed MPE storm (yellow); the green region shows overlap between the observed MPE and transposed storm. Also shown are the OHRFC forecast area of responsibility (light blue shading) and 1000 randomly generated locations for QPF transposition of the 24-h precipitation accumulation for amounts ≥50.8 mm (2.0 in) from June 23, 2016 07 UTC to June 24, 2017 06 UTC. Points identifying transposition locations with Threat Scores ≥0.06 are colored yellow to purple; values <0.06 are filled white. The reference location, identified with a red cross, is the location of the maximum 24-h precipitation, from which storm transpositions are made. The heavy black line indicates the transposition vector.

accumulations in 1970 and 2016, respectively. By doing this we believe we can assess, in relative terms, the gain in hydrologic forecast accuracy reflected by improvements in WPC QPF. It is evident from the Figure 3.17 simulations that hydrograph peaks are, in general, greater for storms with TS values ranging 0.15-0.25 than for storms with TS values 0.06-0.15. However, none of the simulated hydrographs exceed the *Major Flood* level and none approach the observed near-record flood peak level. Moreover, there is considerable variability between the simulated hydrographs within the separate 0.06-0.15 and 0.15-0.25 TS categories.

A better perspective on the problem inherent with using deterministic QPF in hydrologic forecasting is found by looking at simulation results from all 88 storms used in the RDHM *monte carlo* experiment with TS$\geq$0.06. Figure 3.18 shows peak flow and storm TS relative to distance from the reference storm center maximum of the observed 24-h MPE, with points identified by color, reflecting storm TS value ranges. Point size indicates the peak flow magnitude. The high degree of peak flow variability within TS categories is illustrated in Figure 3.19, which is greatest for the 0.30-0.49 TS interval. Also quite evident is that smaller distances of the transposed storm from the reference storm center does not guarantee either higher TS values or peak flows. In fact, the storm with the highest TS and closest to the reference storm center produced a peak flow approximately the same as other transposed storms with significantly lower TS values and at distances much further from the reference storm center. The large degree in the variability of hydrograph response and peak flow relative to the magnitude of the transposed storm TS values, underscores the complex nature of hydrologic prediction using deterministic QPF. Rezacova et al. [153] and Mittermaier and Roberts [124] address the difficulty of QPF verification due to the complex structure of observed precipitation fields due to embedded convection.

WPC QPF for the June 23, 2016 event

For the benefit of completeness, we report results from the RDHM simulation using June 23, 2016 0600 UTC, 24-h WPC QPF. The simulated peak flow was 143 m^3s^{-1} with $TS = 0.30$. Importantly, no 24-h WPC QPF $\geq$50.8 mm (2.0 in), for the June 23, 2016 0700 UTC to June 24, 2016 0600 UTC period, fell with within the Greenbrier River basin. The USGS observed peak flow value, 2285.2 m^3s^{-1}, at Alderson, WV occurred June 24, 2016 0930 UTC. The OHRFC forecast from June 23, 2016 1423 UTC was 180.7 m^3s^{-1}, corresponding to a peak river stage 5.70 feet (1.74 m), which was forecasted to occur June 24, 2016 0600 UTC [41].

3.5 Summary and conclusions

Results presented demonstrate that (1) NOAA/NWS *NEXt generation RADar* (NEXRAD) derived QPE has improved dramatically from 1997-present for the OHRFC area, which is reflected in significantly improved hydrologic simulations over the 1997-2016 hindcast period and that (2) from the perspective of meeting the needs of hydrologic forecasting, QPF improvements have been marginal. The *monte carlo* hydrologic simulation experiment illustrates the sensitivity of hydrologic forecasts to QPF errors, resulting in large peak flow differences within narrow ranges of TS differences. Results from these experiments show that greater QPF *Threat Score (TS)* values do not necessarily produce improved hydrologic forecasts and that considerable variability in hydrologic response should be expected, independent of antecedent basin conditions.

Improved hydrologic simulations resulting from QPE improvements are important in several ways. First, since hydrologic forecasts in large part depend on accurately translating observed precipitation into watershed response through modeling, without consideration of

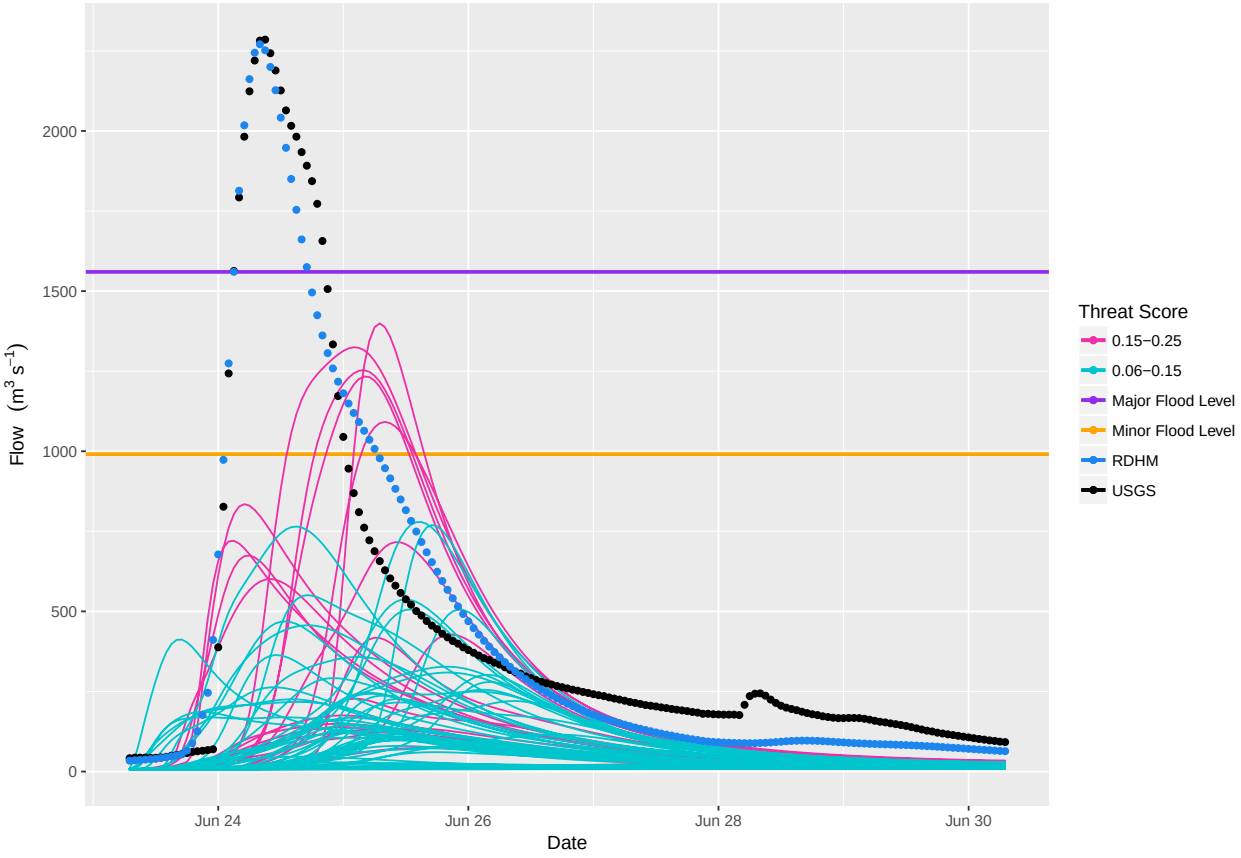

Figure 3.17: Flow hydrographs for the June 23, 2016 07 UTC to June 25, 2017 12 UTC model experiment period for the Greenbrier River at Alderson, WV, showing USGS observed flows (black circles), RDHM simulated hydrographs derived from observed MPE precipitation (blue circles), and the experimental QPF for Threat Score ranges 0.06-0.15 (cyan lines) and 0.15-0.25 (magenta lines). For reference, the Minor and Major Flood levels are shown as horizontal *orange* and *purple* lines, respectively.

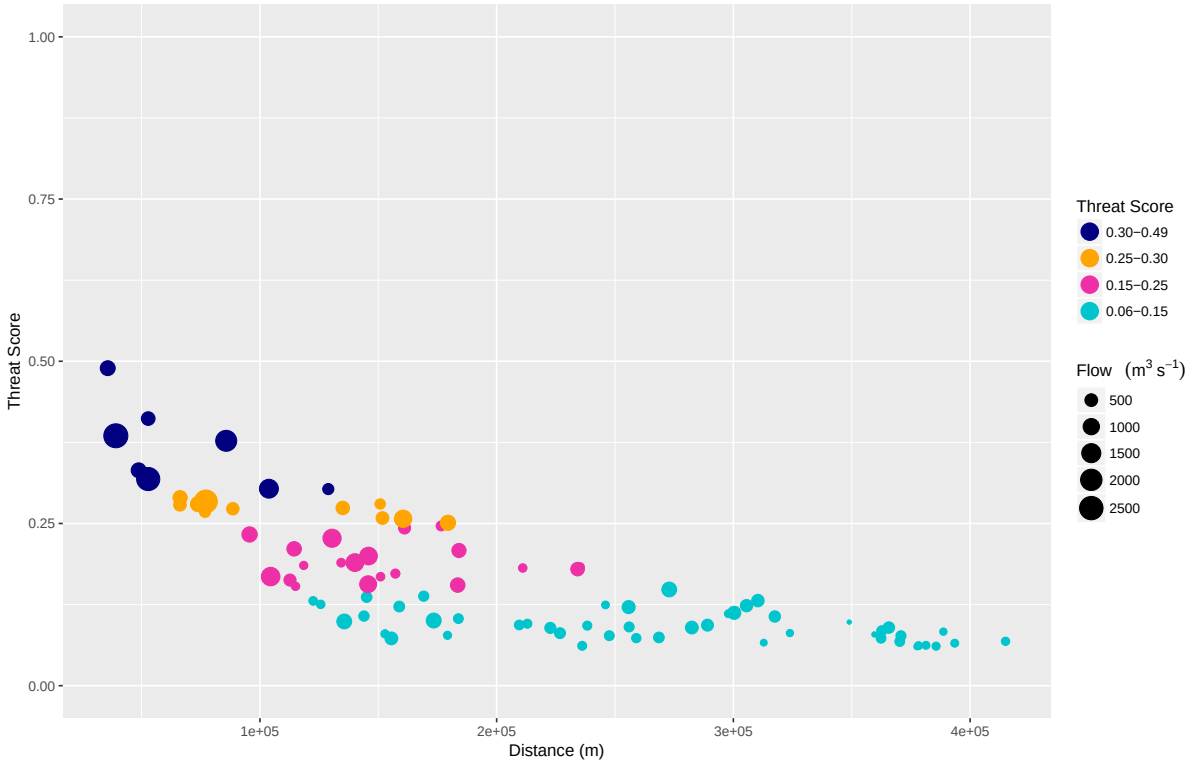

Figure 3.18: Threat Scores (TS≥0.06) of 88 randomly transposed QPF instances with respect to distance from a reference location (see Figure 3.15 (*b*)). Maximum flows derived from RDHM simulations are shown by point size and threat score range, by color, for the Greenbrier River at Alderson, WV, within the June 23, 2016 07 UTC to June 25, 2017 12 UTC model experiment period.

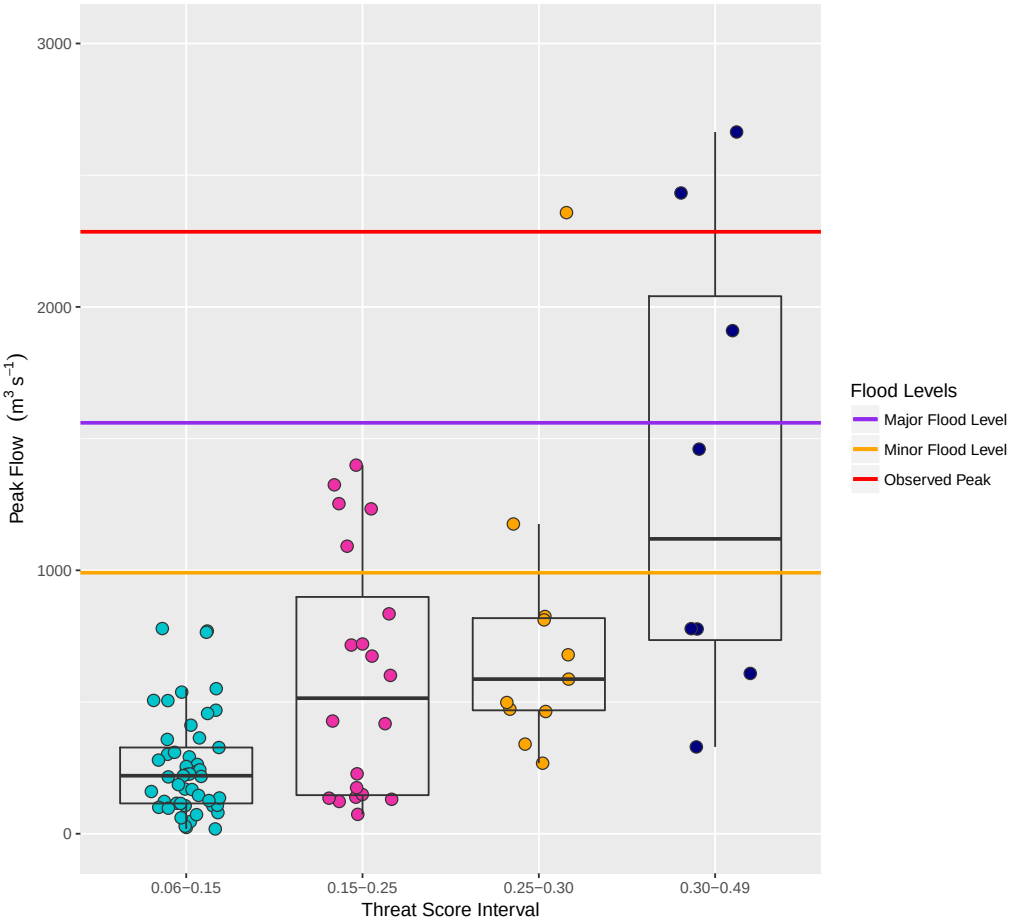

Figure 3.19: Comparison of RDHM simulated peak flows for QPF for Threat Score ranges 0.06-0.15, 0.15-0.25, 0.25-0.30, and 0.30-0.49. For reference, the Minor and Major Flood levels are shown as horizontal *orange* and *purple* lines, respectively for the Greenbrier River at Alderson, WV, within the June 23, 2016 07 UTC to June 25, 2017 12 UTC model experiment period. The USGS observed peak flow is indicated as a *red* line.

QPF, forecasts are improved with improved QPE. Second, improved QPE improves maintenance of model states which are intended to reflect current basin conditions. Consequently, hydrologic forecasts are improved because initial basin conditions are improved and the need for *ad hoc* forecaster adjustments is reduced, which could be made for inappropriate reasons. Namely, an observed difference between simulated and observed flows/stage could simply reflect random model error and not indicate divergence between model states and actual basin conditions. Finally, for well-calibrated models, improved QPE provides greater confidence in the initial model states needed for probabilistic hydrologic forecast systems [40] that run automatically, without direct forecaster intervention.

There are some indications of improved QPF accuracy, shown by WPC with increased *threat score* values from 1970-present. These QPF accuracy gains are reflected by improvements in hydrologic response, which were demonstrated with the RDHM *monte carlo* experiment. However WPC QPF *bias* statistics do not show improvement. In addition, for a variety of QPF sources, NPVU verification statistics do not reveal discernible improvement for the 2001-2009 analysis period. The RDHM *monte carlo* experiments demonstrated significant uncertainty in hydrologic response at current *expected* WPC QPF *threat score* levels. Importantly, significant uncertainty with deterministic QPF has been widely shown to be problematic (Damrath et al. [49]; Ebert [64]; Im et al. [96]; Diomede et al. [60]; Cuo et al. [45]). Significant research and development activity (Cloke and Pappenberger [40]; Adams and Ostrowski [1]; Demargne et al. [58]) in the hydrologic community has been directed at the implementation of probabilistic hydrologic forecast systems with a principal aim to quantify the uncertainties inherent in hydrological forecasting, including uncertainties associated with model forcings, in particular, QPF uncertainties.

The central problem with deterministic QPF is that, *in principle*, the placement, timing, and magnitude of QPF should all be *reasonable* estimates for each of the basins shown, for

example, in Figure 3.2. The RDHM *monte carlo* hydrologic experiment demonstrated how far current deterministic QPF is from consistently providing the needed skill in hydrologic forecasting. It is only for much larger basins where QPF errors are sufficiently masked by spatial and temporal averaging that deterministic QPF has sufficient skill to have value in hydrologic forecasting. The problem is significantly more difficult for flash floods, with affected areas that are considerably smaller than the OHRFC subbasins shown in Figure 3.2.

The methods applied should be applicable to other locations due to the general nature of the techniques used. Data availability is, of course, an issue in many areas where, for instance, radar estimates of precipitation do not exist. In areas where snowmelt flooding dominates results could be different from what is shown where rainfall forced flooding is the principal hydrometeorological flooding process.

Chapter 4

The Effect of QPF on Real-time Deterministic Hydrologic Forecast Uncertainty

4.1 Abstract

The use of *Quantitative Precipitation Forecast (QPF)* in hydrologic forecasting is commonplace, but QPF is subject to considerable error. When QPF is included as a model forcing in the hydrological forecast process, significant error is passed to subsequent hydrologic predictions. Two questions arise: (1) are the resulting observed hydrologic forecast errors sufficiently large to suggest the use of *zero* QPF in the forecast process; if the use of QPF is indicated, (2) how many periods (hours) of QPF (1-, 6-, 12-,..., 72-h...) should be used? Also, do forecast conditions exist under which the use of QPF should be different? This study presents results from two real-time hydrologic forecast experiments, focused on the NOAA/NWS *Ohio River Forecast Center (OHRFC)*. The experiments rely on forecasts

from subbasins at 38 forecast point locations, ranging in drainage area, geographic location within the Ohio River Valley, and watershed response time. Results from an experiment, spanning all flow ranges, for the August 10, 2007 - August 31, 2009 period, show that *non-zero* QPF produces smaller hydrologic forecast error than *zero* QPF. A second experiment, January 23, 2009 through September 15, 2010, suggests that QPF should be limited to 6- to 12-h duration for flood forecasts. Beyond 12-h, hydrologic forecast error increases substantially across all forecast ranges, but errors are much larger for flood forecasts. Increased durations of QPF produce smaller forecast error than shorter QPF durations only for non-flood forecasts. Experimental results are shown to be consistent with NWS, April 2001 to October 2016, forecast verification statistics for the OHRFC.

4.2 Introduction

Single-valued, deterministic *Quantitative Precipitation Forecast (QPF)* is a commonly used model forcing in hydrologic forecasting (Georgakakos and Hudlow [82]; Sokol [167]; Adams [5]; Li et al. [114]). All 13 U.S. *National Oceanic and Atmospheric Administration (NOAA)*, *National Weather Service (NWS) River Forecast Centers (RFCs)*, shown in Figure 4.1, utilize QPF operationally for hydrologic forecasting, ranging in duration from 1- to 10-days. Research has demonstrated that the use of deterministic QPF introduces considerable error into hydrologic forecasting (Cuo et al. [45]; Diomede et al. [61]). Sources of hydrologic forecast uncertainty, including QPF, are recognized by the National Research Council [133]. Adams and Dymond [6] report on the magnitude of QPF uncertainty over the *Conterminous United States (CONUS)* and illustrate, from analyses of a hydrologic modeling experiment for a 3533 km^2 watershed in the Ohio River Valley, that large prediction uncertainties are obtained from the use of deterministic QPF. Despite the advances in our understanding of

QPF errors, unanswered questions remain with the use of QPF.

4.2.1 Background

NWS RFCs are responsible for providing routine river and flood forecasts of stage/flow values to the general public and others on a daily basis, often including evening updates. Additional forecasts are issued during operational periods, as needed, during flooding episodes, subject to changing meteorological conditions, including 24-h per day operational coverage. RFCs currently use the NWS *Community Hydrologic Prediction System (CHPS)* based on the *Flood Early Warning System (FEWS)* [55] as the basis of their modeling system [4]. All RFC CHPS modeling is done within the Linux based NOAA/NWS *Advanced Weather Interactive Processing System (AWIPS)* [140, 141]. The OHRFC uses the Sacramento Soil Moisture Accounting (SAC-SMA) model (Burnash et al. [34]; Burnash [32]), SNOW-17 snow accumulation and ablation model [12], several lumped-parameter hydrologic routing models, and three reservoir simulation models within the CHPS operational environment. All OHRFC CHPS models were migrated from the legacy NWS River Forecast System (NWSRFS) (U.S. Department of Commerce [177]) in 2011.

In addition to QPF, principal hydrologic model forcings are observed precipitation and observed and forecasted temperature. Precipitation observations are obtained from a multisensor estimation process, involving rain gauges, NWS *NEXt-generation RADar (NEXRAD)* doppler radar, and, in some instances, remotely-sensed satellite estimates of precipitation [103]. Forecasted precipitation is derived from numerical weather prediction (NWP) models, usually with meteorological forecaster adjustments made at both the NWS Weather Prediction Center (WPC), formerly the *Hydrometeorological Prediction Center (HPC)*, referred to henceforth as *WPC*, and/or at local RFCs [139]. The hydrologic forecast process in other

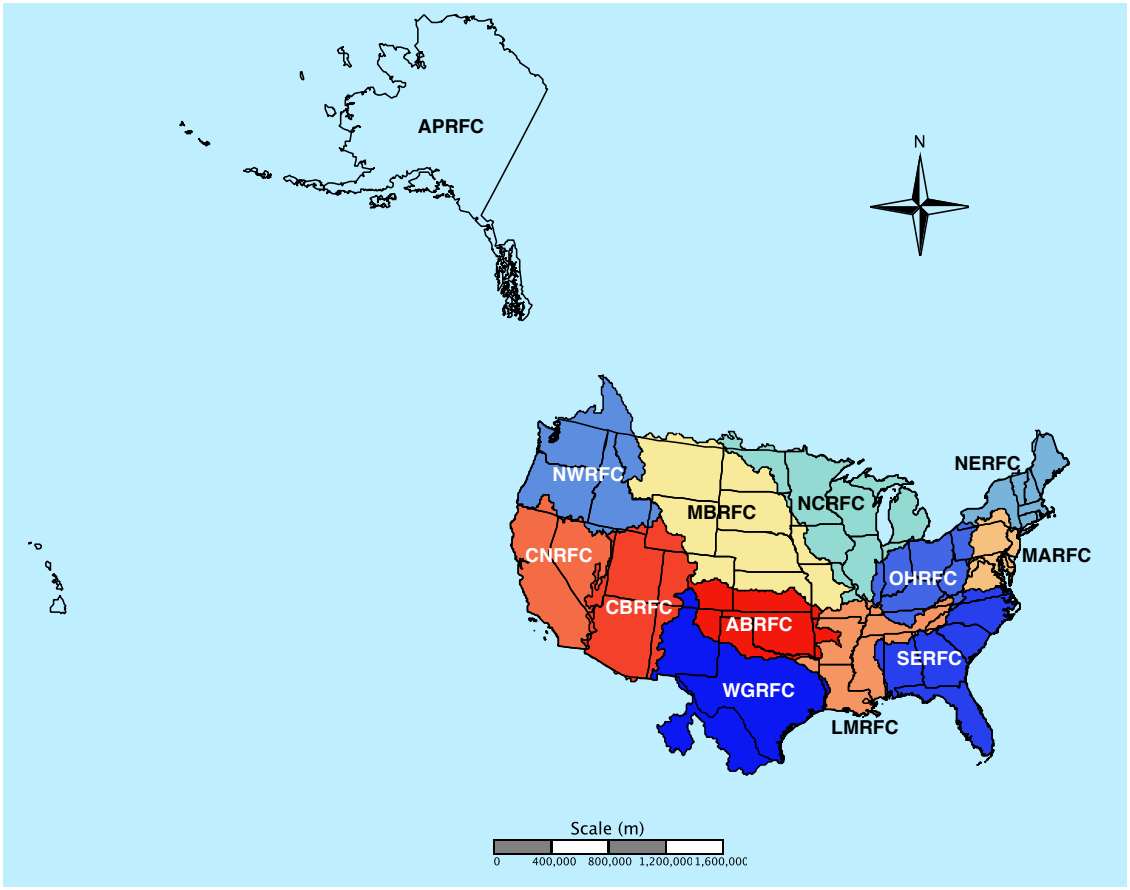

Figure 4.1: The NWS 13 River Forecast Centers (RFCs) – Alaska/Pacific RFC (APRFC), Arkansas-Red RFC (ABRFC), Colorado Basin RFC (CBRFC), California-Nevada RFC(CNRFC), Lower-Mississippi RFC (LMRFC)Middle Atlantic RFC (MARFC), Missouri Basin RFC (MBRFC), North Central RFC (NCRFC), Northwest RFC (NWRFC), Ohio RFC (OHRFC), Southeast RFC (SERFC), and West Gulf RFC (WGRFC). Please note that several RFC boundaries extend beyond the U.S. national boundary into Canada and Mexico.

countries is similar to that used in the U.S. [5].

4.2.2 Research goals

The aim of this research is to frame the potential limits to the use of deterministic QPF for hydrologic forecasting. Section 4.3 of this paper describes two real-time hydrologic forecasting experiments used in this study. We focus model simulations on watersheds in the NOAA/NWS *Ohio River Forecast Center (OHRFC)* area of responsibility, shown in Figure 4.1. The first experiment addresses the issue of whether or not *non-zero* deterministic QPF should be used in hydrological forecasting. The second experiment identifies the expected range of hydrologic forecast error using QPF for OHRFC forecast point locations and a possible limit to the duration of QPF that should be used for flood forecasting before forecast error grows too large for acceptable use. Results of the experiments are presented in section 4.4. The overall context of the experimental results in relation to other RFCs and the conditions under which the use of QPF should possibly be limited are identified in section 4.5. Section 4.6 summarizes the experimental results and presents conclusions.

4.3 Research Approach

Two sets of real-time hydrologic forecast experiments are presented. The experiments, which were made using the legacy NWSRFS, rely on a geographically broad distribution of forecast point locations, with varying basin sizes and hydrologic response times for the OHRFC. Calibrations of SAC-SMA, SNOW-17, channel routing, and reservoir simulation models for operational use for all OHRFC subbasins were completed prior to the time the experiments started, following guidelines presented by Anderson [13] and Smith et al. [166]. All model

experiments were conducted using 1200 UTC model initializations following routine OHRFC operations. Operational and experimental simulations utilize a 6-h time step for model forcings, internally, and output. The two experiments are focused on answering the questions, respectively:

1. should *non-zero* QPF be used in hydrologic forecasting;
2. if the use of *non-zero* QPF is suggested, what duration of QPF seems warranted?

Forecasts are evaluated on the basis of comparisons between U.S. Geological Survey (USGS) observations and simulated river stage values. Results from the experiment are reported in terms of verification statistics using methods proposed by Welles et al. [181] and Demargne et al. [56].

4.3.1 Statistical verification

Verification of hydrologic forecasts are made using:

$$ME = \frac{1}{n}\sum_{k=1}^{n}(y_k - o_k) \tag{4.1}$$

$$MAE = \frac{1}{n}\sum_{k=1}^{n}(|y_k - o_k|) \tag{4.2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(y_k - o_k)^2} \tag{4.3}$$

where we have the *Mean Error (ME)*, *Mean Absolute Error (MAE)*, *Root Mean Square Error (RMSE)*, with quantities y_k and o_k the predicted and observed kth stage values, respectively,

for n total paired values. Units of measure for *stage* are feet, unless reported otherwise. Values for *ME*, *MAE*, and $RMSE = 0$ implies perfect agreement, i.e., no error.

4.3.2 Experiment 1

Watersheds are identified to obtain a range of drainage basin areas and locations to investigate all possible combinations of the use of OHRFC and WPC QPF and interactive forecaster adjustments, termed *runtime modifications*, known as *MODs*. Consequently, 8 independent forecast simulations were made daily, at 12 UTC, for each of the 38 basins, shown in Figure 4.2. The operational and experimental forecasts were made for the period August 10, 2007 - August 31, 2009, using 24-h duration (4, 6-h periods $(24\ \text{h})^{-1}$) QPF. The experimental period spanned 753 days at 38 locations, with 28 forecast periods each (4 6-h periods per day for 7-days), resulting in 801,192 forecast verification pairs for analysis. The experiment was structured to assess:

- if *non-zero* QPF produces smaller error in hydrologic forecasts than *zero* QPF;
- whether or not the use of MODs produces smaller error in hydrologic forecasts than without MODs;
- if the use of local OHRFC QPF or QPF from the *WPC* produces smaller error in hydrologic forecasts in the OHRFC region?

Two experimental forecast scenarios for this experiment are: *with MODs, No QPF* and *with MODs, with WPC*. All experimental forecast results were analyzed using the *R Language and Environment for Statistical Computing* [145] and contributed *verification* package [135] from experimental data stored in the OHRFC *PostgreSQL* verification database.

Table 4.1: Experiment 1 *fast response* basins, listing NWS station identifier (ID), USGS identifier, Station name, Response time category, and basin area.

ID	USGS ID	Name	Area (km^2)	Response
ALDW2	03183500	Alderson WV	3533	Fast
BEAP1	03107500	Beaver Falls PA	8044	Fast
CYCK2	03283500	Clay City KY	938	Fast
DLYW2	03050000	Daily WV	479	Fast
FRAT1	03432350	Franklin TN	497	Fast
GRTW2	03153500	Grantsville WV	2365	Fast
INDI3	03353000	Indianapolis IN	4235	Fast
KILO1	03139000	Kilbuck OH	1202	Fast
MILO1	04199000	Milan OH	961	Fast
MLGO1	03245500	Milford OH	3116	Fast
OLNN6	03010820	Olean NY	3087	Fast
OLPO1	04206000	Old Portage OH	1046	Fast
PSNW2	03069500	Parsons WV	1870	Fast
SWDP1	03041500	Seward PA	1852	Fast

4.3.3 Experiment 2

Basins are categorized as FAST, MEDIUM, and SLOW responding at 38 forecast point locations (Figure 4.2). These include 14 *fast*, 17 *medium*, and 7 *slow* responding forecast point locations. The terms *slow*, *medium*, and *fast* refer to typical time-to-peak response times, from the center-of-mass of the observed precipitation to the hydrograph peak. Response times less than 24 h are classified as *FAST*, response times between 24 h to 60 h are considered *MEDIUM*, and response times greater than 60 h are considered *SLOW*, see [142]. QPF from the NOAA/NWS *Weather Prediction Center (WPC)*, for 6-, 12-, 24-, 36-, 48-, and 72-h durations (at 6-h intervals), was used to generate real-time *experimental* hydrologic forecasts on a daily basis, in parallel with operational forecasts at the OHRFC, for the period January 23, 2009 through September 15, 2010. The experimental period spanned 601 days at 38 locations, with 28 forecast periods each (4 6-h periods per day for 7-days), resulting

Table 4.2: Experiment 1 *medium and slow response* basins, listing NWS station identifier (ID), USGS identifier, Station name, Response time category, and basin area.

ID	USGS ID	Name	Area (km^2)	Response
ATHO1	03159500	Athens OH	2442	Medium
BEDI3	03371500	Bedford IN	10000	Medium
CDIO1	03142000	Cambridge OH	1052	Medium
COLO1	03227500	Columbus OH	4219	Medium
FLRK2	03215000	Fullers Station KY	10093	Medium
FRKP1	03025500	Franklin PA	15493	Medium
FTWI3	04182900	Fort Wayne IN	4988	Medium
KANW2	03193000	Kanawha Falls WV	21681	Medium
LAFI3	03335500	West Lafayette IN	18121	Medium
PARP1	03031500	Parker PA	19868	Medium
PKTO1	03237020	Piketon OH	15115	Medium
PTTP1	03085152	Pittsburgh PA	49471	Medium
SERI3	03365500	Seymour IN	6063	Medium
SHLI3	03373500	Shoals IN	12761	Medium
SPNI3	03357000	Spencer IN	7739	Medium
STRO1	04185000	Stryker OH	1062	Medium
WLBK2	03404000	Williamsburg KY	4162	Medium
DEFO1	04192500	Defiance OH	14361	Slow
EVVI3	03322000	Evansville IN	277600	Slow
HUFI3	03341500	Terre Haute IN	31766	Slow
NHRI3	03378500	New Harmony IN	75716	Slow
NWBI3	03360500	Newberry IN	12142	Slow
PTRI3	03373980	Petersburg IN	28808	Slow
WTVO1	04193500	Waterville OH	16395	Slow

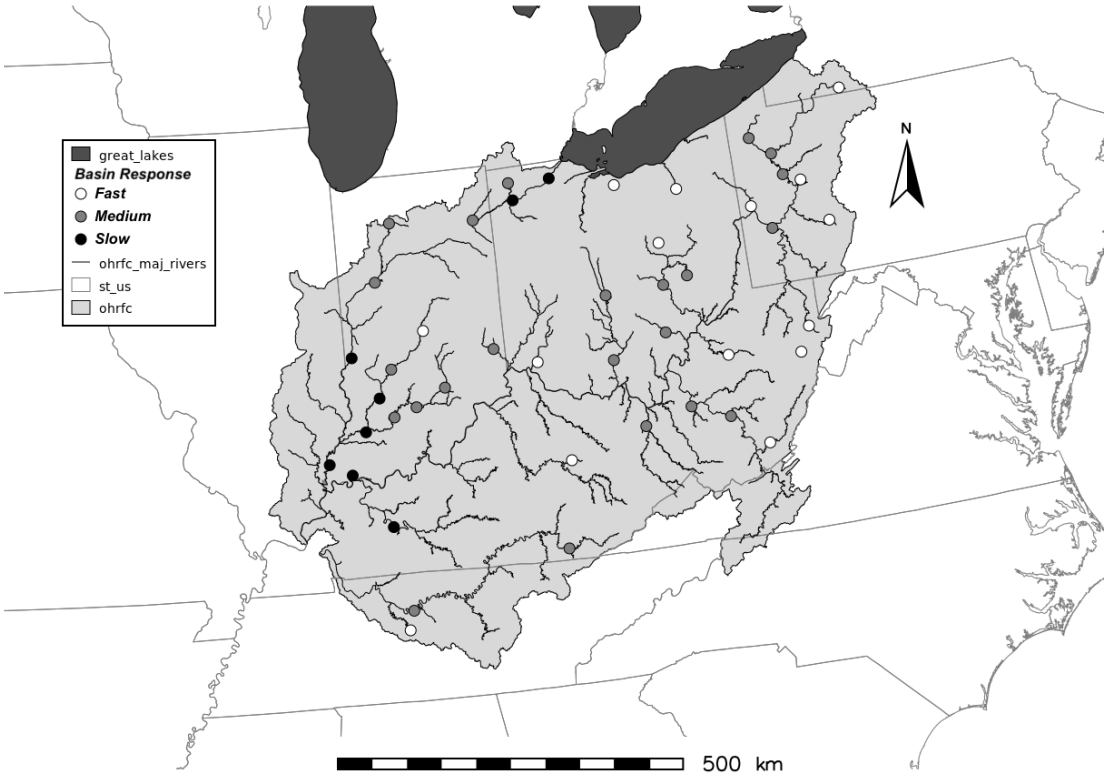

Figure 4.2: Map showing the location of 38 *Experiment 1* and *Experiment 2* forecast point locations used in the OHRFC forecast area, listed in Tables 4.1 and 4.2, identifying *fast*, *medium*, and *slow* responding basins.

in 639,464 forecast verification pairs for analysis. No differences exist in models used or procedures between the *experimental* and *operational* hydrologic forecasts except that, for the *experimental* forecasts:

- the source of the QPF is the WPC, rather than the OHRFC *Hydrometeorological Analysis and Support (HAS)* unit QPF;
- the length of the QPF varies, at 6-, 12-, 24-, 36-, 48-, and 72-h durations (at 6-h intervals), rather than a single fixed 24-h duration (4, 6-h periods $(24\ \text{h})^{-1}$).

4.4 Experimental results

4.4.1 Experiment 1

Statistical forecast verification results from the simulated forecasts for Experiment 1 are presented graphically in Figures 4.2 and 4.4 for the *with* and *without* WPC QPF forecast scenarios, by lead-times ranging from 6- to 168-h, at 6-h intervals. The verification measures shown are *mean error (ME)*, *mean absolute error (MAE)*, and *root mean square error (RMSE)*, given in Equations 5.1 to 5.3. Some immediate findings are that:

1. Referring to Figure 4.3 (b) and (c), for MAE and RMSE, respectively, results indicate that in the first 1-4 forecast periods, little difference is found between forecasts *with* and *without* WPC QPF when forecasts are aggregated across *all* locations;

2. Figure 4.3 (a), for ME values, indicates that differences between forecasts *with* and *without* WPC QPF are evident beginning in the second forecast lead-time period (12-h);

3. As Figure 4.4 illustrates with stratification between *fast*, *medium*, and *slow* responding basins, RMSE values differ little between lead-times 6- through 24-h (periods 1-4) for medium response basins and with all basins lumped together, and through period 8 for slow responding basins;

4. For *fast* responding basins, RMSE values are lower for the *without WPC QPF* forecasts for lead-times 6- through 18-h (periods 1-3) than the *with WPC QPF* forecasts.

The benefit gained from the use of non-zero QPF versus zero-QPF in OHRFC hydrologic forecasts is shown by lower magnitudes of ME, MAE, and RMSE values with *non-zero*

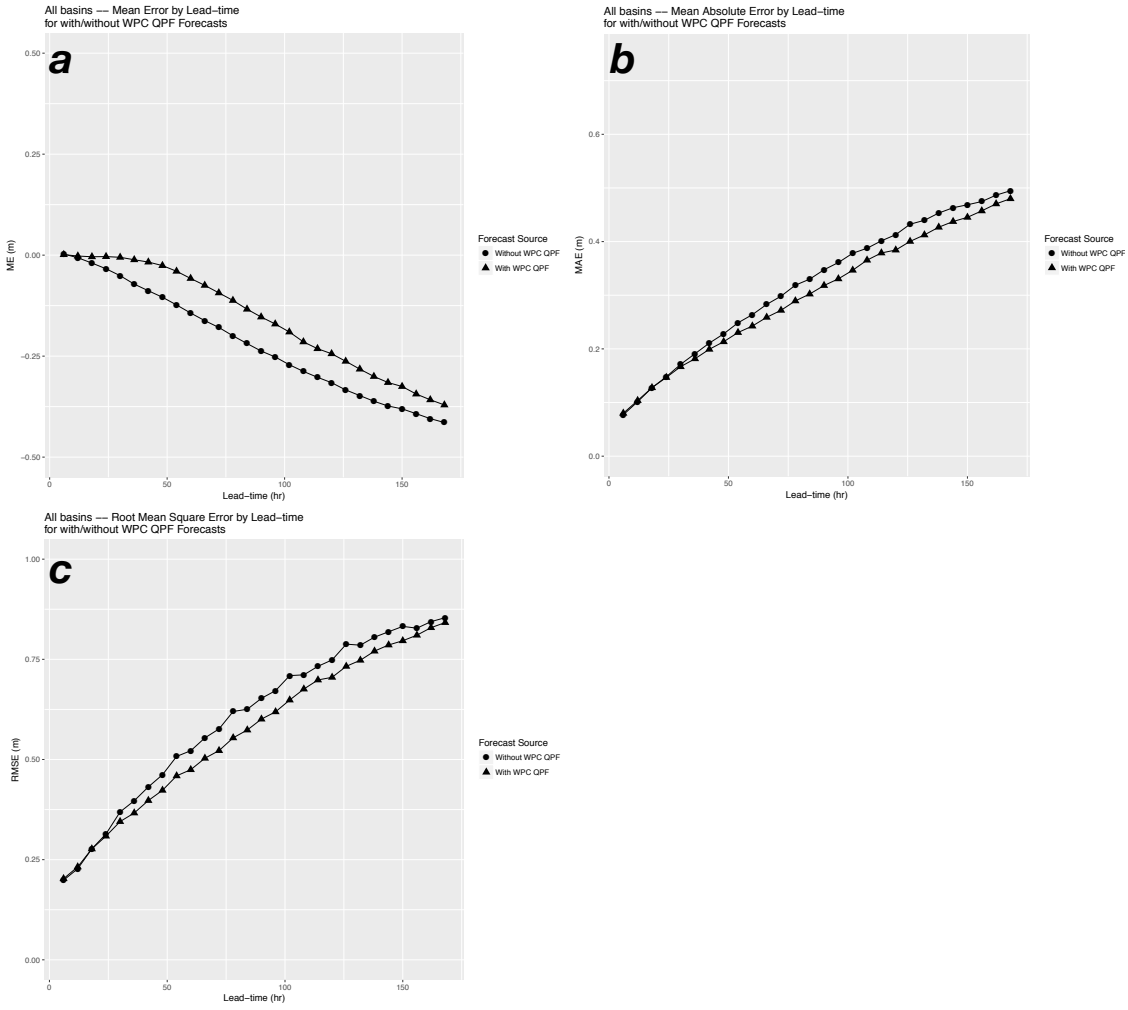

Figure 4.3: Comparison of OHRFC hydrologic forecasts both *with* and *without* WPC QPF, showing ME (a), MAE (b), and RMSE (c) for *all* basins, for all response times, for the OHRFC operational forecast area. Shown for the period August 10, 2007 - August 31, 2009.

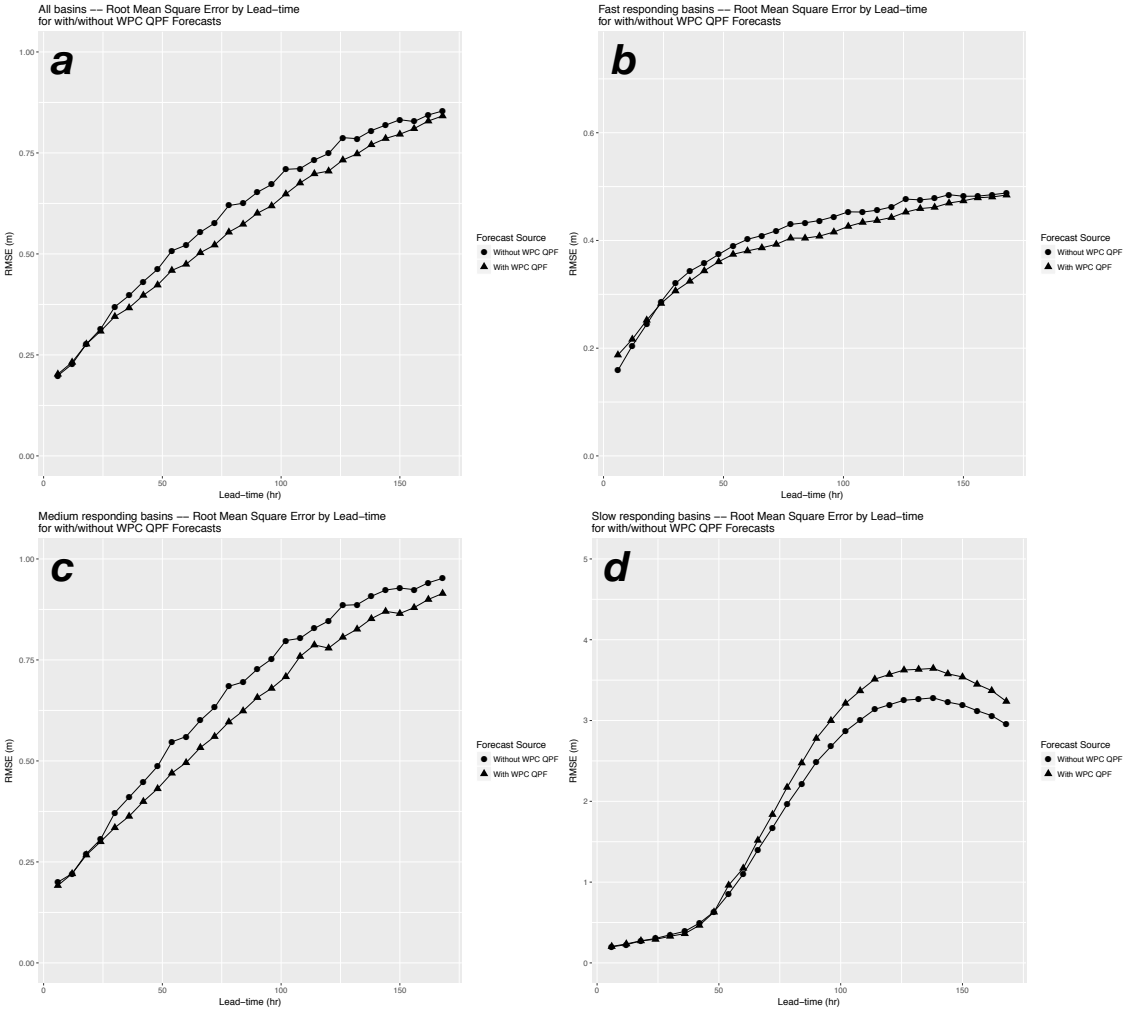

Figure 4.4: RMSE of OHRFC hydrologic forecasts for both *with* and *without* WPC QPF, for All basins (a) and Fast (b), Medium (c) and Slow (d) response basins for the OHRFC operational forecast area. Shown for the period August 10, 2007 - August 31, 2009.

QPF compared to *zero* QPF. That is, Experiment 1 shows that non-zero QPF in hydrologic forecasting produces lower error with longer lead-times than for zero-QPF forecasts. It should be noted that a full 24-h duration (4, 6-h periods $(24\text{ h})^{-1}$) of WPC QPF was used for the experimental QPF scenarios, which was, at the time, the duration of QPF used operationally at the OHRFC. Experiment 2 investigates the influence of the duration of QPF on hydrologic forecast error.

4.4.2 Experiment 2

Experiment 2 results are summarized in Figures 4.5 - 4.8. Table 4.3 provides detailed results of ME, MAE, and RMSE verification measures for *fast* response basins. The following observations can be made:

1. With all basin locations grouped together, verification results show that the magnitude of ME, MAE, and RMSE values increase with longer durations of QPF, which is expected and generally understood;

2. When basins are stratified according to *above* and *below* flood level forecasts categories, *non-flood* (*below* flood) level forecasts exhibit reduced forecast error with longer durations of QPF compared to *above* flood forecasts. In other words, flood forecasts show larger errors than non-flood forecasts, which confirms expectations;

3. Little or no change in hydrologic forecast error is observed with increasing the duration of QPF from 6- to 12-h, irrespective of forecast category or verification measure;

4. For *fast* responding basins, increases in forecast error jump significantly from 6- and 12-h QPF durations to longer QPF durations. Specifically, we see (using Table 4.3) that lengthening the duration of QPF from 6-h to 24-h, increases ME by 114.9% (greater

Table 4.3: *Mean Error (ME)*, *Mean Absolute Error (MAE)* (in (–)), an *Root Mean Square Error (RMSE)* (in [–]) for *Experiment 2* forecasts, averaged across all leadtimes, compared to USGS observed stage values, for 38 NOAA/NWS OHRFC forecast point locations, by QPF Duration (hours), for *Fast*, *Medium*, *Slow*, and combined (*All*) basin response times and *Above* and *Below* flood stage forecast categories. Shown for the period January 23, 2009 - September 15, 2010. Units are expressed in *meters*.

	Fast		**Medium**		**Slow**		**All**	
Duration	Above	Below	Above	Below	Above	Below	Above	Below
6	0.1147	-0.1398	0.0668	-0.2540	-0.0874	-0.31810	0.0382	-0.2222
	(0.3790)	(0.1950)	(0.8157)	(0.3517)	(0.4195)	(0.4137)	(0.5621)	(0.3035)
	[0.2994]	[0.3443]	[0.6828]	[0.5760]	[0.5635]	[0.6477]	[0.5577]	[0.5081]
12	0.1013	-0.1342	0.0736	-0.2466	-0.0814	-0.3063	0.0382	-0.2147
	(0.3754)	(0.1939)	(0.8244)	(0.3489)	(0.4192)	(0.4071)	(0.5643)	(0.3007)
	[0.2981]	[0.3423]	[0.7067]	[0.5700]	[0.5704]	[0.6388]	[0.5705]	[0.5030]
24	0.2465	-0.1267	0.1113	-0.2332	-0.0458	-0.2843	0.1190	-0.2020
	(0.5668)	(0.1944)	(0.8244)	(0.3456)	(0.4229)	(0.3988)	(0.6229)	(0.2979)
	[0.5955]	[0.3424]	[0.7173]	[0.5650]	[0.5793]	[0.6247]	[0.6415]	[0.4981]
36	0.3445	-0.1173	0.1660	-0.2203	-0.0030	-0.2641	0.1894	-0.1892
	(0.6406)	(0.1927)	(0.8049)	(0.3422)	(0.4438)	(0.3906)	(0.6570)	(0.2943)
	[0.6520]	[0.3396]	[0.7666]	[0.5578]	[0.6201]	[0.6153]	[0.6917]	[0.4920]
48	0.3863	-0.1082	0.2146	-0.2073	0.0397	-0.2441	0.2340	-0.1764
	(0.6737)	(0.1914)	(0.8495)	(0.3389)	(0.4659)	(0.3831)	(0.6920)	(0.2910)
	[0.6715]	[0.3371]	[0.8343]	[0.5511]	[0.6398]	[0.6061]	[0.7316]	[0.4863]
72	0.4212	-0.0963	0.2585	-0.1907	0.0767	-0.2202	0.2730	-0.1603
	(0.6985)	(0.1902)	(0.8912)	(0.3350)	(0.4839)	(0.3748)	(0.7218)	(0.2874)
	[0.7080]	[0.3345]	[0.9002]	[0.5436]	[0.6539]	[0.5964]	[0.7745]	[0.4801]

than doubling 6-h ME), MAE by 49.6% (nearly 1.5 times 6-h MAE), and RMSE by 98.9% (nearly doubling 6-h RMSE), which is summarized in Table 4.4.

5. Following the cessation of QPF, error statistics begin to converge with leadtimes >150-h, as shown in Figures 4.3, 4.4, 4.5.

Figure 4.6 (a) and (b) show ME and (c) and (d) show MAE for *medium* and *slow* response basins, respectively, for the six QPF durations, by forecast category. These results show a general trend for both *medium* and *slow* response basins with increasing forecast error for *above* flood forecasts with longer leadtimes, but decreasing error for *below* flood forecasts with increased QPF durations, with longer leadtimes. Figure 4.7 shows (a) ME, (b) MAE, and (c) RMSE for QPF durations, grouped by *above* and *below* flood forecast categories for *fast* response basins. Results for ME, MAE, and RMSE verification measures, which

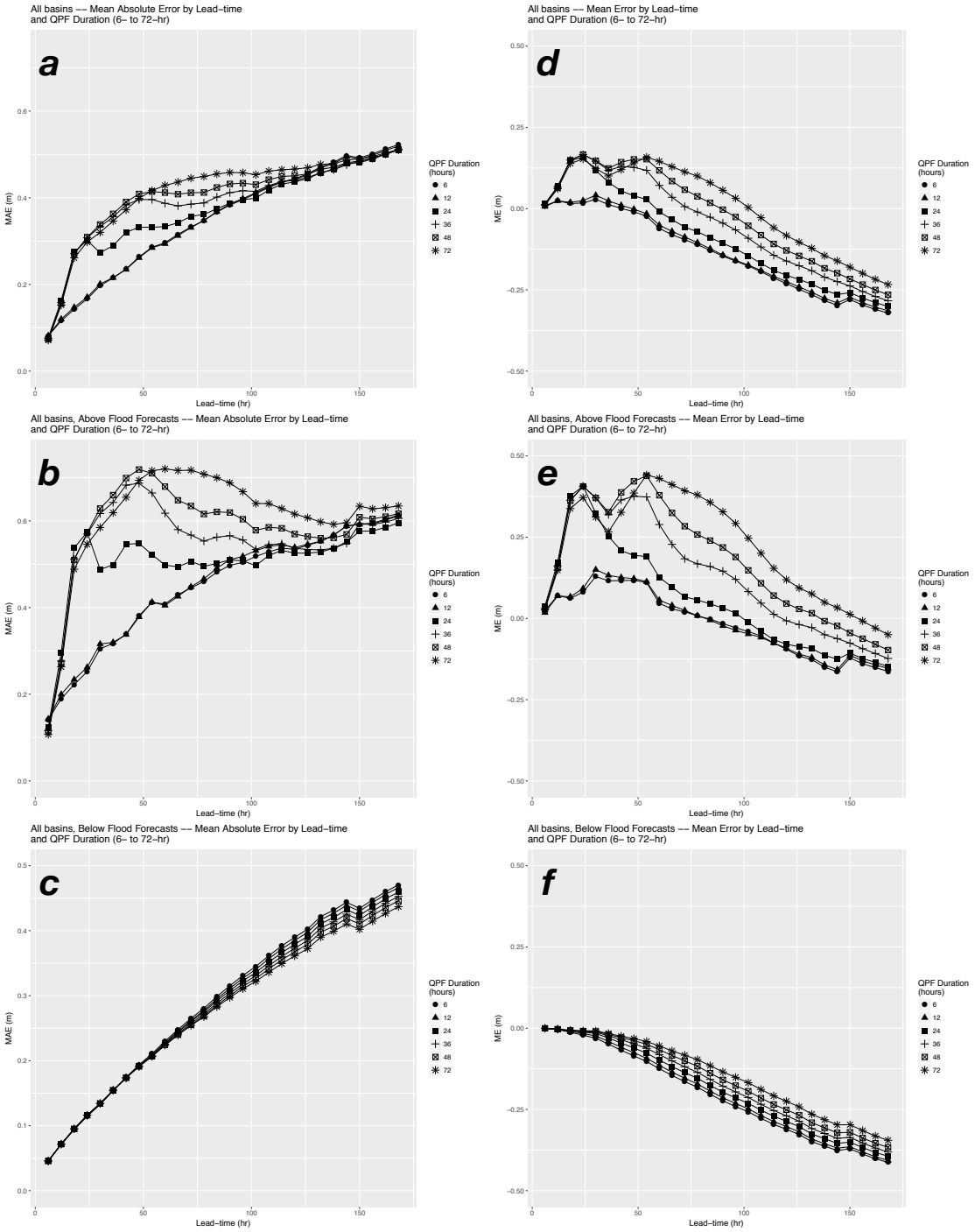

Figure 4.5: MAE for (a) *all* basins (b) *above*, (c) *below* flood stage category forecasts, and ME for (d) *all* basins (e) *above*, (f) *below* flood stage categories, by lead-time for all 38 *Experiment 2* basins, with QPF ranging from 6- to 72-h durations. Shown for the period January 23, 2009 - September 15, 2010.

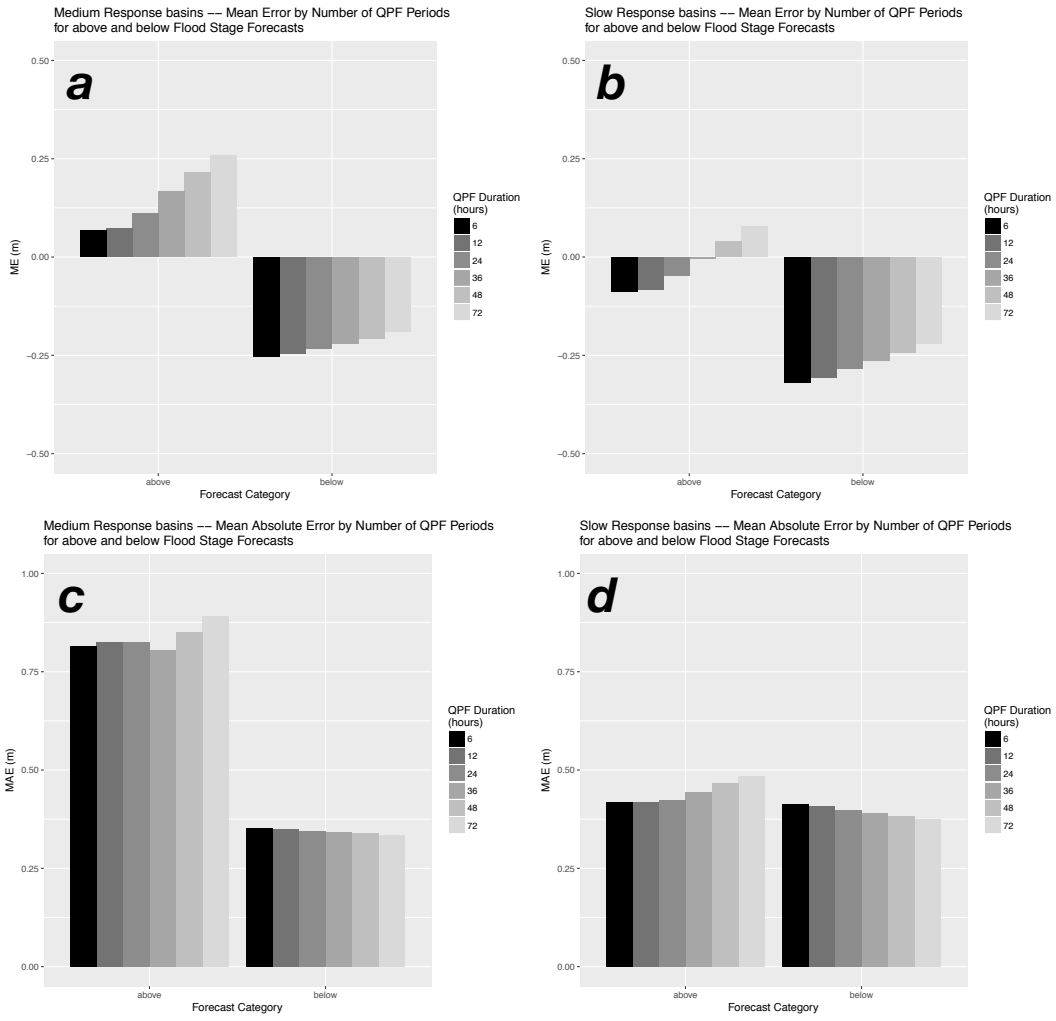

Figure 4.6: ME for (a) *medium* and (b) *slow* responding basins and MAE for (c) *medium* and (d) *slow* responding basins, by QPF durations ranging from 6- to 72-h, for above and below flood stage categories for *Experiment 2*. Results are aggregated across all lead-times. Shown for the period January 23, 2009 - September 15, 2010.

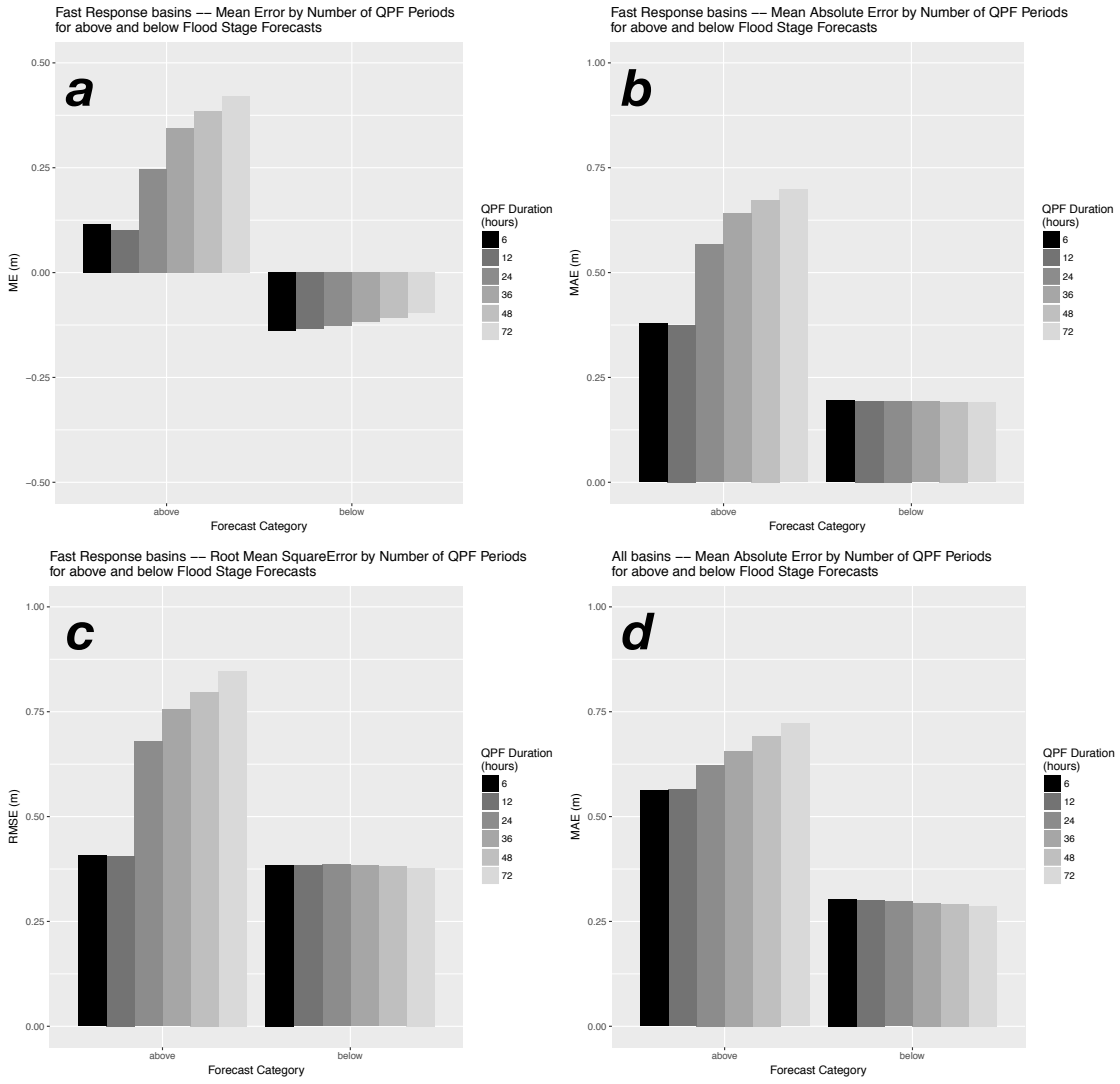

Figure 4.7: Experimental results for *fast* responding *Experiment 2* basins, aggregated across all lead-times, showing (a) ME, (b) MAE , and (c) RMSE for QPF ranging from 6- to 72-h durations. MAE for *all* basins (d) is included for comparison purposes. Shown for the period January 23, 2009 - September 15, 2010.

Table 4.4: Summary of hydrologic forecast error change (%) from 6-h duration WPC QPF to 12-, 24-, 36-, 48-, and 72-h durations for *Fast* responding basins, for *above* flood stage category forecasts, from Table 4.3.

QPF Duration	ME	MAE	RMSE
6	–	–	–
12	-11.7	-0.9	-0.4
24	114.9	49.6	98.9
36	200.3	69.0	117.8
48	236.8	77.8	124.3
72	267.2	84.3	136.5

are summarized in Table 4.3, show significant increases in hydrologic forecast error with longer QPF durations for *above* flood level forecasts. *Below* flood level forecasts reflect slight decreases in hydrologic forecast error with longer QPF durations.

4.5 Discussion

Forecast verification results from Experiments 1 and 2 have demonstrated several key points related to the use of *deterministic* QPF, namely,

1. The use of *non-zero* QPF reduces hydrologic forecast error compared to *zero* QPF;
2. Hydrologic forecast error generally increases with longer durations of QPF. The exception occurs with *below* flood level forecasts, where forecast error diminishes with longer durations of QPF;
3. For *flood* forecasting purposes, the use of QPF beyond either 6- or 12-h is not recommended due to greatly increased hydrologic forecast error.

The latter point is underscored in Figure 4.8, which shows RMSE for all basins and response

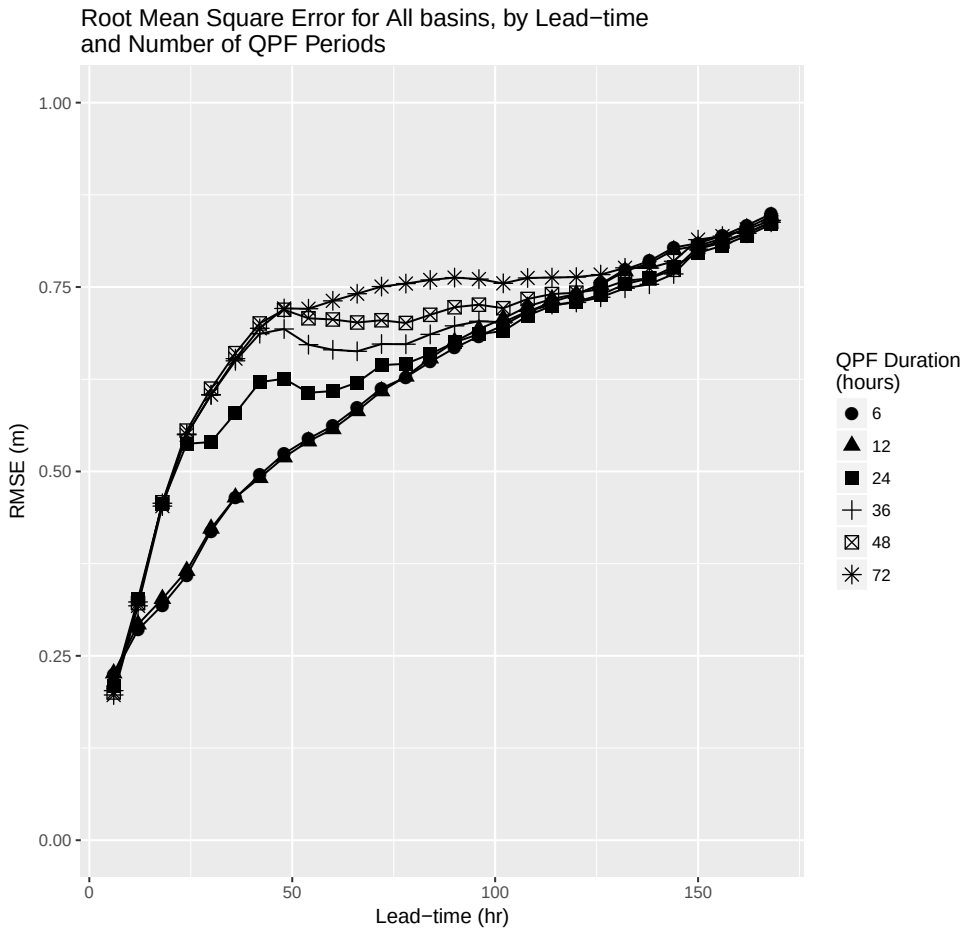

Figure 4.8: RMSE for all basins and response times, by lead-time, for the OHRFC QPF forecast scenarios (6-, 12-, 24-, 36-, 48-, and 72-h). Shown for the period January 23, 2009 - September 15, 2010.

times, with combined *above* and *below* flood category forecasts. Figure 4.8 also supports results presented in Figure 4.7.

We note that Experiment 1 forecast verification statistics are consistent with official NWS RFC forecast verification statistics for the period April 2001 to October 2016, shown in Table 4.5, for the OHRFC, and Table 4.6, for all 13 RFCs. RFC forecast verification statistics can be found at the *Performance Management* website [138] for operational forecasts utilizing the full range of RFC QPF, which does vary between RFCs. A further note is that in 2011,

the OHRFC changed the use of QPF from 24-h duration (4, 6-h periods $(24\ \text{h})^{-1}$) to 48-h duration (8, 6-h periods $(48\ \text{h})^{-1}$) for routine operations.

The findings from Experiment 2 show that the use of longer durations of QPF as a model forcing leads to increased hydrologic forecast error, except for *below flood* forecasts, where errors are reduced with increased durations of QPF. Figure 4.6 shows smaller MAE values at a given lead-time and QPF duration, reflecting smaller error, for *slow* responding basins relative to *medium* and *fast* responding basins, shown in Figure 4.7. In turn, smaller MAE values for a given lead-time and QPF duration are found with *medium* response basins than for *fast* responding basins. This reflects the relative insensitivity of larger watersheds to QPF location errors compared to smaller watersheds. Dramatic increases in flood forecast error for fast responding watersheds was demonstrated for QPF durations >12-h. Because larger watersheds are comprised of many smaller, fast responding basins where flood forecasts are issued, we believe QPF with durations ≤12-h should be recommended for flood forecasting purposes.

However, these results beg the issue of how QPF uncertainty is translated into the expression of uncertainty in hydrologic forecasting. Advancements in probabilistic and ensemble hydrologic forecasting capture the need to quantify hydrologic forecast uncertainty, particularly as it relates to QPF uncertainty (Cloke and Pappenberger [40]; Adams and Ostrowski [1]; Demargne et al. [58]). We believe that the findings from this study, taken together with conclusions drawn from Adams and Dymond [6], which quantifies the magnitude of hydrologic forecast error that can ensue from the use of deterministic QPF, underscores the necessity for the use of ensemble, or, more generally, probabilistic hydrologic forecasting over deterministic forecasting using single-valued deterministic QPF.

Table 4.5: Verification statistics for the NOAA/NWS OHRFC, for *Above* and *Below* flood stage, and *Combined (both above and below)* forecasts, showing *Mean Error (ME)*, *Mean Absolute Error (MAE)*, and *Root Mean Square Error (RMSE)*, expressed in *feet*, for April 2001 to October 2016. *N* is the number of observation-forecast pairs.

	Below Flood Stage				Above Flood Stage				Combined Flood Stage			
Time	**N**	**RMSE**	**MAE**	**ME**	**N**	**RMSE**	**MAE**	**ME**	**N**	**RMSE**	**MAE**	**ME**
6	409970	0.330	0.199	0.013	15328	0.735	0.594	-0.373	425298	0.345	0.213	-0.001
12	417986	0.465	0.294	0.013	16227	1.049	0.813	-0.462	434213	0.487	0.313	-0.005
18	410393	0.587	0.368	0.022	15629	1.275	1.008	-0.572	426022	0.612	0.392	0.000
24	422208	0.685	0.437	0.027	16700	1.398	1.110	-0.629	438908	0.713	0.463	0.002
30	408590	0.758	0.484	0.022	15397	1.623	1.317	-0.787	423987	0.789	0.515	-0.008
36	415890	0.814	0.530	0.013	15654	1.804	1.434	-0.874	431544	0.849	0.562	-0.019
42	408475	0.887	0.575	0.009	14794	1.919	1.546	-0.932	423269	0.923	0.609	-0.024
48	421407	0.953	0.624	-0.009	15823	1.973	1.575	-0.962	437230	0.990	0.658	-0.044
54	407480	1.003	0.655	-0.045	14462	2.194	1.787	-1.164	421942	1.044	0.694	-0.083
60	415397	1.042	0.689	-0.087	14638	2.385	1.912	-1.286	430035	1.088	0.731	-0.128
66	407558	1.105	0.727	-0.128	13862	2.529	2.057	-1.433	421420	1.152	0.771	-0.170
72	420469	1.166	0.773	-0.181	14864	2.645	2.139	-1.527	435333	1.217	0.820	-0.227
All Proj	4965823	0.817	0.530	-0.028	183378	1.775	1.42	-0.903	5149201	0.851	0.562	-0.059
Day 1	1660557	0.518	0.325	0.019	63884	1.120	0.886	-0.511	1724441	0.540	0.346	-0.001
Day 2	1654362	0.854	0.554	0.008	61668	1.830	1.468	-0.889	1716030	0.889	0.586	-0.024
Day 3	1650904	1.080	0.711	-0.111	57826	2.439	1.974	-1.352	1708730	1.126	0.754	-0.153
All Days	4965823	0.817	0.530	-0.028	183378	1.775	1.425	-0.903	5149201	0.851	0.562	-0.059

Table 4.6: Verification statistics for all NOAA/NWS RFCs, for *Above* and *Below* flood stage, and *Combined (both above and below)* forecasts, showing *Mean Error (ME)*, *Mean Absolute Error (MAE)*, and *Root Mean Square Error (RMSE)*, expressed in *feet*, for April 2001 to October 2016. N is the number of observation-forecast pairs.

	Below Flood Stage				Above Flood Stage				Combined Flood Stage			
Time	**N**	**RMSE**	**MAE**	**ME**	**N**	**RMSE**	**MAE**	**ME**	**N**	**RMSE**	**MAE**	**ME**
6	4670053	0.261	0.159	0.016	248967	0.410	0.308	-0.142	4919020	0.268	0.166	0.008
12	4684457	0.367	0.228	0.024	256717	0.672	0.517	-0.220	4941174	0.383	0.243	0.011
18	4694310	0.454	0.285	0.036	248630	0.855	0.662	-0.301	4942940	0.475	0.304	0.019
24	4888601	0.515	0.326	0.042	249815	0.984	0.768	-0.396	5138416	0.538	0.348	0.021
30	4688068	0.578	0.371	0.039	241304	1.127	0.889	-0.489	4929372	0.605	0.396	0.013
36	4687903	0.626	0.406	0.031	239426	1.270	1.012	-0.558	4927329	0.658	0.436	0.002
42	4694345	0.672	0.439	0.026	232476	1.365	1.092	-0.618	4926821	0.704	0.470	-0.005
48	4885648	0.710	0.464	0.019	233029	1.452	1.162	-0.710	5118677	0.744	0.496	-0.014
54	4562969	0.744	0.492	-0.008	222518	1.560	1.254	-0.800	4785487	0.782	0.528	-0.045
60	4288623	0.782	0.521	-0.025	218516	1.691	1.363	-0.869	4507139	0.826	0.562	-0.066
66	4266722	0.813	0.546	-0.045	212815	1.778	1.434	-0.938	4479537	0.859	0.589	-0.087
72	4359841	0.844	0.568	-0.061	213384	1.882	1.522	-1.027	4573225	0.892	0.613	-0.106
All Proj	55371540	0.610	0.397	0.009	2817597	1.228	0.978	-0.572	58189137	0.640	0.425	-0.019
Day 1	18937421	0.401	0.251	0.030	1004129	0.730	0.564	-0.264	19941550	0.417	0.266	0.015
Day 2	18955964	0.647	0.420	0.028	946235	1.302	1.037	-0.592	19902199	0.678	0.450	-0.001
Day 3	17478155	0.795	0.531	-0.034	867233	1.726	1.391	-0.907	18345388	0.839	0.572	-0.076
All Days	55371540	0.610	0.397	0.009	2817597	1.228	0.978	-0.572	58189137	0.640	0.425	-0.019

4.6 Summary and conclusions

Two independent sets of experimental forecasts, initialized with 1200 UTC saved model states from OHRFC daily operational forecasts, were conducted to assess the value of deterministic QPF as a model forcing in hydrologic forecasting. Specifically, in Experiment 1, we investigated if the use of *non-zero* QPF as a model forcing produced hydrologic forecasts exhibiting smaller forecast error than hydrologic forecasts using *zero* QPF. Experiment 2 examined the influence of QPF duration on hydrologic forecast error, if the use of *non-zero* QPF produces hydrologic forecasts with smaller error than *zero* QPF forecasts. Both Experiment 1 and Experiment 2 demonstrate that hydrologic forecast error increases with longer forecast lead-times. Experiment 1 forecast verification statistics suggest that the use of *non-zero* QPF as a hydrologic model forcing produces forecasts with less error than forecasts using *zero* QPF. Experiment 2 forecast verification statistics indicate that, for *above* flood forecasts, longer durations of QPF increases hydrologic forecast error irrespective of basin size and hydrologic response time. For *non-flood* forecasts, longer durations of QPF reduces hydrologic forecast uncertainty. We emphasize the finding from Experiment 2 that shows for fast responding basins, for *above* flood forecasts, the use of QPF beyond 6- or 12-h durations increases hydrologic forecast error dramatically. Consequently, from the perspective of flood forecasting, the use of deterministic QPF should be restricted to 6- to 12-h duration, beyond 6- to 12-h QPF durations hydrologic forecast errors are increased significantly without benefit.

Chapter 5

Use of Central Tendency Measures from an Operational Short Lead-time Hydrologic Ensemble Forecast System

5.1 Abstract

This study presents findings from a real-time forecast experiment that compares legacy deterministic hydrologic stage forecasts to ensemble mean and median stage forecasts from the NOAA/NWS Meteorological Model-based Ensemble Forecast System (MMEFS). The NOAA/NWS *Ohio River Forecast Center (OHRFC)* area of responsibility defines the experimental region. Real-time forecasts from subbasins at 54 forecast point locations, ranging in drainage area, geographic location within the Ohio River Valley, and watershed response time serve as the basis for analyses. In the experiment, operational hydrologic forecasts, with 24-h QPF and forecast temperatures, are compared to MMEFS hydrologic ensemble mean and median forecasts, with model forcings from the NOAA/NWS *National Centers for En-*

vironmental Prediction (NCEP) North American Ensemble Forecast System (NAEFS), over the period, November 30, 2010 through May 24, 2012. Experiments indicate that MMEFS ensemble mean and median forecasts exhibit lower errors beginning at about lead-time 90-h when forecasts at all locations are aggregated. With fast response basins, that peak ≤24-h, ensemble mean and median forecasts exhibit lower errors much earlier, beginning at about lead-time 36-h, which suggests the viability of using MMEFS ensemble forecasts as an alternative to OHRFC legacy forecasts. Analyses show that ensemble *median* forecasts generally exhibit smaller errors than ensemble *mean* forecasts for all stage ranges. Verification results suggest that OHRFC MMEFS NAEFS ensemble forecasts are reasonable, but needed improvements are identified.

5.2 Introduction

The use of hydrologic ensembles to produce probabilistic flood and water resources forecasts, using *Ensemble Prediction Systems (EPSs)*, is rapidly gaining acceptance [40, 58, 133, 141]. However, full adoption of probabilistic forecasts by the public and decision-makers as a replacement to traditional single-valued deterministic hydrologic forecasts is problematic, particularly with how risk-based forecasts are communicated to end-users [59, 123, 143, 147] and because of "institutional conservatism" [149]. National Academies [132] report that with weather related decision making, end-users of weather related forecasts benefit from (1) their understanding of forecasts developed over time, (2) prior experience with severe weather, and (3) other factors, such as family relationships. Related to end-user familiarity with hydrometeorlogical forecasts, National Research Council [133] and Joslyn and Savelli [99] found that end-users understand hydrometeorological forecasts are uncertain, but they make internal adjustments to account for these uncertainties. Morss et al. [128] also found that

end-users of weather forecasts understood forecasts are uncertain and that most preferred the inclusion of uncertainty information with the forecasts. But as Demeritt et al. [59] points out, resistance to the acceptance of EPS forecasts is "not simply cognitive or communicative", there is also the need by decision makers to "shift institutional liability for decisions taken in the face of uncertainty". Murphy [129] and Krzysztofowicz [110] argue for the adoption of probabilistic hydrometeorological forecasting, pointing out that rational decision-making in such a system necessarily shifts decision making from the forecaster to end-users of forecasts. An intuitive understanding of this undoubtedly helps to shape the reluctance by end users to adopt probabilistic hydrometeorological forecasts. In other words, resistance to the adoption of forecasts derived from EPSs, in the form of a probabilistic forecast, by both individuals and many decision makers, is complex, even with prior understanding that single-valued deterministic forecasts are uncertain. There is the added issue pointed to by Stern and Easterling [169] that addresses the need for *climate forecasts* to be relevant to make them useful. This need applies to weather and hydrologic forecasts as well, which points to the broad issue, not addressed in this paper, of how to best convey forecast uncertainty to end-users in ways that are relevant to them. We might ask, however, if there is an interim step with the use of EPSs, that can be taken that addresses two issues related to flood forecasting and the eventual adoption of probabilistic hydrologic forecasts, namely:

1. Improving flood forecast accuracy over current deterministic hydrologic forecasting methods that rely on single-valued *Quantitative Precipitation Forecast (QPF)*;

2. Softening the landscape for end-users for eventual adoption of forecasts derived from EPSs in the form of probabilistic forecasts.

In this paper we explore the use of ensemble mean and median hydrologic forecasts from an EPS as alternatives to deterministic predictions that depend on single-valued QPF. The

study region in this paper is the forecast area of responsibility of the *National Oceanic and Atmospheric Administration (NOAA)*, *National Weather Service (NWS)*, *Ohio River Forecast Center (OHRFC)*, shown in Figure 5.1, which is one of thirteen NOAA/NWS *River Fortecast Centers (RFCs)*. Single-valued, deterministic QPF is a commonly used model forcing in hydrologic forecasting (Georgakakos and Hudlow [82]; Sokol [167]; Adams [5]; Li et al. [114]), and used by all NWS RFCs. Research has demonstrated that the use of deterministic QPF introduces considerable error into hydrologic forecasting (Cuo et al. [45]; Diomede et al. [61]; Adams and Dymond [6]; [7]). We hypothesize that ensemble mean or median forecasts have smaller error than deterministic hydrologic forecasts that rely on single-valued QPF, suggested by Du et al. [63], Mylne et al. [130] with numerical weather prediction (NWP) ensemble modeling systems.

5.2.1 Background

NWS RFCs are responsible for providing routine river stage/flow forecast guidance to NWS *Weather Forecast Offices (WFOs)* following procedures described in Adams [4] and Adams and Dymond [7]. The central responsibility of most RFCs is flood prediction, although for RFCs in western States, water supply forecasting, largely for reservoir inflows is, perhaps, of greater importance. RFCs utilize the NWS *Community Hydrologic Prediction System (CHPS)* [4], based on the *Flood Early Warning System (FEWS)* [55]. CHPS modeling is predominantly interactive, as described by Adams and Smith [2], within the Linux based NOAA/NWS *Advanced Weather Interactive Processing System (AWIPS)* [140, 141]. The OHRFC employs several models within the CHPS operational environment, including the Sacramento Soil Moisture Accounting (SAC-SMA) model Burnash [32], Burnash et al. [34], SNOW-17 snow accumulation and ablation model [12], several lumped-parameter hydrologic routing models, and three reservoir simulation models. All OHRFC CHPS models were mi-

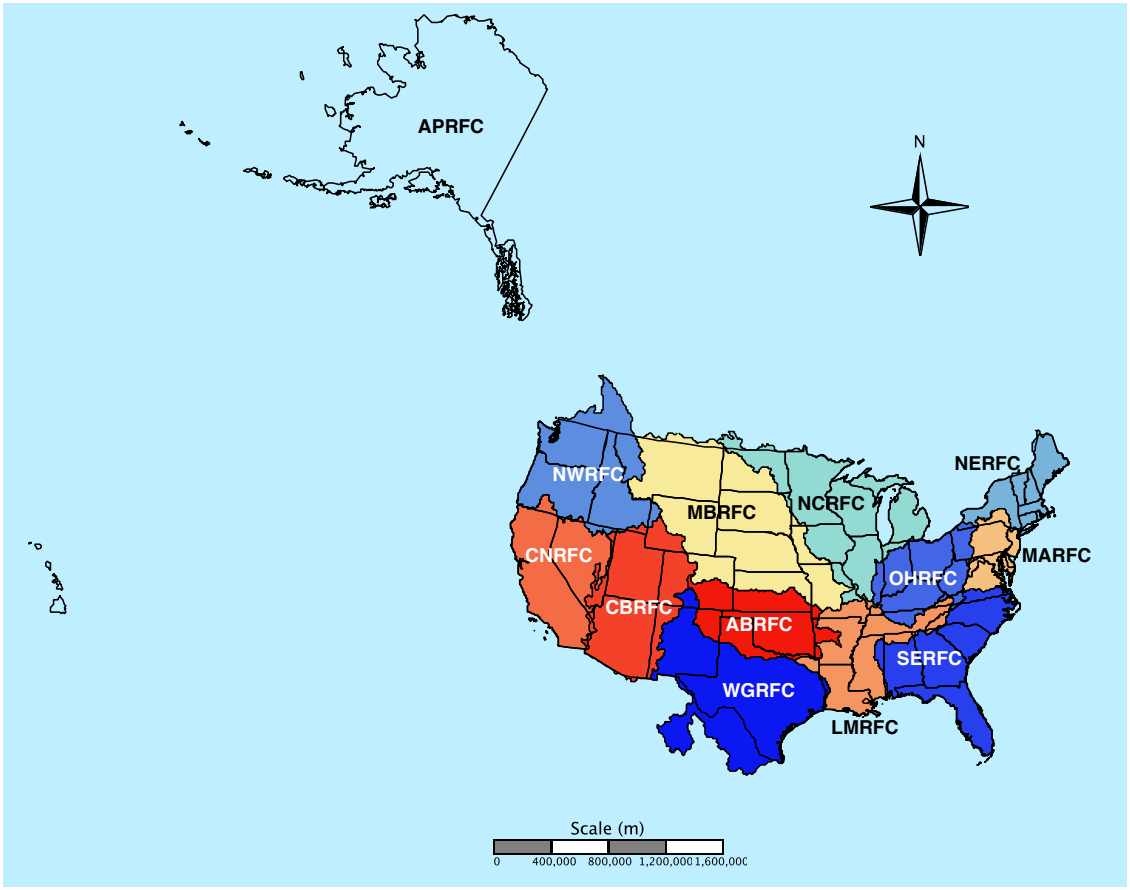

Figure 5.1: The NWS 13 River Forecast Centers (RFCs) – Alaska/Pacific RFC (APRFC), Arkansas-Red RFC (ABRFC), Colorado Basin RFC (CBRFC), California-Nevada RFC(CNRFC), Lower-Mississippi RFC (LMRFC)Middle Atlantic RFC (MARFC), Missouri Basin RFC (MBRFC), North Central RFC (NCRFC), Northwest RFC (NWRFC), Ohio RFC (OHRFC), Southeast RFC (SERFC), and West Gulf RFC (WGRFC). Please note that several RFC boundaries extend beyond the U.S. national boundary into Canada and Mexico.

grated from the legacy NWS River Forecast System (NWSRFS) (U.S. Department of Commerce [177]) in 2011, with parallel NWSRFS and CHPS modeling through 2012. In addition to QPF, principal hydrologic model forcings are observed precipitation and observed and forecasted temperature. Observed precipitation forcings are obtained from a multisensor estimation process, involving rain gauges, NWS *NEXt-generation RADar (NEXRAD)* doppler radar, and, at some RFCs, remotely-sensed satellite estimates of precipitation [92, 103, 202]. Forecasted precipitation is derived from numerical weather prediction (NWP) models, usually with meteorological forecaster adjustments made at both the NWS Weather Prediction Center (WPC) and/or at local RFCs [139].

5.2.2 Research goals

The aim of this research is to determine the utility of using hydrologic ensemble mean or median forecasts of river stage from the NOAA/NWS *Meteorological Model-based Ensemble Forecast System (MMEFS)*, described in Adams and Ostrowski [1], as an alternative to current, operational, single-valued deterministic hydrologic stage forecasts at the OHRFC and, possibly, elsewhere. Section 5.3 of this paper describes the real-time hydrologic forecasting experiment used in this study. Model simulations are restricted to watersheds in the OHRFC area of responsibility, shown in Figure 5.1. The experiment consists of concurrent generation of OHRFC operational river stage forecasts and MMEFS ensemble forecasts for the November 30, 2010 through May 24, 2012 period. Verification results of the ensemble median and mean forecasts relative to the OHRFC operational forecasts are presented in section 5.4. Experimental results are discussed in relation to verification of the MMEFS ensemble forecasts in section 5.5. Section 5.6 summarizes the experimental results and presents conclusions.

5.3 Research Approach

The approach of this study is to compare OHRFC operational forecasts to MMEFS ensemble mean and median forecasts that use numerical weather prediction (NWP) model precipitation and temperature output from the NOAA/NWS *National Centers for Environmental Prediction (NCEP)*, *North American Ensemble Forecast System (NAEFS)* [35] as hydrologic model forcings. The study period was November 30, 2010 through May 24, 2012. The NAEFS consists of 42 ensemble members. The research methodology includes:

1. capturing OHRFC operational forecasts initialized at 1200 UTC (daily), with a 5-day forecast horizon;

2. capturing automated MMEFS NAEFS hydrologic ensemble forecasts based on OHRFC 1200 UTC saved model states (daily), with a 7-day forecast horizon;

3. deterministic verification of operational forecasts and MMEFS NAEFS ensemble mean and median forecasts (after May 24, 2012);

4. verification of MMEFS NAEFS ensemble forecasts (after May 24, 2012).

The real-time hydrologic forecasts were made using the legacy NWSRFS, relying on a geographically broad distribution of forecast point locations, with varying basin sizes and hydrologic response times for the OHRFC. All model forcing inputs and internal and output time-steps are 6-hourly. A total of 54 basins, shown in Figure 5.2, were selected for the study. Calibrations of SAC-SMA, SNOW-17, channel routing, and reservoir simulation models for operational use for all OHRFC subbasins were completed long before the experiments started, following guidelines presented by Anderson [13] and Smith et al. [166]. Operational and MMEFS simulations utilize a 6-h time step for model forcings, internally, and output.

Forecasts are evaluated on the basis of comparisons between U.S. Geological Survey (USGS) observed stages and model estimated river stage values, which were transformed from simulated flow values using USGS station rating curves. Deterministic verification followed methods proposed by Welles et al. [181] and Demargne et al. [56].

The 54 study basins (Figure 5.2) are categorized as FAST (Table 5.1), MEDIUM (Table 5.2), and SLOW (Table 5.3) responding. These include 26 *fast*, 20 *medium*, and 8 *slow* responding forecast point locations. The terms *slow*, *medium*, and *fast* refer to hydrograph time-to-peak response times, from the center-of-mass of the observed precipitation to the hydrograph peak. Response times less than 24 h are classified as *FAST*, response times between 24 h to 60 h are considered *MEDIUM*, and response times greater than 60 h are considered *SLOW*, see OHD [142].

5.3.1 Operational legacy forecasts

The study relied on operational forecasts, using the OHRFC operational modeling system outlined in Section 5.2.1, covering the period November 30, 2010 - May 24, 2012. All operational forecasts used 24-h duration (4, 6-h periods $(24\ \text{h})^{-1}$) QPF. The experimental period spans 541 days at 54 locations, with 28 forecast periods each (4 6-h periods per day for 5-days), resulting in 817,992 forecast verification pairs for analysis. It should be pointed out that the operational forecasts used in this study include modeling of all 696 subbasins in the OHRFC area, an approximately 450,000 km^2 region, shown in Figure 5.2. Operational forecast horizons are 5-days.

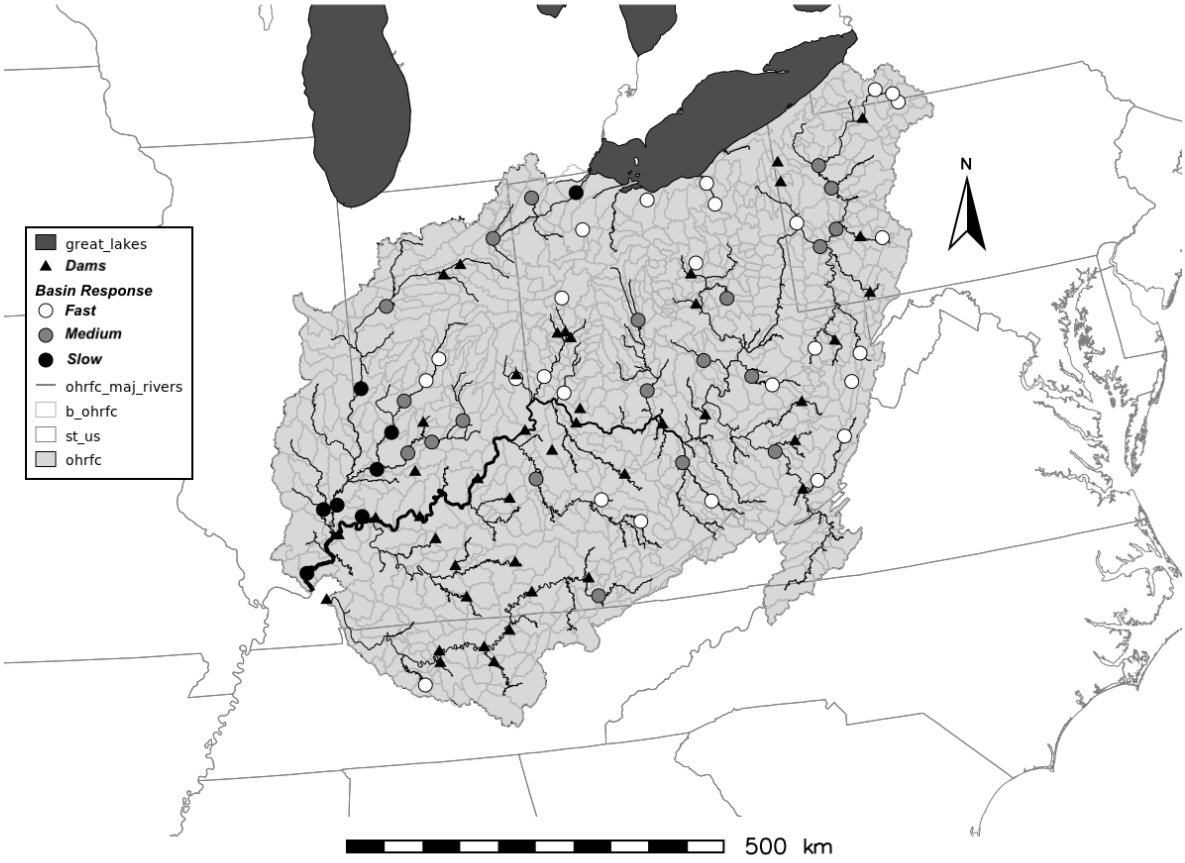

Figure 5.2: Map showing the location of 54 *Experiment* forecast point locations used in the OHRFC forecast area, listed in Tables 5.1, 5.2, and 5.3, identifying *fast*, *medium*, and *slow* responding basins. Locations of dams are shown with maximum storage capacities ≥250,000 ac-ft (308,370,000 m^3). Gray outlined polygons are 696 modeled subbasins.

Table 5.1: Fast response basins used in the study, listing NWS station identifier (ID), USGS identifier, Station name, basin area, and response time category.

ID	**USGS ID**	**Name**	**Area (km^2)**	**Response**
ALDW2	03183500	Greenbrier River at Alderson, WV	3533	Fast
BEAP1	03107500	Beaver River at Beaver Falls, PA	8044	Fast
BRKI3	03276000	East Fork Whitewater River at Brookville, IN	984	Fast
BUCW2	03182500	Greenbrier River at Buckeye, WV	1399	Fast
CLKW2	03059000	West Fork River at Clarksburg, WV	995	Fast
CNTI3	03354000	White River near Centerton, IN	6330	Fast
CYCK2	03283500	Red River at Clay City, KY	938	Fast
DLYW2	03050000	Tygart Valley River at Daily, WV	479	Fast
ELRP1	03010500	Allegheny River at Eldred, PA	1424	Fast
FDYO1	04189000	Blanchard River near Findlay, OH	896	Fast
FRAT1	03432350	Harpeth River at Franklin, TN	497	Fast
GRTW2	03153500	Little Kanawha River at Grantsville, WV	2365	Fast
HAMO1	03274000	Great Miami River at Hamilton OH	9402	Fast
INDI3	03353000	White River at Indianapolis, IN	4235	Fast
INDO1	04208000	Cuyahoga River at Independence OH	1831	Fast
JKNK2	03280000	North Fork Kentucky River at Jackson, KY	2852	Fast
KILO1	03139000	Kilbuck Creek at Kilbuck, OH	1202	Fast
MILO1	04199000	Huron River at Milan, OH	961	Fast
MLGO1	03245500	Little Miami River at Milford, OH	3116	Fast
OLNN6	03010820	Allegheny River at Olean, NY	3087	Fast
OLPO1	04206000	Cuyahoga River at Old Portage, OH	1046	Fast
PSNW2	03069500	Cheat River at Parsons, WV	1870	Fast
SIDO1	03261500	Great Miami River at Sidney, OH	1401	Fast
SLMN6	03011020	Allegheny River at Salamanca, NY	4165	Fast
SWDP1	03041500	Conemaugh River at Seward, PA	1852	Fast
WILW2	03213700	Tug Fork at Williamson, WV	2424	Fast

Table 5.2: Same as Table 5.1 but for *medium* response basins.

ID	USGS ID	Name	Area (km^2)	Response
ATHO1	03159500	Hocking River at Athens, OH	2442	Medium
BEDI3	03371500	East Fork White River near Bedford, IN	10000	Medium
CDIO1	03142000	Wills Creek at Cambridge, OH	1052	Medium
COLO1	03227500	Scioto River at Columbus, OH	4219	Medium
ELZW2	03155000	Little Kanawha River at Palestine, WV	3926	Medium
FFTK2	03287500	Kentucky River at Lock 4 at Frankfort, KY	13706	Medium
FLRK2	03215000	Big Sandy River at Fullers Station, KY	10093	Medium
FRKP1	03025500	Allegheny River at Franklin, PA	15493	Medium
FTWI3	04182900	Maumee River at Fort Wayne, IN	4988	Medium
KANW2	03193000	Kanawha River at Kanawha Falls, WV	21681	Medium
LAFI3	03335500	Wabash River at West Lafayette, IN	18121	Medium
NATP1	03049500	Allegheny River at Natrona, PA	29552	Medium
PARP1	03031500	Allegheny River at Parker, PA	19868	Medium
PKTO1	03237020	Scioto River at Piketon, OH	15115	Medium
PTTP1	03085152	Monongahela R. at Point State Park, Pittsburgh, PA	49471	Medium
SERI3	03365500	East Fork White River at Seymour, IN	6063	Medium
SHLI3	03373500	East Fork White River at Shoals, IN	12761	Medium
SPNI3	03357000	White River at Spencer, IN	7739	Medium
STRO1	04185000	Tiffin River at Stryker, OH	1062	Medium
WLBK2	03404000	Cumberland River at Williamsburg, KY	4162	Medium

Table 5.3: Same as Table 5.1 but for *slow* response basins.

ID	USGS ID	Name	Area (km^2)	Response
CARI2	03381500	Little Wabash River at Carmi, IL	8034	Slow
EVVI3	03322000	Ohio River at Evansville, IN	277600	Slow
GOLI2	03384500	Ohio River at Dam 51 at Golconda, IL	372699	Slow
HUFI3	03341500	Wabash River at Terre Haute, IN	31766	Slow
NHRI3	03378500	Wabash River at New Harmony, IN	75716	Slow
NWBI3	03360500	White River at Newberry, IN	12142	Slow
PTRI3	03373980	White River above Petersburg, IN	28808	Slow
WTVO1	04193500	Maumee River at Waterville, OH	16395	Slow

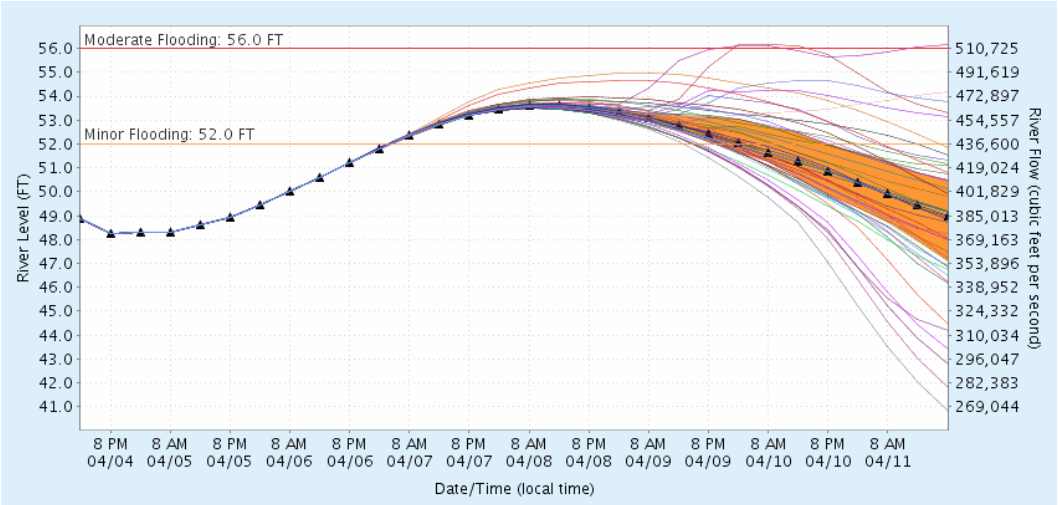

Figure 5.3: Example MMEFS NAEFS ensemble forecast, showing 42 individual ensemble model members (various colors), ensemble median (black line identified with triangles), and the 75% to 25% probability of exceedance confidence band is shown as the *orange* region. The *Minor* and *Moderate* flood levels are indicated for reference.

5.3.2 MMEFS ensemble forecasts

The automated MMEFS NAEFS hydrologic model ensemble simulations exactly parallel OHRFC operational forecasts. All simulations begin with 1200 UTC initializations, but they are not run until about 1800 UTC when NAEFS data is available from NCEP. Simulations utilize the full operational suite of models and follow identical operational workflows as the legacy deterministic OHRFC model forecast runs. Forecast horizons are 7-days. An example MMEFS NAEFS forecast is shown in Figure 5.3.

5.3.3 Forecast verification

Verification of the operational legacy forecasts use the *R Language and Environment for Statistical Computing* [145] and contributed *verification* package [135]. MMEFS NAEFS ensemble mean and median forecast verification statistics were obtained from ensemble anal-

yses utilizing the NOAA/NWS *Ensemble Verification Service (EVS)* [29, 57]. Operational forecast data are stored in the OHRFC *PostgreSQL* verification database and MMEFS simulations are written to NWSRFS *Ensemble Streamflow Prediction (ESP)* [51] format files. Verification measures used are *Mean Error (ME)*, *Mean Absolute Error (MAE)*, *Root Mean Square Error (RMSE)*, given in Equations 5.1, 5.2, and 5.3:

$$ME = \frac{1}{n}\sum_{k=1}^{n}(y_k - o_k) \tag{5.1}$$

$$MAE = \frac{1}{n}\sum_{k=1}^{n}(|y_k - o_k|) \tag{5.2}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(y_k - o_k)^2} \tag{5.3}$$

where the quantities y_k and o_k are the predicted and observed kth stage values, respectively, for n total paired values. Units of measure for *stage* are meters, unless reported otherwise. Values for *ME*, *MAE*, and $RMSE = 0$ implies perfect agreement, i.e., no error.

5.4 Study results

Verification results from the experiment are summarized in Figures 5.4 and 5.5. Mean error (ME), mean absolute error (MAE), and root mean square error (RMSE), based on *predicted* and *observed* stage pairs, are shown by forecast leadtime, in hours. Figure 5.4 compares results from *fast response* basins to the results for *all* basins. Figure 5.5 shows MAE for *medium response* and *slow response* basins. Several observations can be made from

Figure 5.4, namely,

1. In most instances, there is little difference between ensemble mean and median values, by leadtime. With the exception of RMSE, where ensemble mean values are smaller than ensemble median values, ensemble *median* values are always smaller in magnitude than ensemble *mean* values, which suggests that ensemble median forecasts should be preferred over ensemble mean forecasts, since less error is incurred;

2. With the the aggregation of all 54 basins, with respect to ME, little difference exists between the ensemble median forecast (and mean) and the OHRFC operational forecast (OHRFC 24-h QPF) through leadtime 54-h. Beginning with leadtime 60-h, OHRFC operational forecast become increasingly more negatively biased with increased leadtimes, whereas the MMEFS ensemble median forecasts remain unbiased through leadtime 168-h;

3. With *fast* response basins, the ensemble median forecast always shows ME values equal to or smaller in magnitude than the OHRFC operational forecasts, which get increasingly more negative with longer leadtimes after leadtime 72-h compared to ensemble median forecasts that very slowly become more negative with longer leadtimes;

4. OHRFC operational forecasts have smaller MAE values compared to MMEFS ensemble median and mean forecasts until leadtime 96-h, with all basins aggregated; however, for *fast response* basins, MMEFS ensemble median and mean forecasts have MAE values equal to or smaller than OHRFC operational forecasts beginning at about leadtime 36-h;

5. With respect to RMSE, with all basins aggregated, OHRFC operational forecasts exhibit smaller error compared to MMEFS ensemble median and mean forecasts until

leadtime 90-h, after which MMEFS ensemble median and mean forecast RMSE values are smaller than OHRFC operational forecast RMSE values; however, for *fast response* basins, MMEFS ensemble median and mean forecast RMSE values are approximately equal to or less than OHRFC operational forecast RMSE values beginning with leadtime 72-h.

Figure 5.5 shows that for *medium* and *slow* response basins, OHRFC operational forecasts tend to exhibit smaller forecast error compared to MMEFS ensemble median and mean forecasts until longer leadtimes are reached, $\geq$102-h. An explanation for this finding is discussed in Section 5.5.

5.5 Discussion

Three topics of discussion follow, (1) MMEFS ensemble median and mean forecasts compared to OHRFC operational forecasts, (2) MMEFS ensemble verification, and (3) needed improvements to the MMEFS. A discussion of MMEFS ensemble verification is needed to demonstrate that MMEFS NAEFS ensemble median and mean forecasts are derived from a system that has the properties of acceptable forecast skill, reliability, sharpness, and discrimination [185]. A demonstration of acceptable ensemble forecast verification results will provide a degree of confidence that the ensemble median and mean forecast are derived from a reasonably robust ensemble forecast system.

5.5.1 MMEFS ensemble median and mean forecasts

Results presented in Figure 5.4 clearly show, in Section 5.4, with all 54 basins aggregated, that MMEFS ensemble median and mean forecast add considerable value at longer leadtimes

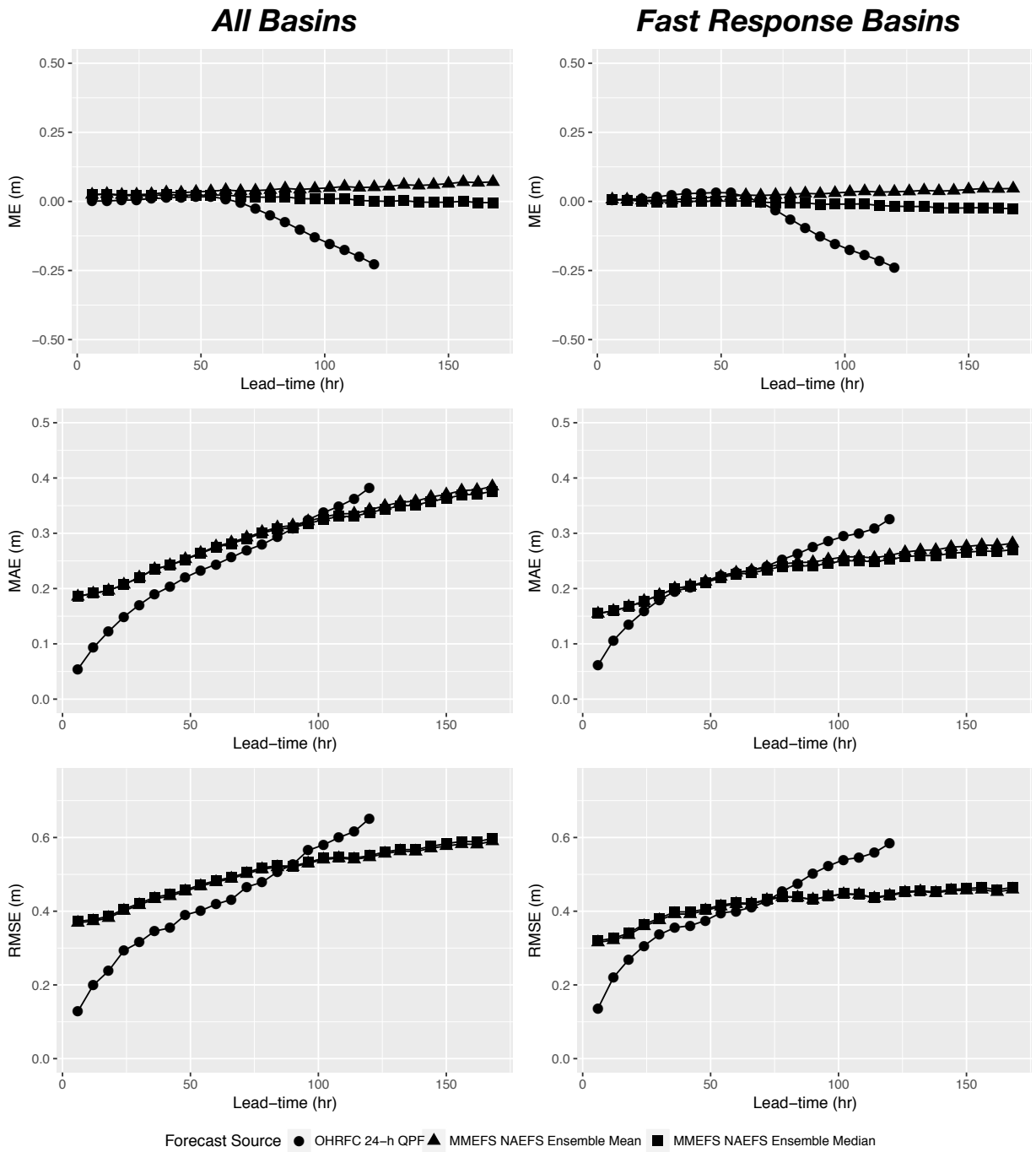

Figure 5.4: ME, MAE, and RMSE by leadtime for *All* and *Fast Response* basins identified in Figure 5.2 and in Tables 5.1, 5.2, and 5.3. Results are shown for operational forecast (OHRFC 24-h QPF) and MMEFS NAEFS ensemble mean and median forecasts, November 30, 2010 through May 24, 2012. Units are meters.

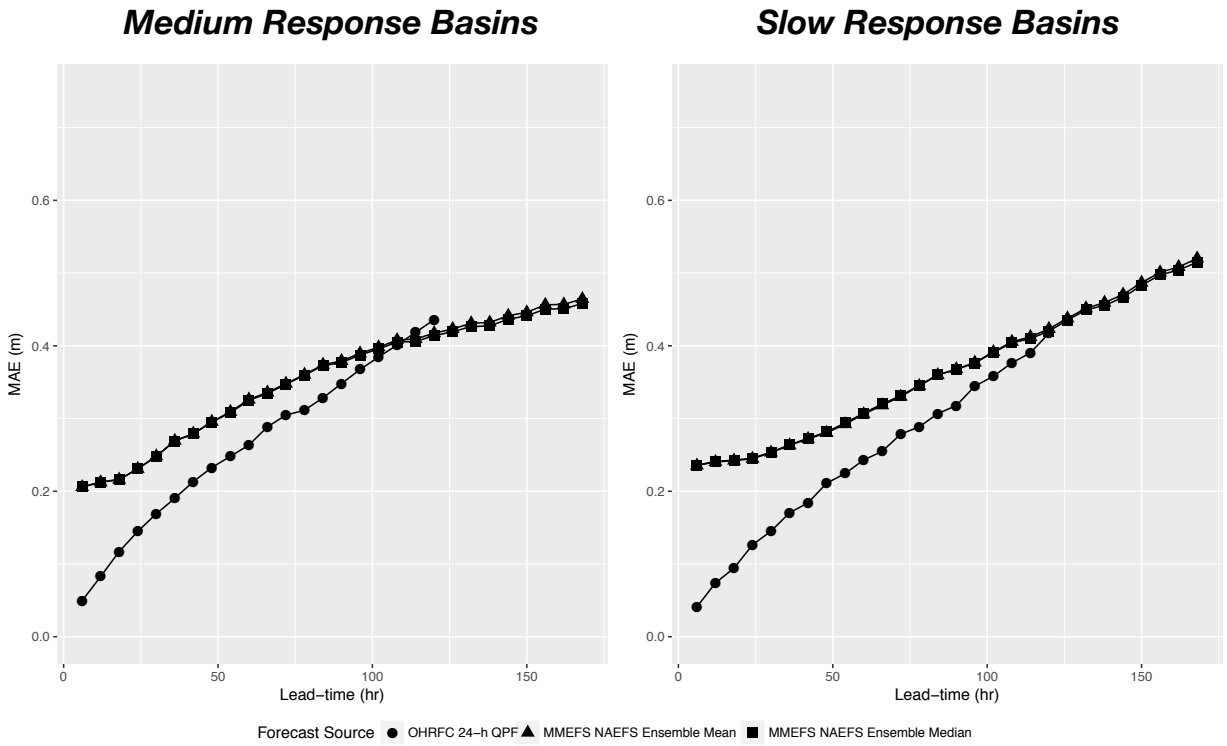

Figure 5.5: MAE by leadtime for *Medium* and *Slow* response basins identified in Figure 5.2 and in Tables 5.2 and 5.3. Results are shown for operational forecast (OHRFC 24-h QPF) and MMEFS NAEFS ensemble mean and median forecasts, November 30, 2010 through May 24, 2012. Units are meters.

above OHRFC operational forecasts, at leadtimes ≥90-h. An important result is that the MMEFS ensemble median and mean forecasts for *fast response* basins have smaller forecast errors than basins with longer basin response times beginning with leadtimes ≥36-h for MAE and for ME, MMEFS ensemble median and mean forecasts, for *fast response* basins, display smaller error immediately, at leadtime 6-h. This finding is important because it demonstrates the viability of using MMEFS ensemble median and mean forecast as an alternative to the continued use of OHRFC operational forecasts relying on 24-h or, currently, 48-h QPF.

Several factors influence the disparity in MMEFS ensemble median and mean forecast verification results for *medium* and *slow* response basins compared to *fast* response basins. These include:

1. Model calibrations are good for *fast* response, headwater basins, but calibrations for *medium* and *slow* response basins are more problematic. Calibration for non-headwater basins is more difficult due to, in part, overwhelming upstream flow influences that can not be accurately separated from total observed streamflow at local non-headwater gauges. The separation of flows routed from upstream basins from the observed *total streamflow* is necessary to estimate the *local observed streamflow*. With the *local observed streamflow*, model parameter adjustments to the local basin can be made appropriately to reflect the hydrologic response of the local watershed being calibrated. The consequence is that hydrologic model parameters used for *downstream* basins usually can only be estimated, not calibrated. The outcome is that local downstream basin simulations can be considerably more in error than *fast response* headwater basins that have been calibrated. Also, overall modeling for downstream *medium* and *slow* response basin is more complex, involving the use of flow routing and reservoir simulation models, which adds to modeling uncertainty and error;

2. For some *medium* and all *slow* response basins, complex channel flow dynamics, including hydrodynamic backwater effects, adversely affects stage-discharge relationships and, on the mainstem of the Ohio River, lock & dam controls influence stage-discharge relationships. The implication here is that, while flow simulations *may* be good, stage-discharge relationships are non-unique because of *hysteresis*. This non-unique stage-discharge relationship precludes unique conversion from flow to stage. Consequently, considerable error is incurred, which adversely influences verification scores due to the introduction of erroneous stage values. Manual forecaster adjustments in the OHRFC operational workflow are made, in part, to minimize stage-discharge ratings *hysteresis* effects. Such adjustments are not possible in MMEFS NAEFS simulations. The need for dynamic flow routing modeling served as the basis for the development and operational implementation of the *Ohio River Community HEC-RAS Model* into CHPS [8]. Unfortunately, this model was not used in the MMEFS NAEFS simulations;

3. Reservoir releases significantly alter downstream flows. OHRFC operational forecasts benefit from the inclusion of U.S. Army Corps of Engineers (USACE) deterministic reservoir release schedules, including flow releases from locks & dams, which are incorporated into the operational forecasts. MMEFS modeling relies on NWSRFS (now CHPS) based reservoir model simulations, which, under many scenarios, can be quite erroneous compared to actual USACE reservoir releases, which involves human decision making that is difficult to capture in a reservoir simulation model. Figure 5.2 shows that all *slow* and all but two *medium* response basin locations are downstream of significant reservoirs with maximum storage capacities $\geq$250,000 ac-ft (308,370,000 m^3). All of the *fast* response basins are upstream of these reservoirs. Consequently, OHRFC operational forecast errors are reduced in the near-term relative to MMEFS NAEFS ensemble median and mean forecasts that rely solely on model simulations of

reservoir outflows.

5.5.2 Ensemble verification

Demargne et al. [57] discuss the need for verification of hydrologic ensemble forecasts, identifying the need to improve research and operations, specifically aimed at (1) monitoring changes in forecast quality over time, (2) analyzing sources of forecast error, and (3) evaluating forecast skill improvements resulting from the introduction of new science and technology. Consequently, we present some ensemble verification results to serve as a baseline evaluation of MMEFS forecast quality to identify areas of needed system improvement, which should further reduce MMEFS ensemble median and mean prediction errors.

Many statistical measures have been used to evaluate probabilistic forecasts. Johnson and Bowler [97] suggests that, for example, *resolution*, the property that there should be large variability of observed frequencies associated with different forecast probabilities around the climatological value, is desirable. Also, ensemble forecasts should be *reliable*, that is, forecast probabilities should give an estimate of the expected frequencies of the event occurring. Welles et al. [181] and Demargne et al. [56] explain the necessity of hydrologic forecast verification. Hydrologic ensemble forecast verification methods, as recommended and discussed by Brown et al. [29] and Demargne et al. [57] are used to assess the MMEFS NAEFS ensemble forecasts.

The statistical measures to evaluate MMEFS ensemble forecasts are discussed below.

Continuous ranked probability skill score (CRPSS)

The ranked probability score (RPS), shown in Equation 5.4, is a measure of how well forecasts, that are expressed as probability distributions, are in matching observed outcomes

$$\text{RPS} = \frac{1}{r-1} \sum_{i=1}^{r} \left(\sum_{j=1}^{i} p_j - \sum_{j=1}^{i} e_j \right)^2, \tag{5.4}$$

where r is the number of outcomes, p_j is the forecasted probability of outcome j and e_j is the actual probability of outcome j. As a note, the special case where $r = 2$ gives the Brier score [185]. The RPS applies to probability forecasts for discrete categories and the *continuous ranked probability score (CRPS)* extends the measure to continuous forecasts. Bradley and Schwartz [24] show that the *continuous ranked probability skill score (CRPSS)* (Equation 5.7), is a summary measure representing the weighted-average skill score, using climatology as the reference forecast, $CRPS_{clim}$ in Equation 5.7, over the continuous range of outcomes y. The CRPSS, which can be derived from the *mean square error (MSE)*, is given by :

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^{r} \left(Y_i - \hat{Y}_i \right)^2, \tag{5.5}$$

$$\overline{\text{CRPS}} = \int_{-\infty}^{\infty} MSE(y) dy, \tag{5.6}$$

$$\overline{\text{CRPSS}} = 1 - \frac{\overline{CRPS}}{\overline{CRPS_{clim}}}, \tag{5.7}$$

where Y_i and $\hat{Y}_i$ are the *observed* and *predicted* values, respectively, of a verification pair. The overbar ($\overline{CRPS}$) refers to averaging of CRPS values across the sample of events. CRPSS values can range from $-\infty$ to 1, with perfect skill equal to 1 and negative values when the forecast has worse CRPS than the reference forecast.

Figure 5.6 summarizes the forecast skill of all 54 basins, with aggregation across all forecast

stage ranges (a) and for stage ranges ≥0.90 probability of non-exceedance (b). The results show that MMEFS NAEFS forecasts are skillful relative to *sample climatology* and that forecasts for stage ranges ≥0.90 probability of non-exceedance are more skillful than forecasts which are aggregated across all stage ranges, which is encouraging since RFCs emphasize flood forecasting. Also depicted in Figure 5.6 is differentiation between *fast*, *medium*, and *slow* response basins, which shows MMEFS ensemble forecasts to be most skillful for *slow* response basins and that forecast skill is most variable for *medium* response basins. Forecast skill, as expected, declines with increased leadtimes.

An interesting point relates to some of the lowest CRPSS values in Figure 5.6(a) which correspond to the forecast point location at Pittsburgh, PA (PTTP1). Referring to Table 5.4, CRPSS values aggregated across all forecast stage ranges demonstrate little to no skill. However, for stage ranges ≥0.90 probability of non-exceedance, MMEFS ensemble forecasts show reasonable skill. The reason for this difference relates to the complex physical setting described by Adams et al. [8] involving downstream control at Dashields, PA Lock & Dam that regulates the pool level at the PTTP1 streamgauge at low flows in a manner that is not well-captured by the OHRFC modeling system. With higher flows, particularly flood flows, this control does not exist.

Reliability diagram

Reliability diagrams [90] represent, graphically, the *observed* frequency of an event plotted against the *forecast* probability of an event. This expresses how often (as a relative frequency) a forecast probability actually occurred. The hit rate is calculated from the sets of forecasts for each probability separately. Consequently, the *hit rate* for each probability bin, n, is given by:

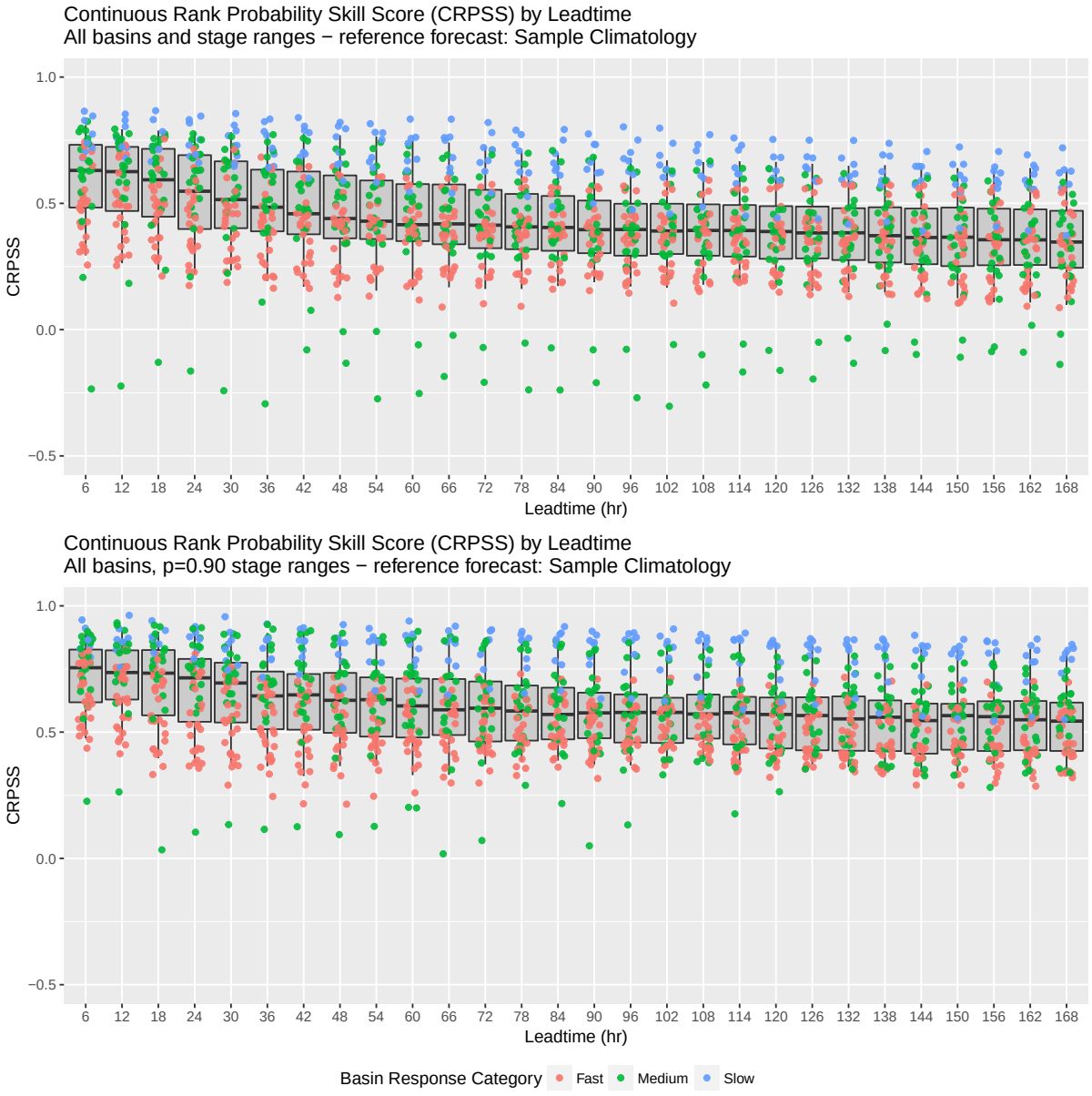

Figure 5.6: CRPSS by leadtime for all forecast point locations identified in Figure 5.2, for all forecast stage ranges and stage ranges ≥0.90 probability of non-exceedance. Point shading identifies basin response category. Units are dimensionless.

Table 5.4: CRPSS for Pittsburgh, PA (PTTP1) for all stage ranges and for stages with probability of exceedance, p=0.90, by Leadtime.

Leadtime (hr)	**CRPSS (all)**	**CRPSS (p=0.90)**
6	0.2172	0.5758
24	0.2294	0.6261
48	-0.0265	0.4014
72	-0.0751	0.3930
96	-0.1014	0.3686
120	-0.0834	0.3810
144	-0.0800	0.3527
168	-0.1182	0.3351

$$\text{Hit Rate}_\text{n} = \frac{O_n}{O_n + N_n}, \tag{5.8}$$

$$\text{F}_\text{n} = \frac{O_n + N_n}{T}, \tag{5.9}$$

where F is the forecast frequency, O is the number of observed instances, N is the number of non-observed instances, and T is the total number of forecasts.

Figure 5.7 shows the reliability diagram for stage ranges $\geq$0.90 probability of non-exceedance, for leadtimes 24-, 48-, 96-, 120-, and 168-h, aggregated across all 54 basins. Generally, the ensemble forecasts show reasonable reliability, but that forecast over-confidence exists between forecast probability ranges 0.50 to 0.75. Results for leadtime 24-h is, most likely, representative of small sample size problems, which is a general concern since the study period was short, only 541 days, November 30, 2010 through May 24, 2012.

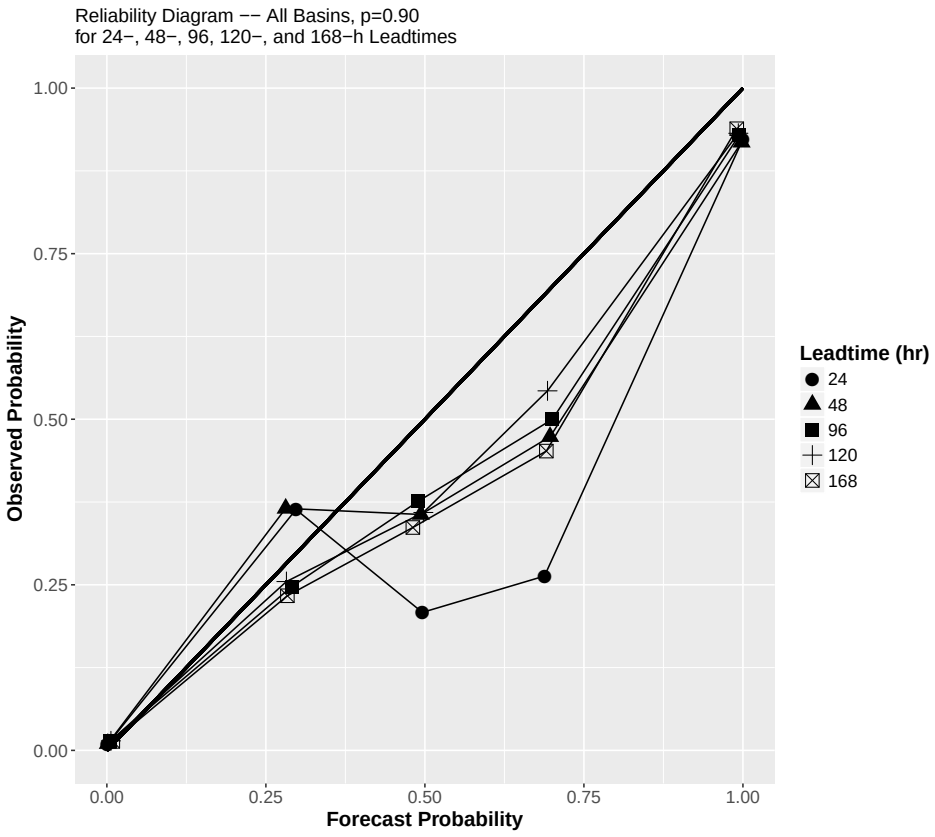

Figure 5.7: Reliability Diagram for all 54 basins, for lead-times 24-, 48-, 96-, 120-, and 168-h. Shown for stage ranges $\geq$0.90 probability of non-exceedance.

	Observed	
Forecast	yes	no
yes	a	b
no	c	d

Table 5.5: Contingency table.

Relative Operating Characteristic (ROC)

The *Relative Operating Characteristic (ROC)* [100] provides information on the hit rates and false alarm rates that can be expected from use of different probability thresholds. The ROC is a summary score used to describe the ability of forecasts to discriminate between events and non-events. From Table 5.5:

$$\text{Hit Rate} = \frac{a}{a+c} \tag{5.10}$$

$$\text{False Alarm Rate} = \frac{b}{b+d} \tag{5.11}$$

Figure 5.8 shows the ROC diagram, aggregated across all 54 basins, for stage ranges $\geq$0.90 probability of non-exceedance, for forecast leadtimes 24-, 96-, 120-, and 168-h, indicating that MMEFS NAEFS ensemble forecasts discriminate between events and non-events very well. This result is representative of similar analyses, done for individual basins and for all forecast stage ranges.

Rank Histogram

Rank histograms are useful for evaluating ensemble forecasts because they can efficaciously assess the reliability and errors in the mean and spread of ensemble forecasts [87]. Rank his-

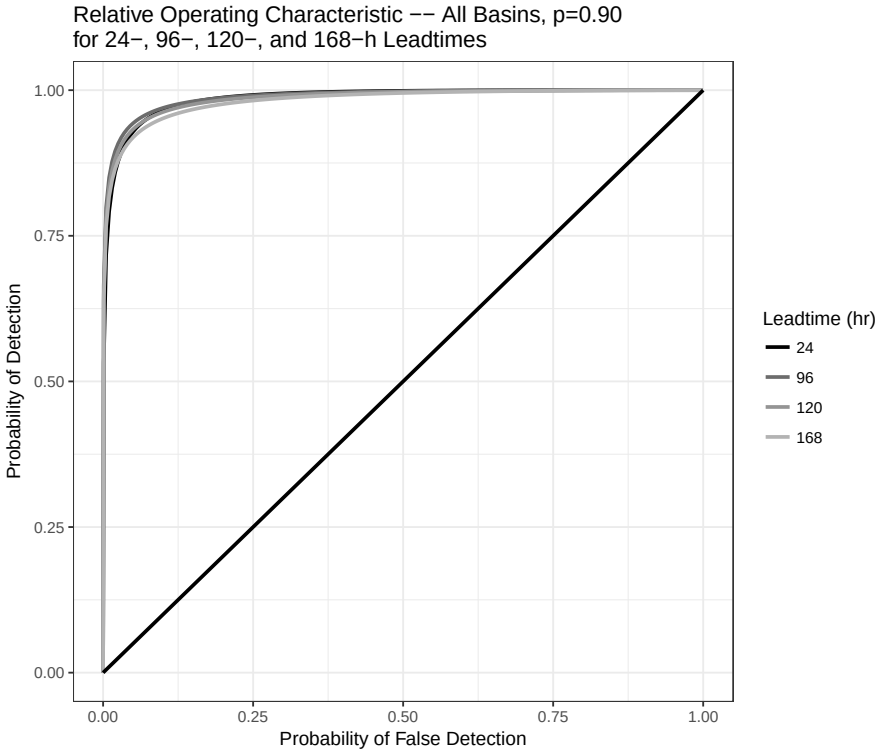

Figure 5.8: ROC for all 54 basins, for lead-times 24-, 96-, 120-, and 168-h. Shown for stage ranges $\geq$0.90 probability of non-exceedance.

tograms are created by tallying the rank of observations relative to values from an ensemble sorted from lowest to highest rank, which produces, ideally, a uniform distribution across the ranks.

Rank histogram analyses are summarized by Figure 5.9, which shows severe under-spread of MMEFS ensemble forecasts, particularly at shorter leadtimes, illustrated for leadtime 24-h. An explanation for this result is shown in Figure 5.3, which was specifically selected to illustrate the under-spread problem. In this example, the 42 MMEFS NAEFS ensembles appear as a single-valued forecast from the beginning of the forecast at 2 PM 04/04 through 8 AM 04/07. This occurred because there were (1) no model forcings to substantially perturbe the hydrologic models over the 2 PM 04/04 through 8 AM 04/07 period and (2) because inherent model error is not included in the MMEFS.

5.5.3 MMEFS improvements

MMEFS ensemble forecasts do not currently make use of ensemble model-error correction methods suggested by, for example, Bogner and Kalas [19], Li et al. [115] or post-processing bias correction of raw ensemble forecasts proposed by Brown and Seo [28], Hashino et al. [91], Yuan and Wood [197] and others to reduce uncertainties arising from model inputs and outputs, initial and boundary conditions, and the structure and parameter estimates of models. Wentao et al. [182] recently reviewed statistical postprocessing methods for hydrometeorological ensemble forecasting, citing the need for further work on many fronts. These concerns include the need to address stationarity assumptions, handle extreme events, including the timing of flood peaks, in the case streamflow modeling, further investigate methods proposed to make adjustments at un-gauged locations, and continue research into methods that attempt to address total uncertainty, including model structure, parameter estimation, and

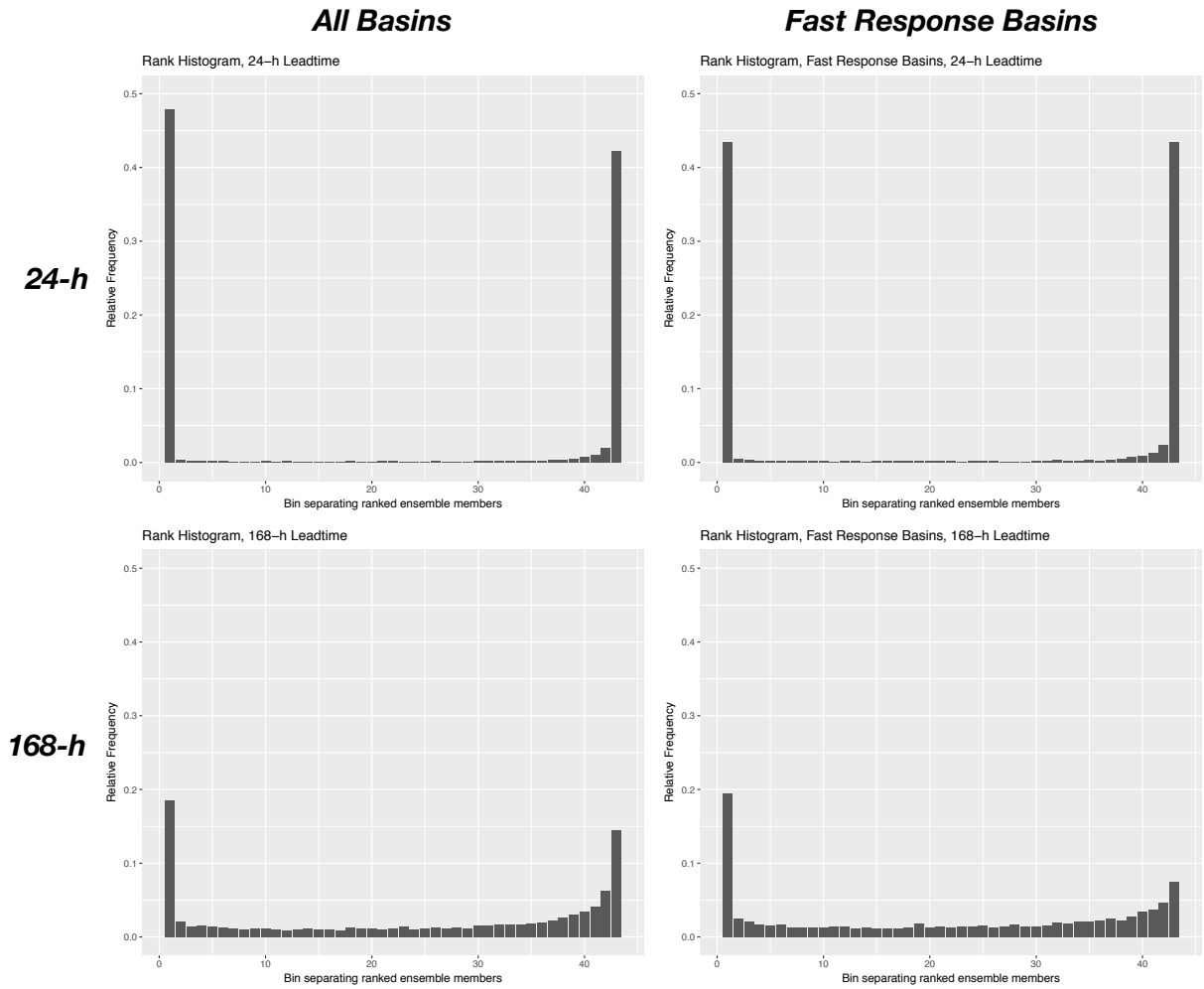

Figure 5.9: Rank histograms for all forecast point locations identified in Figure 5.2 and for *fast* response basins, for 24- and 168-h leadtimes.

model initial and boundary conditions. The need for the use of such techniques is illustrated in Section 5.5.2 where MMEFS ensemble verification results are discussed related to ensemble under-spread and apparent biases. Although not strictly an ensemble modeling related issue, MMEFS ensemble forecasts would benefit from improved hydrologic model calibrations for many basins. Model simulation error would be reduced further and MMEFS verification would also improve by incorporating the *Ohio River Community HEC-RAS Model* directly into MMEFS ensemble simulations for slow-response rivers where complex hydrodynamics are inadequately handled by simple streamflow routing models. MMEFS ensemble verification metrics are clearly better for basins known to have good calibrations than for basins with sub-optimal model calibrations.

5.6 Summary and conclusions

Experimental results from this study demonstrate that NAEFS based MMEFS ensemble median forecasts have smaller forecast error than ensemble mean forecasts based on ME and MAE verification measures. Even though RMSE values suggest slightly lower forecast error with ensemble mean compared to ensemble median forecasts, analyses, overall, suggest that ensemble median forecasts should be preferred over ensemble mean forecasts. More importantly, when forecasts spanning all ranges in stage and basin response times are aggregated, MMEFS ensemble mean and median forecasts show lower forecast error than legacy OHRFC operational forecasts, based on 24-h deterministic QPF, at long forecast leadtimes, that is, beginning at approximately $\geq$90-h. This result shows the viability of using ensemble mean/median forecasts for extended forecasts beyond the 4-day forecast horizon. When analyses are restricted to *fast* response basins only, MMEFS ensemble mean and median forecasts have smaller forecast error than legacy OHRFC operational forecasts beginning at

about leadtime 36-h. This finding has potentially significant implications on forecast operations at the OHRFC and other hydrologic forecast centers. Specifically, lower MMEFS ensemble mean/median forecast errors compared to the current operational OHRFC deterministic forecasts suggests the feasibility of changing operational workflows from manually intensive, interactive forecasting procedures to a more automated operational environment using ensemble forecasting methodologies. Of course, the generation of full probabilistic forecasts from ensemble methods is greatly preferred over the use of the ensemble mean or median because the former conveys forecast uncertainty to end-users, where the latter does not [133]. In addition, a wide range of water resources applications requires the use of flow hydrographs; the proposed ensemble median/mean *stage* forecasts should not be translated to flow hydrographs because mass balance is not preserved. Consequently, for most other water resources applications, full hydrologic ensemble forecasts are needed.

MMEFS ensemble verification results covering the November 30, 2010 through May 24, 2012 study period demonstrate forecast skill and reasonable forecast discrimination, reliability, and sharpness. However, verification results also identify needed areas of improvement, such as the need to account for model error, utilizing *a priori* hindcast experiments to quantify model error, and the adoption of a post-forecast, ensemble bias correction methodology.

Chapter 6

Conclusions

The results of each Chapter of this dissertation are presented within the individual Chapters, with their associated conclusions. The conclusions presented here summarize the results and conclusions drawn from the entire dissertation (Section 6.1). Section 6.2 discusses the significance of the research within the broader hydrologic community, especially how the findings could impact decision-making by those establishing policy. Finally, the *Future Work* section, Section 6.3, summarizes future research that is needed to fill knowledge gaps identified or implied in the individual Chapters within the dissertation.

6.1 Summary

The main objective of the research was to demonstrate the necessity for using probabilistic hydrologic forecasting in place of current forecast methods that rely on the use single-valued deterministic QPF. This goal was reached, first, by identifying that QPE improvements have been significant, which has lead directly to significantly improved hydrologic modeling simulations, indicated by accepted measures of model error estimation. Second, Chapter 3

illustrated that QPF improvements have been marginal over the period 1970-2015. Moreover, a hydrologic *monte carlo* simulation experiment demonstrated that the small gains in QPF accuracy did not produce significant reductions in hydrologic forecast error. On the basis of QPF *threat score (TS)*, an important finding was that very large hydrologic forecast errors could still result from significantly reduced QPF error. Moreover, results from the *monte carlo* experiment showed that there should be no expectation that a high QPF TS value should be associated with a smaller hydrologic forecast error than a lower TS value. The implication is that there are spatial scale effects that relate to watershed size and the intra-storm variability of areas of strong convection that complicate the interaction between cores of heavy rainfall and watershed response.

Experiment-1 of Chapter 4 demonstrated that hydrologic forecasts using *non-Zero QPF* had smaller errors than forecasts that used *zero-QPF*. Experiment-2 showed that for non-flood forecasts, forecast error diminished with longer durations of QPF. However, for *flood forecasts*, this finding was reversed; that is, increased durations of QPF increased hydrologic forecast error. Moreover, for *fast response* basins, Experiment-2 showed that forecast error increased dramatically when QPF durations changed from 12-h to 24-h or greater. This is problematic because one does not know beforehand whether or not flooding will occur. Consequently, the findings suggest that QPF should be restricted to 6- or 12-h durations. This implies that hydrologic forecasts for non-flood events, which is by far, most common, will experience reduced forecast accuracy. This presents a paradox with the use of QPF for *flood* and *non-flood* forecasting.

The answer to the QPF paradox, of course, is that ensemble hydrologic forecasting should be preferred over current deterministic forecasting approaches, which is demonstrated in Chapter 5. In Chapter 5, the use of hydrologic ensemble median and mean forecasts was explored as a mechanism to reduce hydrologic forecast uncertainty. The operational experi-

ment illustrated that when results from all 54 experimental basins are aggregated, forecast verification showed that the MMEFS NAEFS ensemble median and mean forecasts had less error than OHRFC operational forecasts beginning at lead-time 90-h. However, with *fast response* basins, ensemble median and mean forecasts had less error than OHRFC operational forecasts beginning at lead-time 36-h, based on *mean absolute error* and at lead-time 6-h for *mean error*. This demonstrates the viability of using ensemble median and mean forecasts as a replacement for OHRFC operational forecasts. Moreover, the use of hydrologic ensemble median or mean forecasts includes the notion of uncertainty in hydrologic forecasts in terms of *expectation* or "best estimate". Since an ensemble median or mean *stage* forecast *looks like* a single-valued, deterministic hydrologic forecast, acceptance by the general public and decision makers should be achievable. A minor finding was that ensemble median forecasts had smaller error than ensemble mean forecasts, based on *mean absolute error* and *mean error*. Differences in *root mean square error* were nearly negligible.

6.2 Engineering Significance

The most significant outcome of the research presented in this dissertation is a demonstration of the *necessity* for using ensemble hydrologic forecasting methods over current operational methods that rely on single-valued QPF. The research demonstrated that very large hydrologic forecast errors are introduced by single-valued QPF, particularly for rainfall driven events. The conclusion is that traditional QPF should not be used in hydrologic forecasting. While the hydrologic forecast community has been moving in the direction of ensemble forecasting methods, national hydrologic services (NHSs), such as the NWS in the U.S., have been very slow to adopt ensemble methods operationally to *replace* legacy forecasting methods and forecast products that convey forecast uncertainty. The general public and decision

makers need to receive flood alerts and warnings in the form of probabilistic statements and graphics. Research presented in this dissertation shows that ensemble median forecasts are a viable replacement for current operational forecasts that rely on deterministic QPF. Ensemble median and mean forecast have the added benefit of conveying a measure of forecast uncertainty in much the same manner that, with hurricane or tropical storm prediction of storm tracks, meteorological forecasters display both the *expected* storm track (the ensemble mean) and the individual storm tracks from numerous numerical weather prediction models.

Finally, research presented in this dissertation showed that significant improvements in the accuracy of hydrologic modeling have resulted from scientific advances in remote sensing of precipitation by radar and development of improved radar-precipitation estimation systems.

6.3 Future Work

Research presented in this dissertation underscores the need for improved hydrologic model calibrations to minimize forecast errors. The research also demonstrates the value of improved precipitation estimation methods that reduce hydrologic simulation hindcast and forecast errors. Further reduction of modeling and forecast error will result from the incorporation of hydrodynamic and improved reservoir simulation models into hydrologic forecast systems, which will also improve the reliability of ensemble forecast systems.

The MMEFS ensemble forecast experimental period was too short, resulting in small sample sizes that affected verification statistics. Future verification efforts of the MMEFS are needed to better evaluate MMEFS strengths and areas of needed improvement. Consequently, future analyses should include longer experimental periods, on the order of years. This requirement is especially needed for smaller, fast responding basins that experience flooding infrequently. MMEFS verification results show clear under-spread of ensemble members, largely because

model error is not represented in MMEFS forecasts. Consequently, future MMEFS development should include methods to account for model error, utilizing *a priori* hindcast experiments to quantify model error, and the adoption of a post-forecast, ensemble bias correction methodology.

The greatest need for future work falls into three categories:

1. Education of forecasters, NHS managers and policymakers, the public, decision-makers, emergency managers, and politicians to promote an understanding of probabilistic forecasts and the necessity of adopting probabilistic hydrologic forecasting over current single-valued, deterministic forecasts;

2. Creation of forecast *products* in the form of graphics and digital and written statements that convey probabilistic forecasts that are meaningful to and usable by end-users. On this point, it would be a mistake to believe that there is, say, a single graphical representation of a probabilistic forecast that will be meaningful to all. A range of representations are needed that convey probabilistic forecasts that meet the needs of end-users, ranging from sophisticated to naïve;

3. Raw and bias-corrected ensemble forecast time-series data should be made publicly available to researchers and emergency and water resources managers for use in decision support systems. This implies the need for the creation of a standardized format for ensemble time-series data transmission and storage.

Bibliography

[1] T. Adams and J. Ostrowski. Short lead-time hydrologic ensemble forecasts from numerical weather prediction model ensembles. In *Proceedings World Environmental and Water Resources Congress 2010*, Providence, R.I., 2010. EWRI.

[2] T. Adams and G. Smith. National Weather Service Interactive River Forecasting Using State, Parameter, and Data Modifications. In *Proceedings of the International Symposium on Engineering Hydrology*, San Francisco, CA, 1993. EWRI.

[3] Thomas E. Adams. An assessment of the intra-storm spatial variability of floods in the Piedmont region of Maryland. Master's thesis, Virginia Tech, Blacksburg, VA, 1986.

[4] Thomas E. Adams. Flood Forecasting in the United States NOAA/National Weather Service. In Thomas E. Adams and Thomas C. Pagano, editors, *Flood Forecasting: a global perspective*, chapter 12, pages 275–325. Elsevier/Academic Press., New York, NY, first edition, 2016.

[5] Thomas E. Adams. Flood Forecasting in the United States NOAA/National Weather Service. In Thomas E. Adams and Thomas C. Pagano, editors, *Flood Forecasting: a global perspective*, chapter 12, pages 275–325. Elsevier/Academic Press., New York, NY, first edition, 2016.

[6] Thomas E. Adams and Randle Dymond. Hydrometeorological Forcing Errors for a Real-time Flood Forecast System in the Ohio River Valley, USA. *Journal of Hydrometeorology*, 2018, submitted to the AMS J Hydrometeorology.

[7] Thomas E. Adams and Randle Dymond. The Effect of QPF on Real-time Deterministic Hydrologic Forecast Uncertainty. *Journal of Hydrometeorology*, 2018b, submitted to the AMS J Hydrometeorology.

[8] Thomas E. Adams, Sherry Chen, and Randle Dymond. Results from Operational Hydrologic Forecasts using the NOAA/NWS OHRFC Ohio River Community HEC-RAS Model. *Journal of Hydrologic Engineering*, in press 2018.

[9] L. Alfieri, D. Velasco, and J. Thielen. Flash flood detection through a multi-stage probabilistic warning system for heavy precipitation events. *Advances in Geosciences*, 29:69–75, 2011. doi: 10.5194/adgeo-29-69-2011. URL https://www.adv-geosci.net/29/69/2011/.

[10] L. Alfieri, P. Burek, E. Dutra, B. Krzeminski, D. Muraro, J. Thielen, and F. Pappenberger. Glofas – global ensemble streamflow forecasting and flood early warning. *Hydrology and Earth System Sciences*, 17(3):1161–1175, 2013. doi: 10.5194/hess-17-1161-2013. URL https://www.hydrol-earth-syst-sci.net/17/1161/2013/.

[11] E.N. Anagnostou, W.F. Krajewski, D.-J. Seo, and E.R. Johnson. Mean-field radar rainfall bias studies for WSR-88D. *ASCE Journal of Engineering Hydrology*, 3(3): 149–159, 1998.

[12] E. A. Anderson. National Weather Service River Forecast System-Snow Accumulation and Ablation Model. Technical Report NWS-HYDRO-17, U.S. National Weather Service, Office of Hydrology, Hydrology Laboratory. Technical Memo., 1325 East West Highway, Room 8372 Silver Spring, Maryland 20910, 1973.

[13] Eric A. Anderson. Calibration of conceptual hydrologic models for use in river forecasting. Technical report, U.S. National Weather Service, Office of Hydrology, Hydrology Laboratory, 2002.

[14] D. Anghileri, N. Voisin, A. Castelletti, F. Pianosi, B. Nijssen, and D. P. Lettenmaier. Value of long-term streamflow forecasts to reservoir operations for water supply in snow-dominated river catchments. *Water Resour. Res.*, 52, 2016. doi: doi:10.1002/2015WR017864.

[15] Mary Lynn Baeck and James A. Smith. Rainfall estimation by the wsr-88d for heavy rainfall events. *Weather and Forecasting*, 13(2):416–436, 1998. doi: 10.1175/1520-0434(1998)013⟨0416:REBTWF⟩2.0.CO;2.

[16] Satish Bastola, Vasubandhu Misra, and Haiqin Li. Seasonal hydrological forecasts for watersheds over the southeastern united states for the boreal summer and fall seasons. *Earth Interactions*, 17(25):1–22, 2013. doi: 10.1175/2013EI000519.1. URL https://doi.org/10.1175/2013EI000519.1.

[17] J. V. L. Beckers, A. H. Weerts, E. Tijdeman, and E. Welles. Enso-conditioned weather resampling method for seasonal ensemble streamflow prediction. *Hydrology and Earth System Sciences*, 20(8):3277–3287, 2016. doi: 10.5194/hess-20-3277-2016. URL https://www.hydrol-earth-syst-sci.net/20/3277/2016/.

[18] Kurt K. Benke, Kim E. Lowell, and Andrew J. Hamilton. Parameter uncertainty, sensitivity analysis and prediction error in a water-balance hydrological model. *Mathematical and Computer Modelling*, 47(11):1134 – 1149, 2008. ISSN 0895-7177. doi: https://doi.org/10.1016/j.mcm.2007.05.017. URL http://www.sciencedirect.com/science/article/pii/S0895717707002373.

[19] K. Bogner and M. Kalas. Error?correction methods and evaluation of an ensemble based hydrological forecasting system for the upper danube catchment. *Atmospheric Science Letters*, 9(2):95–102, 2008. doi: 10.1002/asl.180. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/asl.180.

[20] K. Bogner and F. Pappenberger. Multiscale error analysis, correction, and predictive uncertainty estimation in a flood forecasting system. *Water Resources Research*, 47 (7):n/a–n/a, 2011. ISSN 1944-7973. doi: 10.1029/2010WR009137. URL http://dx.doi.org/10.1029/2010WR009137. W07524.

[21] M. Borga, E.N. Anagnostou, G. Blschl, and J.-D. Creutin. Flash floods: Observations and analysis of hydro-meteorological controls. *Journal of Hydrology*, 394(12):1 – 3, 2010. ISSN 0022-1694. doi: https://doi.org/10.1016/j.jhydrol.2010.07.048. URL https://www.sciencedirect.com/science/article/pii/S0022169410004877. Flash Floods: Observations and Analysis of Hydrometeorological Controls.

[22] Philippe Bougeault, Zoltan Toth, Craig Bishop, Barbara Brown, David Burridge, De Hui Chen, Beth Ebert, Manuel Fuentes, Thomas M. Hamill, Ken Mylne, Jean Nicolau, Tiziana Paccagnella, Young-Youn Park, David Parsons, Baudouin Raoult, Doug Schuster, Pedro Silva Dias, Richard Swinbank, Yoshiaki Takeuchi, Warren Tennant, Laurence Wilson, and Steve Worley. The thorpex interactive grand global ensemble. *Bulletin of the American Meteorological Society*, 91(8):1059–1072, 2010. doi: 10.1175/2010BAMS2853.1. URL https://doi.org/10.1175/2010BAMS2853.1.

[23] L. C. Bowling, J. W. Pomeroy, and D. P. Lettenmaier. Parameterization of blowing-snow sublimation in a macroscale hydrology model. *Journal of Hydrometeorology*, 5 (5):745–762, 2004. doi: 10.1175/1525-7541(2004)005⟨0745:POBSIA⟩2.0.CO;2.

[24] A. Allen Bradley and Stuart S. Schwartz. Summary verification measures and their interpretation for ensemble forecasts. *Monthly Weather Review*, 139(9):3075–3089, 2011. doi: 10.1175/2010MWR3305.1. URL https://doi.org/10.1175/2010MWR3305.1.

[25] A. Allen Bradley, Mohamed Habib, and Stuart S. Schwartz. Climate index weighting of ensemble streamflow forecasts using a simple bayesian approach. *Water Resources Research*, 51(9):7382–7400, 2015. ISSN 1944-7973. doi: 10.1002/2014WR016811. URL http://dx.doi.org/10.1002/2014WR016811.

[26] J. Breidenbach and J. Bradberry. Multisensor precipitation estimates produced by National Weather Service forecast centers for hydrologic applications. In *Proc., 2001 Georgia Water Resources Conf.*, pages 179–182, Athens, GA, 2001.

[27] J. Breidenbach, D.-J. Seo, P. Tilles, and K. Roy. Accounting for radar beam blockage patterns in radar-derived precipitation mosaics for River Forecast Centers. In *Proc., 15th Conf. on Interactive Information Processing Systems*, pages 179–182, Boston, MA, 1999.

[28] James D. Brown and Dong-Jun Seo. A nonparametric postprocessor for bias correction of hydrometeorological and hydrologic ensemble forecasts. *Journal of Hydrometeorology*, 11(3):642–665, 2010. doi: 10.1175/2009JHM1188.1. URL https://doi.org/10.1175/2009JHM1188.1.

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

[30] Patrick Broxton, Peter A. Troch, Mike Schaffner, Carl Unkrich, and David Goodrich. An all-season flash flood forecasting system for real-time operations. *Bulletin of the American Meteorological Society*, 95(3):399–407, 2014. doi: 10.1175/BAMS-D-12-00212.1.

[31] Roberto Buizza. Potential forecast skill of ensemble prediction and spread and skill distributions of the ECMWF ensemble prediction system. *Mon. Wea. Rev.*, 125:99–119, 1997.

[32] R.J. Burnash. *The NWS River Forecast System - Catchment Model*. Water Resources Publications, first edition, 1995.

[33] R.J. Burnash. *The NWS River Forecast System - Catchment Model*. Water Resources Publications, first edition, 1995.

[34] R.J. Burnash, R.L. Ferral, and R.A. McGuire. A generalized streamflow simulation system: Conceptual modeling for digital computers. Technical report, U.S. Department of Commerce National Weather Service and State of California Department of Water Resources, 1973.

[35] Guillem Candille. The multiensemble approach: The naefs example. *Monthly Weather Review*, 137(5):1655–1665, 2009. doi: 10.1175/2008MWR2682.1. URL https://doi.org/10.1175/2008MWR2682.1.

[36] Francesca Cecinati, Miguel Angel Rico-Ramirez, Gerard B.M. Heuvelink, and Dawei Han. Representing radar rainfall uncertainty with ensembles based on a time-variant geostatistical error modelling approach. *Journal of Hydrology*, 548(Supplement C):391 – 405, 2017. ISSN 0022-1694. doi: https://doi.org/10.1016/j.jhydrol.2017.02.053. URL http://www.sciencedirect.com/science/article/pii/S0022169417301324.

[37] Jerome P. Charba, David W. Reynolds, Brett E. McDonald, and Gary M. Carter. Comparative Verification of Recent Quantitative Precipitation Forecasts in the National Weather Service: A Simple Approach for Scoring Forecast Accuracy. *Weather and Forecasting*, 18(2):161–183, 2003.

[38] Grzegorz J. Ciach. Local random errors in tipping-bucket rain gauge measurements. *Journal of Atmospheric and Oceanic Technology*, 20(5):752–759, 2003. doi: 10.1175/1520-0426(2003)20⟨752:LREITB⟩2.0.CO;2.

[39] Grzegorz J. Ciach and Witold F. Krajewski. Radar-Rain Gauge Comparisons under Observational Uncertainties. *Journal of Applied Meteorology*, 38(10):1519–1525, 1999. doi: 10.1175/1520-0450(1999)038⟨1519:RRGCUO⟩2.0.CO;2.

[40] H.L. Cloke and F. Pappenberger. Ensemble flood forecasting: A review. *Journal of Hydrology*, 375(3):613 – 626, 2009. ISSN 0022-1694. doi: https://doi.org/10.1016/j.jhydrol.2009.06.005. URL http://www.sciencedirect.com/science/article/pii/S0022169409003291.

[41] P. Corrigan. Email June 29. personal communication, 2016. Weather Forecast Office - Blacksburg, VA, NOAA, National Weather Service.

[42] L. Crochemore, M.H. Ramos, F. Pappenberger, and C. Perrin. Seasonal streamflow forecasting by conditioning climatology with precipitation indices. *Hydrology and Earth System Sciences*, 21(3):1573–1591, 2017. doi: 10.5194/hess-21-1573-2017. URL https://www.hydrol-earth-syst-sci.net/21/1573/2017/.

[43] Timothy D. Crum and Ron L. Alberty. The wsr-88d and the wsr-88d operational support facility. *Bulletin of the American Meteorological Society*, 74(9):1669–1687, 1993. doi: 10.1175/1520-0477(1993)074⟨1669:TWATWO⟩2.0.CO;2.

[44] Luciana K. Cunha, James A. Smith, Mary Lynn Baeck, and Witold F. Krajewski. An early performance evaluation of the nexrad dual-polarization radar rainfall estimates for urban flood applications. *Weather and Forecasting*, 28(6):1478–1497, 2013. doi: 10.1175/WAF-D-13-00046.1.

[45] Lan Cuo, Thomas C. Pagano, and Q. J. Wang. A review of quantitative precipitation forecasts and their use in short- to medium-range streamflow forecasting. *Journal of Hydrometeorology*, 12(5):713–728, 2011. doi: 10.1175/2011JHM1347.1.

[46] C. Daly, R.P. Neilson, and D.L. Phillips. A Statistical-Topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteor.*, 33:140–158, 1994.

[47] Christopher Daly, Michael Halbleib, Joseph I. Smith, Wayne P. Gibson, Matthew K. Doggett, George H. Taylor, Jan Curtis, and Phillip P. Pasteris. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous united states. *International Journal of Climatology*, 28(15):2031–2064, 2008. ISSN 1097-0088. doi: 10.1002/joc.1688. URL http://dx.doi.org/10.1002/joc.1688.

[48] Christopher Daly, Joseph I. Smith, and Keith V. Olson. Mapping atmospheric moisture climatologies across the conterminous united states. *PLOS ONE*, 10(10):1–33, 10 2015. doi: 10.1371/journal.pone.0141140. URL https://doi.org/10.1371/journal.pone.0141140.

[49] U. Damrath, G. Doms, D. Fruehwald, E. Heise, B. Richter, and J. Steppeler. Operational quantitative precipitation forecasting at the german weather service. *J. Hydrology*, 239:260–285, 2000.

[50] D.R. Dawdy and J.M. Bergmann. Effect of rainfall variability on streamflow simulation. *Water Resour. Res.*, 5(5):140–158, 1969.

[51] G. N. Day. Extended streamflow forecasting using NWSRFS. *ASCE, J. Water Resour. Plann. Manage.*, 3:157–170, 1985.

[52] C. M. DeChant and H. Moradkhani. Improving the characterization of initial condition for ensemble streamflow prediction using data assimilation. *Hydrology and Earth System Sciences*, 15(11):3399–3410, 2011. doi: 10.5194/hess-15-3399-2011. URL https://www.hydrol-earth-syst-sci.net/15/3399/2011/.

[53] Caleb M. DeChant and Hamid Moradkhani. Toward a reliable prediction of seasonal forecast uncertainty: Addressing model and initial condition uncertainty with ensemble data assimilation and sequential bayesian combination. *Journal of Hydrology*, 519(Part D):2967 – 2977, 2014. ISSN 0022-1694. doi: https://doi.org/10.1016/j.jhydrol.2014.05.045. URL http://www.sciencedirect.com/science/article/pii/S0022169414004090.

[54] Caleb M. DeChant and Hamid Moradkhani. On the assessment of reliability in probabilistic hydrometeorological event forecasting. *Water Resources Research*, 51 (6):3867–3883, 2015. ISSN 1944-7973. doi: 10.1002/2014WR016617. URL http://dx.doi.org/10.1002/2014WR016617.

[55] Deltares. Flood Early Warning System (FEWS). Online: http://www.deltares.nl/en/software/479962/delft-fews, 3 2018.

[56] J. Demargne, M. Mulluski, K. Werner, T. Adams, S. Lindsey, N. Schwein, W. Marosi, and E. Welles. Application of forecast verification science to operational river forecasting in the u.s. national weather service. *Bulletin of the American Meteorological Society*, 90(6):779–784, 2009.

[57] J. Demargne, J.D. Brown, D-J. Seo, L. Wu, Z. Toth, and Y. Zhu. Diagnostic verification

of hydrometeorological and hydrologic ensembles. *Atmospheric Science Letters*, 11(2): 114–122, 2010.

[58] Julie Demargne, Limin Wu, Satish K. Regonda, James D. Brown, Haksu Lee, Minxue He, Dong-Jun Seo, Robert Hartman, Henry D. Herr, Mark Fresch, John Schaake, and Yuejian Zhu. The science of noaa's operational hydrologic ensemble forecast service. *Bulletin of the American Meteorological Society*, 95(1):79–98, 2014. doi: 10.1175/BAMS-D-12-00081.1. URL https://doi.org/10.1175/BAMS-D-12-00081.1.

[59] David Demeritt, Sbastien Nobert, Hannah Cloke, and Florian Pappenberger. Challenges in communicating and using ensembles in operational flood forecasting. *Meteorological Applications*, 17(2):209–222, 2010. doi: 10.1002/met.194. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/met.194.

[60] T. Diomede, F. Nerozzi, T. Paccagnella, and E. Todini. The use of meteorological analogues to account for lam qpf uncertainty. *Hydrology and Earth System Sciences*, 12(1):141–157, 2008. doi: 10.5194/hess-12-141-2008. URL https://www.hydrol-earth-syst-sci.net/12/141/2008/.

[61] Tommaso Diomede, Chiara Marsigli, Andrea Montani, Fabrizio Nerozzi, and Tiziana Paccagnella. Calibration of limited-area ensemble precipitation forecasts for hydrological predictions. *Monthly Weather Review*, 142(6):2176–2197, 2014. doi: 10.1175/MWR-D-13-00071.1.

[62] J. Du and M. S. Tracton. Implementation of a real-time short-range ensemble forecasting system at ncep: an update. In *Preprints, 9th Conference on Mesoscale Processes, Ft. Lauderdale,Florida*, pages 355–356. Amer. Meteor. Soc., 2001. URL http://www.emc.ncep.noaa.gov/mmb/SREF/reference.html.

[63] J. Du, G. DiMego, M. S. Tracton, and B. Zhou. NCEP short-range ensemble forecasting (SREF) system: multi-IC, multi-model and multi-physics approach. Technical Report TD No. 1161, Research Activities in Atmospheric and Oceanic Modeling, J. Cote (ed.), 2003.

[64] E. E. Ebert. Ability of a poor man's ensemble to predict the probability and distribution of precipitation. *Mon. Wea. Rev.*, 129:2461–2480, 2001.

[65] E. E. Ebert, U. Damrath, W. Wergen, and J.L. McBride. The WGNE assessment of short-term quantitative precipitation forecasts. *Bull. Am. Meteorol. Soc.*, 84(4): 481–492, 2003.

[66] Mohammad Ebtehaj, Hamid Moradkhani, and Hoshin V. Gupta. Improving robustness of hydrologic parameter estimation by the use of moving block bootstrap resampling. *Water Resources Research*, 46(7):n/a–n/a, 2010. ISSN 1944-7973. doi: 10.1029/2009WR007981. URL http://dx.doi.org/10.1029/2009WR007981. W07515.

[67] Hisham Eldardiry, Emad Habib, Yu Zhang, and Jeffrey Graschel. Artifacts in Stage IV NWS Real-Time Multisensor Precipitation Estimates and Impacts on Identification of Maximum Series. *Journal of Hydrologic Engineering*, 22(5):E4015003, 2017. doi: 10.1061/(ASCE)HE.1943-5584.0001291.

[68] John F. England, Pierre Y. Julien, and Mark L. Velleux. Physically-based extreme flood frequency with stochastic storm transposition and paleoflood data on large watersheds. *Journal of Hydrology*, 510(Supplement C):228 – 245, 2014. ISSN 0022-1694. doi: https://doi.org/10.1016/j.jhydrol.2013.12.021. URL http://www.sciencedirect.com/science/article/pii/S0022169413009189.

[69] Richard Essery and John Pomeroy. Vegetation and topographic control of wind-blown snow distributions in distributed and aggregated simulations for an arctic tundra basin. *Journal of Hydrometeorology*, 5(5):735–744, 2004. doi: 10.1175/1525-7541(2004)005⟨0735:VATCOW⟩2.0.CO;2.

[70] Richard Essery, Long Li, and John Pomeroy. A distributed model of blowing snow over complex terrain. *Hydrological Processes*, 13(14-15):2423–2438, 1999. ISSN 1099-1085. doi: 10.1002/(SICI)1099-1085(199910)13:14/15⟨2423::AID-HYP853⟩3.0.CO;2-U. URL http://dx.doi.org/10.1002/(SICI)1099-1085(199910)13:14/15<2423::AID-HYP853>3.0.CO;2-U.

[71] Gerald I. Evenden. Cartographic Projection Procedures for the UNIX Environment – A User's Manual. Technical Report Open-File Report 90-284, U.S. Dept. of the Interior, Geological Survey, 1990.

[72] J.-M. Faurès, D. C. Goodrich, D. A. Woolhiser, and S. Sorooshian. Impact of small-scale spatial rainfall variability on runoff modeling. *Journal of Hydrology*, 173:309–326, December 1995. doi: 10.1016/0022-1694(95)02704-S.

[73] B.D. Finnerty, M.B. Smith, D-J Seo, V. Koren, and G. Moglen. Space-time scale sensitivity of the sacramento model to radar-gage precipitation inputs. *Journal of Hydrology*, 203:21–38, 1997.

[74] M.M. Fogel. Effect of storm rainfall variability on runoff from small semiarid watersheds. *Transactions of the ASAE*, 12(6):808–812, 1969.

[75] G. Forzieri, L. Feyen, R. Rojas, M. Flörke, F. Wimmer, and A. Bianchi. Ensemble projections of future streamflow droughts in europe. *Hydrology and Earth System Sciences*, 18(1):85–108, 2014. doi: 10.5194/hess-18-85-2014. URL https://www.hydrol-earth-syst-sci.net/18/85/2014/.

[76] E. Foufoula-Georgiou. A probabilistic storm transposition approach for estimating exceedance probabilities of extreme precipitation depths. *Water Resources Research*, 25(5):799–815, 1989. ISSN 1944-7973. doi: 10.1029/WR025i005p00799. URL http://dx.doi.org/10.1029/WR025i005p00799.

[77] Kristie J. Franz, Holly C. Hartmann, Soroosh Sorooshian, and Roger Bales. Verification of National Weather Service ensemble streamflow predictions for water supply forecasting in the Colorado River Basin. *Journal of Hydrometeorology*, 4(12):1105–1118, 2003.

[78] R. A. Fulton. WSR-88D Polar-to-HRAP Mapping. Technical Report Tech. Memo., Hydrologic Research Laboratory, Office of Hydrology, National Weather Service, 1998.

[79] Richard A. Fulton, Jay P. Breidenbach, Dong-Jun Seo, Dennis A. Miller, and Timothy O'Bannon. The WSR-88D Rainfall Algorithm. *Weather and Forecasting*, 13(2):377–395, 1998. doi: 10.1175/1520-0434(1998)013⟨0377:TWRA⟩2.0.CO;2.

[80] Richard A. Fulton, Feng Ding, and Dennis A. Miller. Truncation errors in historical WSR-88D rainfall products. In *31st Conference on Radar Meteorology*. Amer. Meteor. Soc., 2003.

[81] Brian Gaudet and William R. Cotton. Statistical characteristics of a real-time precipitation forecasting model. *Wea. Forecasting*, 13:966–982, 1998.

[82] Konstantine P. Georgakakos and Michael D. Hudlow. Quantitative precipitation forecast techniques for use in hydrologic forecasting. *Bulletin of the American Meteorological Society*, 65(11):1186–1200, 1984. doi: 10.1175/1520-0477(1984)065⟨1186:QPFTFU⟩2.0.CO;2.

[83] Konstantine P. Georgakakos and Michael D. Hudlow. Quantitative precipitation forecast techniques for use in hydrologic forecasting. *Bull. Amer. Meteor. Soc.*, 65:1186–1200, 1984.

[84] David C. Goodrich, Jean-Marc Faurs, David A. Woolhiser, Leonard J. Lane, and Soroosh Sorooshian. Measurement and analysis of small-scale convective storm rainfall variability. *Journal of Hydrology*, 173(1?4):283 – 308, 1995. ISSN 0022-1694. doi: http://dx.doi.org/10.1016/0022-1694(95)02703-R. URL http://www.sciencedirect.com/science/article/pii/002216949502703R.

[85] Jonathan J. Gourley and Chris M. Calvert. Automated detection of the bright band using wsr-88d data. *Weather and Forecasting*, 18(4):585–599, 2003. doi: 10.1175/1520-0434(2003)018⟨0585:ADOTBB⟩2.0.CO;2.

[86] GRASS Development Team. *Geographic Resources Analysis Support System (GRASS GIS) Software*. Open Source Geospatial Foundation, USA, 2016. URL http://grass.osgeo.org.

[87] Thomas M. Hamill. Interpretation of rank histograms for verifying ensemble forecasts. *Monthly Weather Review*, 129(3):550–560, 2001. doi: 10.1175/1520-0493(2001)129⟨0550:IORHFV⟩2.0.CO;2. URL https://doi.org/10.1175/1520-0493(2001)129<0550:IORHFV>2.0.CO;2.

[88] Alan F. Hamlet, Daniel Huppert, and Dennis P. Lettenmaier. Economic value of long-lead streamflow forecasts for columbia river hydropower. *Journal of Water Resources Planning and Management*, 128(2):91–101, 2002. doi: 10.1061/(ASCE)0733-9496(2002)128:2(91). URL http://ascelibrary.org/doi/abs/10.1061/%28ASCE%290733-9496%282002%29128%3A2%2891%29.

[89] M.J. Hamlin. The significance of rainfall in the study of hydrological processes at basin scale. *Journal of Hydrology*, 65(1?3):73 – 94, 1983. ISSN 0022-1694. doi: http://dx.doi.org/10.1016/0022-1694(83)90211-1. URL http://www.sciencedirect.com/science/article/pii/0022169483902111. Scale Problems in Hydrology.

[90] Holly C. Hartmann, Thomas C. Pagano, S. Sorooshian, and R. Bales. Confidence builders. *Bulletin of the American Meteorological Society*, 83(5):683–698, 2002. doi: 10.1175/1520-0477(2002)083⟨0683:CBESCF⟩2.3.CO;2. URL https://doi.org/10.1175/1520-0477(2002)083<0683:CBESCF>2.3.CO;2.

[91] T. Hashino, A. A. Bradley, and S. S. Schwartz. Evaluation of bias-correction methods for ensemble streamflow volume forecasts. *Hydrology and Earth System Sciences*, 11 (2):939–950, 2007. doi: 10.5194/hess-11-939-2007. URL https://www.hydrol-earth-syst-sci.net/11/939/2007/.

[92] Yuxiang He, Yu Zhang, Robert Kuligowski, Robert Cifelli, and David Kitzmiller. Incorporating satellite precipitation estimates into a radar-gauge multi-sensor precipitation estimation algorithm. *Remote Sensing*, 10(1), 2018. ISSN 2072-4292. doi: 10.3390/rs10010106. URL http://www.mdpi.com/2072-4292/10/1/106.

[93] F.A. Huff. Sampling errors in measurement of mean precipitation. *J. Appl. Met.*, 9 (1):35–44, 1970.

[94] M. D. Humphrey, J. D. Istok, J. Y. Lee, J. A. Hevesi, and A. L. Flint. A new method for automated dynamic calibration of tipping-bucket rain gauges. *Journal of Atmospheric and Oceanic Technology*, 14(6):1513–1519, 1997. doi: 10.1175/1520-0426(1997)014⟨1513:ANMFAD⟩2.0.CO;2.

[95] Steven M. Hunter and Edmond W. Holroyd. Demonstration of improved operational water resources management through the use of better snow water equivalent information. Technical Report R-02-02, U.S. Dept. of the Interior, Bureau of Reclamation, Technical Service Center, River Systems and Meteorology Group, Water Resources Services Div., September 2002.

[96] Jung-Sun Im, Keith Brill, and Edwin Danaher. Confidence interval estimation for quantitative precipitation forecasts (qpf) using short-range ensemble forecasts (sref). *Weather and Forecasting*, 21(1):24–41, 2006. doi: 10.1175/WAF902.1.

[97] Christine Johnson and Neill Bowler. On the reliability and calibration of ensemble forecasts. *Monthly Weather Review*, 137(5):1717–1720, 2009. doi: 10.1175/2009MWR2715.1. URL https://doi.org/10.1175/2009MWR2715.1.

[98] D.M.A. Jones and W.M. Wendland. Some statistics of instantaneous precipitation. *Journal of Climate and Applied Meteorology*, 23(9):1273–1285, 1984.

[99] Susan Joslyn and Sonia Savelli. Communicating forecast uncertainty: public perception of weather forecast uncertainty. *Meteorological Applications*, 17(2):180–195, 2010. doi: 10.1002/met.190. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/met.190.

[100] Viatcheslav V. Kharin and Francis W. Zwiers. On the roc score of probability forecasts. *Journal of Climate*, 16(24):4145–4150, 2003. doi: 10.1175/1520-0442(2003)016⟨4145:OTRSOP⟩2.0.CO;2. URL https://doi.org/10.1175/1520-0442(2003)016<4145:OTRSOP>2.0.CO;2.

[101] G.S. Kirk and J.E. Raven. *The Presocratic Philosophers: A Critical History with a Selection of Texts*. University Press, 1957. URL https://books.google.com/books?id=2pzWAAAAMAAJ.

[102] David Kitzmiller, Suzanne Van Cooten, Feng Ding, Kenneth Howard, Carrie Langston, Jian Zhang, Heather Moser, Yu Zhang, Jonathan J. Gourley, Dongsoo Kim, and David Riley. Evolving multisensor precipitation estimation methods: Their impacts on flow prediction using a distributed hydrologic model. *Journal of Hydrometeorology*, 12(6): 1414–1431, 2011. doi: 10.1175/JHM-D-10-05038.1.

[103] David Kitzmiller, Dennis Miller, Richard Fulton, and Feng Ding. Radar and multisensor precipitation estimation techniques in national weather service hydrologic operations. *Journal of Hydrologic Engineering*, 18(2):133–142, 2013. doi: 10.1061/(ASCE)HE.1943-5584.0000523.

[104] V. I. Koren. Parameterization of frozen ground effects: sensitivity to soil properties. predictions in ungauged basins: Promises and progress. In *Symposium S7, Seventh IAHS Scientific Assembly*, pages 125–133, Foz do Iquacu, Brazil, 2006. IAHS Publication 303.

[105] V. I. Koren, M. Smith, D. Wang, and Z. Zhang. Use of soil property data in the derivation of conceptual rainfall-runoff model parameters. In *Conference on Hydrology*, Long Beach, CA, 2000. AMS.

[106] Victor Koren, Seann Reed, Michael Smith, Ziya Zhang, and Dong-Jun Seo. Hydrology laboratory research modeling system (hl-rms) of the us national weather service. *Journal of Hydrology*, 291(3):297 – 318, 2004. ISSN 0022-1694. doi: https://doi.org/10.1016/j.jhydrol.2003.12.039. URL http://www.sciencedirect.com/science/article/pii/S002216940400037X. Catchment modelling: Towards an improved representation of the hydrological processes in real-world model applications.

[107] Victor Koren, Michael Smith, Zhengtao Cui, Brian Cosgrove, Kevin Werner, and

Robert Zamora. Modification of Sacramento Soil Moisture Accounting Heat Transfer Component (SAC-HT) for Enhanced Evapotranspiration. Technical Report NOAA NWS Technical Report NWS 53, Department of Commerce, NOAA/NWS, October 2010.

[108] W.F. Krajewski and G.J. Ciach. Towards probabilistic quantitative precipitation wsr-88d algorithms: Preliminary studies and problem formulation. Technical report, Final Report for the Office of Hydrologic Development, NOAA/NWS, Washington, D.C., May 2003.

[109] Tobias Krueger, Jim Freer, John N. Quinton, Christopher J. A. Macleod, Gary S. Bilotta, Richard E. Brazier, Patricia Butler, and Philip M. Haygarth. Ensemble evaluation of hydrological model hypotheses. *Water Resources Research*, 46(7):n/a–n/a, 2010. ISSN 1944-7973. doi: 10.1029/2009WR007845. URL http://dx.doi.org/10.1029/2009WR007845. W07516.

[110] Roman Krzysztofowicz. Probabilistic hydrometeorological forecasts: Toward a new era in operational forecasting. *Bulletin of the American Meteorological Society*, 79(2):243–252, 1998. doi: 10.1175/1520-0477(1998)079⟨0243:PHFTAN⟩2.0.CO;2. URL https://doi.org/10.1175/1520-0477(1998)079<0243:PHFTAN>2.0.CO;2.

[111] Roman Krzysztofowicz. Bayesian theory of probabilistic forecasting via deterministic hydrologic model. *Water Resources Research*, 35(9):2739–2750, 1999. ISSN 1944-7973. doi: 10.1029/1999WR900099. URL http://dx.doi.org/10.1029/1999WR900099.

[112] Roman Krzysztofowicz. The case for probabilistic forecasting in hydrology. *Journal of Hydrology*, 249(1?4):2 – 9, 2001. ISSN 0022-1694. doi: http://dx.doi.org/10.1016/

S0022-1694(01)00420-6. URL http://www.sciencedirect.com/science/article/pii/S0022169401004206.

[113] C.L. Larson and B.M. Reich. Relationship of observed rainfall and runoff recurrence intervals. In *Proc. 2nd Int. Hydrol. Symp.*, volume 1, pages 34–43, Fort Collins, Colorado, 1972.

[114] J. Li, Y. Chen, H. Wang, J. Qin, J. Li, and S. Chiao. Extending flood forecasting lead time in a large watershed by coupling wrf qpf with a distributed hydrological model. *Hydrology and Earth System Sciences*, 21(2):1279–1294, 2017. doi: 10.5194/hess-21-1279-2017. URL https://www.hydrol-earth-syst-sci.net/21/1279/2017/.

[115] M. Li, Q. J. Wang, J. C. Bennett, and D. E. Robertson. Error reduction and representation in stages (erris) in hydrological modelling for ensemble streamflow forecasting. *Hydrology and Earth System Sciences*, 20(9):3561–3579, 2016. doi: 10.5194/hess-20-3561-2016. URL https://www.hydrol-earth-syst-sci.net/20/3561/2016/.

[116] Wentao Li, Qingyun Duan, Chiyuan Miao, Aizhong Ye, Wei Gong, and Zhenhua Di. A review on statistical postprocessing methods for hydrometeorological ensemble forecasting. *Wiley Interdisciplinary Reviews: Water*, pages e1246–n/a, 2017. ISSN 2049-1948. doi: 10.1002/wat2.1246. URL http://dx.doi.org/10.1002/wat2.1246. e1246.

[117] Y. Lin and K. E. Mitchell. The NCEP Stage II/IV hourly precipitation analyses: Development and applications. In *Pre-prints, 19th Conf. on Hydrology, 1.2. [Available online at https://ams.confex.com/ ams/pdfpapers/83847.pdf.]*, San Diego, CA, 2005. Amer. Meteor. Soc.

[118] Glen E. Liston and Kelly Elder. A distributed snow-evolution modeling system (snowmodel). *Journal of Hydrometeorology*, 7(6):1259–1276, 2006. doi: 10.1175/JHM548.1.

[119] Edwin P. Maurer and Dennis P. Lettenmaier. Predictability of seasonal runoff in the Mississippi River basin. *Journal of Geophysical Research: Atmospheres*, 108(D16):n/a–n/a, 2003. ISSN 2156-2202. doi: 10.1029/2002JD002555. URL http://dx.doi.org/10.1029/2002JD002555. 8607.

[120] J. McEnery, J. Ingram, Q. Duan, T. Adams, and L. Anderson. NOAAs Advanced Hydrologic Prediction Service: Building pathways for better science in water forecasting. *Bull. of the American Meteorological Society*, 24(3):375–385, 2005.

[121] Hilary McMillan, Bethanna Jackson, Martyn Clark, Dmitri Kavetski, and Ross Woods. Rainfall uncertainty in hydrological modelling: An evaluation of multiplicative error models. *Journal of Hydrology*, 400(1):83–94, 2011.

[122] P. A. Mendoza, A. W. Wood, E. Clark, E. Rothwell, M. P. Clark, B. Nijssen, L. D. Brekke, and J. R. Arnold. An intercomparison of approaches for improving operational seasonal streamflow forecasts. *Hydrology and Earth System Sciences*, 21(7):3915–3935, 2017. doi: 10.5194/hess-21-3915-2017. URL https://www.hydrol-earth-syst-sci.net/21/3915/2017/.

[123] Sarah Michaels. Probabilistic forecasting and the reshaping of food risk management. *Journal of Natural Resources Policy Research*, 7(1):41?51, 2015. doi: 10.1080/19390459.2014.970800.

[124] Marion Mittermaier and Nigel Roberts. Intercomparison of spatial forecast verification methods: Identifying skillful spatial scales using the fractions skill score. *Weather and Forecasting*, 25(1):343–354, 2010. doi: 10.1175/2009WAF2222260.1.

[125] Naoki Mizukami, Victor Koren, Michael Smith, David Kingsmill, Ziya Zhang, Brian Cosgrove, and Zhengtao Cui. The impact of precipitation type discrimination on hydrologic simulation: Rain?snow partitioning derived from hmt-west radar-detected brightband height versus surface temperature data. *Journal of Hydrometeorology*, 14 (4):1139–1158, 2013. doi: 10.1175/JHM-D-12-035.1.

[126] Nicolas Le Moine, Frederic Hendrickx, and Joel Gailhard. Rainfall-runoff modelling as a tool for constraining the reanalysis of daily precipitation and temperature fields in mountainous regions. In *Cold and Mountain Region Hydrological Systems Under Climate Change Towards Improved Projections 1 Proceedings of H02*, 2013.

[127] Hamid Moradkhani and Matthew Meier. Long-lead water supply forecast using large-scale climate predictors and independent component analysis. *Journal of Hydrologic Engineering*, 15(10):744–762, 2010. doi: 10.1061/(ASCE)HE.1943-5584.0000246. URL http://ascelibrary.org/doi/abs/10.1061/%28ASCE%29HE.1943-5584.0000246.

[128] Rebecca E. Morss, Julie L. Demuth, and Jeffrey K. Lazo. Communicating uncertainty in weather forecasts: A survey of the u.s. public. *Weather and Forecasting*, 23(5):974–991, 2008. doi: 10.1175/2008WAF2007088.1. URL https://doi.org/10.1175/2008WAF2007088.1.

[129] Allan H. Murphy. Probabilities, odds, and forecasts of rare events. *Weather and Forecasting*, 6(2):302–307, 1991. doi: 10.1175/1520-0434(1991)006⟨0302:POAFOR⟩2.0.CO;2. URL https://doi.org/10.1175/1520-0434(1991)006<0302:POAFOR>2.0.CO;2.

[130] Kenneth R. Mylne, Ruth E. Evans, and Robin T. Clark. Multi-model multi-analysis ensembles in quasi-operational medium-range forecasting. *Quarterly Jour-*

nal of the Royal Meteorological Society, 128(579):361–384, 2006. doi: 10.1256/00359000260498923. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1256/00359000260498923.

[131] J.E. Nash and J.V. Sutcliffe. River flow forecasting through conceptual models part i - a discussion of principles. *Journal of Hydrology*, 10(3):282–290, 1970. ISSN 0022-1694. doi: https://doi.org/10.1016/0022-1694(70)90255-6. URL http://www.sciencedirect.com/science/article/pii/0022169470902556.

[132] National Academies. *Integrating Social and Behavioral Sciences Within the Weather Enterprise*. The National Academies Press, Washington, DC, 2018. ISBN 978-0-309-46422-2. doi: 10.17226/24865. URL https://www.nap.edu/catalog/24865/integrating-social-and-behavioral-sciences-within-the-weather-enterprise.

[133] National Research Council. Completing the Forecast: Characterizing and Communicating Uncertainty for Better Decisions Using Weather and Climate Forecasts. Technical report, Committee on Estimating and Communicating Uncertainty in Weather and Climate Forecasts, Washington, D.C., 2006.

[134] National Weather Service. National Weather Service Strategic Plan 2011 – Building A Weather-Ready Nation. Online: https://www.weather.gov/media/wrn/strategic_plan.pdf, 3 2011.

[135] NCAR. *verification – Weather Forecast Verification Utilities*. NCAR, Research Applications Laboratory, 2015. URL https://CRAN.R-project.org/package=verification. R package version 1.42.

[136] Brian R. Nelson, Olivier P. Prat, D.-J. Seo, and Emad Habib. Assessment and implications of ncep stage iv quantitative precipitation estimates for product intercomparisons.

Weather and Forecasting, 31(2):371–394, 2016. doi: 10.1175/WAF-D-14-00112.1. URL https://doi.org/10.1175/WAF-D-14-00112.1.

[137] Andrew J. Newman, Martyn P. Clark, Jason Craig, Bart Nijssen, Andrew Wood, Ethan Gutmann, Naoki Mizukami, Levi Brekke, and Jeff R. Arnold. Gridded ensemble precipitation and temperature estimates for the contiguous united states. *Journal of Hydrometeorology*, 16(6):2481–2500, 2015. doi: 10.1175/JHM-D-15-0026.1.

[138] NOAA/NWS Performance Management. NOAA/NWS Verification. Online: https://verification.nws.noaa.gov/services/public/index.aspx, 3 2018. Password access reqired.

[139] David R. Novak, Christopher Bailey, Keith F. Brill, Patrick Burke, Wallace A. Hogsett, Robert Rausch, and Michael Schichtel. Precipitation and temperature forecast performance at the weather prediction center. *Weather and Forecasting*, 29(3):489–504, 2014. doi: 10.1175/WAF-D-13-00066.1.

[140] NRC. An Assessment of the Advanced Weather Interactive Processing system: Operational Test and Evaluation of the First System Build. Technical report, National Research Council, Washington, D.C., 1997.

[141] NRC. Toward a New Advanced Hydrologic Prediction Service (AHPS). Technical report, Committee to Assess the National Weather Service Advanced Hydrologic Prediction Service Initiative, Water Science and Technology Board, Washington, D.C., 2006.

[142] OHD. National Weather Service Verification Software Users' Manual. Technical report, NOAA/NWS, Office of Hydrologic Development, 1325 East West Highway, Room 8372 Silver Spring, Maryland 20910, 2000.

[143] Florian Pappenberger, Elisabeth Stephens, Jutta Thielen, Peter Salamon, David Demeritt, Schalk Jan Andel, Fredrik Wetterhall, and Lorenzo Alfieri. Visualizing probabilistic flood forecast information: expert preferences and perceptions of best practice in uncertainty communication. *Hydrological Processes*, 27(1):132–146, 2012. doi: 10.1002/hyp.9253. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.9253.

[144] Florian Pappenberger, Elisabeth Stephens, Jutta Thielen, Peter Salamon, David Demeritt, Schalk Jan van Andel, Fredrik Wetterhall, and Lorenzo Alfieri. Visualizing probabilistic flood forecast information: expert preferences and perceptions of best practice in uncertainty communication. *Hydrological Processes*, 27(1):132–146, 2013. ISSN 1099-1085. doi: 10.1002/hyp.9253. URL http://dx.doi.org/10.1002/hyp.9253.

[145] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2017. URL https://www.R-project.org/.

[146] David Raff, Levi Brekke, Kevin Werner, Andy Wood, and Kathleen White. Short-term water management decisions: User needs for improved climate, weather, and hydrologic information. Technical report, U.S. Army Corps of Engineers, Bureau of Reclamation, National Oceanic and Atmospheric Administration, Washington, D.C., January 2013.

[147] M. H. Ramos, S. J. van Andel, and F. Pappenberger. Do probabilistic forecasts lead to better decisions? *Hydrology and Earth System Sciences*, 17(6):2219–2232, 2013. doi: 10.5194/hess-17-2219-2013. URL https://www.hydrol-earth-syst-sci.net/17/2219/2013/.

[148] R. Rausch. Email July 28. personal communication, 2016. Weather Prediction Center, NOAA, National Weather Service.

[149] Steve Rayner, Denise Lach, and Helen Ingram. Weather forecasts are for wimps: Why water resource managers do not use climate forecasts. *Climatic Change*, 69(2): 197–227, Apr 2005. ISSN 1573-1480. doi: 10.1007/s10584-005-3148-z. URL https://doi.org/10.1007/s10584-005-3148-z.

[150] Seann M. Reed and David R. Maidment. A GIS procedure for merging NEXRAD precipitation data and digital elevation models to determine rainfall-runoff modeling parameters. Technical report, University of Texas, 1995. URL http://www.ce.utexas.edu/prof/maidment/gishyd97/library/nexrad/rep95_3.htm.

[151] Seann M. Reed and David R. Maidment. Coordinate transformations for using nexrad data in gis-based hydrologic modeling. *Journal of Hydrologic Engineering*, 4(2):174–182, 1999. doi: 10.1061/(ASCE)1084-0699(1999)4:2(174).

[152] Brian M. Reich. Flood series compared to rainfall extremes. *Water Resources Research*, 6(6):1655–1667, 1970. ISSN 1944-7973. doi: 10.1029/WR006i006p01655. URL http://dx.doi.org/10.1029/WR006i006p01655.

[153] Daniela Rezacova, Petr Zacharov, and Zbynek Sokol. Uncertainty in the area-related qpf for heavy convective precipitation. *Atmospheric Research*, 93(1):238 – 246, 2009. ISSN 0169-8095. doi: https://doi.org/10.1016/j.atmosres.2008.12.005. URL http://www.sciencedirect.com/science/article/pii/S0169809508003475. 4th European Conference on Severe Storms.

[154] O. Rössler, P. Froidevaux, U. Börst, R. Rickli, O. Martius, and R. Weingartner. Retrospective analysis of a nonforecasted rain-on-snow flood in the alps: a mat-

ter of model limitations or unpredictable nature? *Hydrology and Earth System Sciences*, 18(6):2265–2285, 2014. doi: 10.5194/hess-18-2265-2014. URL https://www.hydrol-earth-syst-sci.net/18/2265/2014/.

[155] Peter Salamon and Luc Feyen. Disentangling uncertainties in distributed hydrological modeling using multiplicative error models and sequential data assimilation. *Water Resources Research*, 46(12):n/a–n/a, 2010. ISSN 1944-7973. doi: 10.1029/2009WR009022. URL http://dx.doi.org/10.1029/2009WR009022. W12501.

[156] J. Schaake, A. Henkel, and S. Cong. Application of PRISM climatologies for hydrologic modeling and forecasting in the western U.S. In *Proc., 18th Conf. on Hydrology*, 2004.

[157] J. Schaake, T. M. Hamill, and R. Buizza. Hepex: The hydrological ensemble prediction experiment. *Bull. Amer. Meteor. Soc.*, 88:1541–1547, 2007.

[158] D.-J. Seo. Real-time estimation of rainfall fields using radar rainfall and rain gage data. *Journal of Hydrology*, 208(1):37 – 52, 1998. ISSN 0022-1694. doi: https://doi.org/10.1016/S0022-1694(98)00141-3. URL http://www.sciencedirect.com/science/article/pii/S0022169498001413.

[159] D.-J Seo, J.P Breidenbach, and E.R Johnson. Real-time estimation of mean field bias in radar rainfall data. *Journal of Hydrology*, 223(3):131 – 147, 1999. ISSN 0022-1694. doi: https://doi.org/10.1016/S0022-1694(99)00106-7. URL http://www.sciencedirect.com/science/article/pii/S0022169499001067.

[160] D.-J. Seo, J. Breidenbach, R. Fulton, and D. Miller. Real-time adjustment of range-dependent biases in WSR-88D rainfall estimates due to nonuniform vertical profile of reflectivity. *J. Hydrometeorology*, 1:222–240, 2000.

[161] S. Sharma, R. Siddique, S. Reed, P. Ahnert, P. Mendoza, and A. Mejia. Relative

effects of statistical preprocessing and postprocessing on a regional hydrological ensemble prediction system. *Hydrology and Earth System Sciences Discussions*, 2017:1–30, 2017. doi: 10.5194/hess-2017-514. URL https://www.hydrol-earth-syst-sci-discuss.net/hess-2017-514/.

[162] R. Siddique and A. Mejia. Ensemble streamflow forecasting across the U.S. Middle Atlantic region with a distributed hydrological model forced by GEFS reforecasts. *J. Hydrometeor*, 2017. doi: doi:10.1175/JHM-D-16-0243.1.

[163] James A. Smith, Mary Lynn Baeck, Matthias Steiner, and Andrew J. Miller. Catastrophic rainfall from an upslope thunderstorm in the central appalachians: The rapidan storm of june 27, 1995. *Water Resources Research*, 32(10):3099–3113, 1996. ISSN 1944-7973. doi: 10.1029/96WR02107. URL http://dx.doi.org/10.1029/96WR02107.

[164] James A. Smith, Mary Lynn Baeck, Julia E. Morrison, and Paula Sturdevant-Rees. Catastrophic rainfall and flooding in texas. *Journal of Hydrometeorology*, 1(1):5–25, 2000. doi: 10.1175/1525-7541(2000)001⟨0005:CRAFIT⟩2.0.CO;2.

[165] M. Smith, V. Koren, B. Finnerty, and D. Johnson. Distributed Modeling: Phase 1 Results. Technical report, Department of Commerce, NOAA Technical Report NWS 44, Washington, D.C., 1999.

[166] Michael B. Smith, Donald P. Laurine, Victor I. Koren, Seann M. Reed, and Ziya Zhang. *Hydrologic Model Calibration in the National Weather Service*, pages 133–152. American Geophysical Union, 2013. ISBN 9781118665671. doi: 10.1029/WS006p0133. URL http://dx.doi.org/10.1029/WS006p0133.

[167] Zbyněk Sokol. Mos-based precipitation forecasts for river basins. *Weather and Forecasting*, 18(5):769–781, 2003. doi: 10.1175/1520-0434(2003)018⟨0769:MPFFRB⟩2.0.CO;2.

[168] D. J. Stensrud, H. E. Brooks, J. Du, M. S. Tracton, and E. Rogers. Using ensembles for short-range forecasting. *Mon. Wea. Rev.*, 127:433–446, 1999.

[169] Paul C. Stern and William E. Easterling, editors. *Making Climate Forecasts Matter*. The National Academies Press, Washington, DC, 1999. ISBN 978-0-309-06475-0. doi: 10.17226/6370. URL https://www.nap.edu/catalog/6370/making-climate-forecasts-matter.

[170] O. Sungmin, U. Foelsche, G. Kirchengast, and J. Fuchsberger. Validation and correction of rainfall data from the WegenerNet high density network in southeast Austria. *Journal of Hydrology*, 2016. ISSN 0022-1694. doi: https://doi.org/10.1016/j.jhydrol.2016.11.049. URL http://www.sciencedirect.com/science/article/pii/S0022169416307648.

[171] G.H. Taylor, C. Daly, and W.P. Gibson. Development of an isohyetal analysis for Oregon using the PRISM model. In *8th Conf. on Applied Climatology*, pages 126–127, Anaheim, CA, 1993.

[172] G.H. Taylor, C. Daly, and W.P. Gibson. Development of a model for use in estimating the spatial distribution of precipitation. In *9th Conf. on Applied Climatology*, pages 92–93, Dallas, TX, 1995.

[173] D. Tetzlaff and S. Uhlenbrook. Significance of spatial variability in precipitation for process-oriented modelling: results from two nested catchments using radar and ground station data. *Hydrology and Earth System Sciences*, 9(1/2):29–41, 2005. doi: 10.5194/hess-9-29-2005. URL https://www.hydrol-earth-syst-sci.net/9/29/2005/.

[174] A. Thiboult, F. Anctil, and M.-A. Boucher. Accounting for three sources of uncertainty in ensemble hydrological forecasting. *Hydrology and Earth System Sciences*, 20(5):

1809–1825, 2016. doi: 10.5194/hess-20-1809-2016. URL https://www.hydrol-earth-syst-sci.net/20/1809/2016/.

[175] S. W. D. Turner, J. C. Bennett, D. E. Robertson, and S. Galelli. Complex relationship between seasonal streamflow forecast skill and value in reservoir operations. *Hydrology and Earth System Sciences*, 21(9):4841–4859, 2017. doi: 10.5194/hess-21-4841-2017. URL https://www.hydrol-earth-syst-sci.net/21/4841/2017/.

[176] T. M. Twedt, J. C. Schaake Jr., and E. L. Peck. National Weather Service extended streamflow prediction. In *Proceedings Western Snow Conference*, pages 52–57, 1977.

[177] U.S. Department of Commerce. National Weather Service River Forecast System (NWSRFS-Model). Technical report, NOAA Technical Memorandum NWS-Hydro-14, Washington, D.C., 1972.

[178] M. S. Wandishin, S. L. Mullen, D. J. Stensrud, and H.E. Brooks. Evaluation of short-range multimodel ensemble system. *Mon. Wea. Rev.*, 129:729–747, 2001.

[179] Nicholas E. Wayand. *Observation, Simulation, and Evaluation of Snow Dynamics in the Transitional Snow Zone*. PhD dissertation, University of Washington, 2016.

[180] Nicholas E. Wayand, Martyn P. Clark, and Jessica D. Lundquist. Diagnosing snow accumulation errors in a rain-snow transitional environment with snow board observations. *Hydrological Processes*, 31(2):349–363, 2017. ISSN 1099-1085. doi: 10.1002/hyp.11002. URL http://dx.doi.org/10.1002/hyp.11002. HYP-16-0027.R2.

[181] Edwin Welles, Soroosh Sorooshian, Gary Carter, and Billy Olsen. Hydrologic verification: A call for action and collaboration. *Bull. of the American Meteorological Society*, 88:503–511, 2007.

[182] Li Wentao, Duan Qingyun, Miao Chiyuan, Ye Aizhong, Gong Wei, and Di Zhenhua. A review on statistical postprocessing methods for hydrometeorological ensemble forecasting. *Wiley Interdisciplinary Reviews: Water*, 4(6):e1246, 2017. doi: 10.1002/wat2.1246. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/wat2.1246.

[183] Kevin Werner, David Brandon, Martyn Clark, and Subhrendu Gangopadhyay. Incorporating medium-range numerical weather model output into the ensemble streamflow prediction system of the national weather service. *Journal of Hydrometeorology*, 6 (2):101–114, 2005. doi: 10.1175/JHM411.1. URL https://doi.org/10.1175/JHM411.1.

[184] F. Wetterhall, F. Pappenberger, L. Alfieri, H. L. Cloke, J. Thielen-del Pozo, S. Balabanova, J. Daňhelka, A. Vogelbacher, P. Salamon, I. Carrasco, A. J. Cabrera-Tordera, M. Corzo-Toscano, M. Garcia-Padilla, R. J. Garcia-Sanchez, C. Ardilouze, S. Jurela, B. Terek, A. Csik, J. Casey, G. Stanknavičius, V. Ceres, E. Sprokkereef, J. Stam, E. Anghel, D. Vladikovic, C. Alionte Eklund, N. Hjerdt, H. Djerv, F. Holmberg, J. Nilsson, K. Nyström, M. Sušnik, M. Hazlinger, and M. Holubecka. Hess opinions "forecaster priorities for improving probabilistic flood forecasts". *Hydrology and Earth System Sciences*, 17(11):4389–4399, 2013. doi: 10.5194/hess-17-4389-2013. URL https://www.hydrol-earth-syst-sci.net/17/4389/2013/.

[185] D.S. Wilks. *Statistical Methods in the Atmospheric Sciences*. Academic Press, second edition, 2006.

[186] Charles B. Wilson, Juan B. Valdes, and Ignacio Rodriguez-Iturbe. On the influence of the spatial distribution of rainfall on storm runoff. *Water Resources Research*, 15

(2):321–328, 1979. ISSN 1944-7973. doi: 10.1029/WR015i002p00321. URL http://dx.doi.org/10.1029/WR015i002p00321.

[187] Adam Winstral, Kelly Elder, and Robert E. Davis. Spatial snow modeling of wind-redistributed snow using terrain-based parameters. *Journal of Hydrometeorology*, 3(5): 524–538, 2002. doi: 10.1175/1525-7541(2002)003⟨0524:SSMOWR⟩2.0.CO;2.

[188] Andrew W. Wood and Dennis P. Lettenmaier. A test bed for new seasonal hydrologic forecasting approaches in the western united states. *Bulletin of the American Meteorological Society*, 87(12):1699–1712, 2006. doi: 10.1175/BAMS-87-12-1699. URL https://doi.org/10.1175/BAMS-87-12-1699.

[189] Andrew W. Wood and Dennis P. Lettenmaier. An ensemble approach for attribution of hydrologic prediction uncertainty. *Geophysical Research Letters*, 35(14):n/a–n/a, 2008. ISSN 1944-8007. doi: 10.1029/2008GL034648. URL http://dx.doi.org/10.1029/2008GL034648. L14401.

[190] Andrew W. Wood and John C. Schaake. Correcting errors in streamflow forecast ensemble mean and spread. *Journal of Hydrometeorology*, 9(1):132–148, 2008. doi: 10.1175/2007JHM862.1. URL https://doi.org/10.1175/2007JHM862.1.

[191] World Meteorological Organization. WMO Statement on the Scientific Basis for, and Limitations of, River Discharge and Stage Forecasting. Online: http://www.wmo.int/pages/prog/hwrp/publications/statements/stmnt_limitations08042010.pdf, 3 2018.

[192] Daniel B. Wright, James A. Smith, and Mary Lynn Baeck. Flood frequency analysis using radar rainfall fields and stochastic storm transposition. *Water Resources Research*, 50(2):1592–1615, 2014. ISSN 1944-7973. doi: 10.1002/2013WR014224. URL http://dx.doi.org/10.1002/2013WR014224.

[193] Youlong Xia, Kenneth Mitchell, Michael Ek, Brian Cosgrove, Justin Sheffield, Lifeng Luo, Charles Alonge, Helin Wei, Jesse Meng, Ben Livneh, Qingyun Duan, and Dag Lohmann. Continental-scale water and energy flux analysis and validation for north american land data assimilation system project phase 2 (nldas-2): 2. validation of model-simulated streamflow. *Journal of Geophysical Research: Atmospheres*, 117(D3):n/a–n/a, 2012. ISSN 2156-2202. doi: 10.1029/2011JD016051. URL http://dx.doi.org/10.1029/2011JD016051. D03110.

[194] Youlong Xia, Kenneth Mitchell, Michael Ek, Justin Sheffield, Brian Cosgrove, Eric Wood, Lifeng Luo, Charles Alonge, Helin Wei, Jesse Meng, Ben Livneh, Dennis Lettenmaier, Victor Koren, Qingyun Duan, Kingtse Mo, Yun Fan, and David Mocko. Continental-scale water and energy flux analysis and validation for the north american land data assimilation system project phase 2 (nldas-2): 1. intercomparison and application of model products. *Journal of Geophysical Research: Atmospheres*, 117(D3):n/a–n/a, 2012. ISSN 2156-2202. doi: 10.1029/2011JD016048. URL http://dx.doi.org/10.1029/2011JD016048. D03109.

[195] Jingbing Xiao, Richard Bintanja, Stephen J. Déry, Graham W. Mann, and Peter A. Taylor. An Intercomparison Among Four Models Of Blowing Snow. *Boundary-Layer Meteorology*, 97(1):109–135, 2000. URL https://doi.org/10.1023/A:1002795531073.

[196] C. Bryan Young, A. Allen Bradley, Witold F. Krajewski, Anton Kruger, and Mark L. Morrissey. Evaluating nexrad multisensor precipitation estimates for operational hydrologic forecasting. *Journal of Hydrometeorology*, 1(3):241–254, 2000. doi: 10.1175/1525-7541(2000)001⟨0241:ENMPEF⟩2.0.CO;2.

[197] Xing Yuan and Eric F. Wood. Downscaling precipitation or bias?correcting stream-

flow? some implications for coupled general circulation model (cgcm)?based ensemble seasonal hydrologic forecast. *Water Resources Research*, 48(12), 2012. doi: 10.1029/2012WR012256. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012WR012256.

[198] I. Zalachori, M.-H. Ramos, R. Garçon, T. Mathevet, and J. Gailhard. Statistical processing of forecasts for hydrological ensemble prediction: a comparative study of different bias correction strategies. *Advances in Science and Research*, 8(1):135–141, 2012. doi: 10.5194/asr-8-135-2012. URL https://www.adv-sci-res.net/8/135/2012/.

[199] Mauricio Zambrano-Bigiarini. *hydroGOF: Goodness-of-fit functions for comparison of simulated and observed hydrological time series*, 2014. URL https://CRAN.R-project.org/package=hydroGOF. R package version 0.3-8.

[200] Jian Zhang, Kenneth Howard, Carrie Langston, Brian Kaney, Youcun Qi, Lin Tang, Heather Grams, Yadong Wang, Stephen Cocks, Steven Martinaitis, Ami Arthur, Karen Cooper, Jeff Brogden, and David Kitzmiller. Multi-Radar Multi-Sensor (MRMS) Quantitative Precipitation Estimation: Initial Operating Capabilities. *Bulletin of the American Meteorological Society*, 97(4):621–638, 2016. doi: 10.1175/BAMS-D-14-00174.1.

[201] Yu Zhang, Thomas Adams, and James V. Bonta. Subpixel-Scale Rainfall Variability and the Effects on Separation of Radar and Gauge Rainfall Errors. *Journal of Hydrometeorology*, 8(6):1348–1363, 2007. doi: 10.1175/2007JHM835.1.

[202] Yu Zhang, Seann Reed, and David Kitzmiller. Effects of retrospective gauge-based readjustment of multisensor precipitation estimates on hydrologic simulations. *Jour-*

nal of Hydrometeorology, 12(3):429–443, 2011. doi: 10.1175/2010JHM1200.1. URL https://doi.org/10.1175/2010JHM1200.1.

[203] L. Zhao, Q. Duan, J. Schaake, A. Ye, and J. Xia. A hydrologic post-processor for ensemble streamflow predictions. *Advances in Geosciences*, 29:51–59, 2011. doi: 10.5194/adgeo-29-51-2011. URL https://www.adv-geosci.net/29/51/2011/.

Appendices

Appendix A

Data Sources

Data sources used in Chapters 3 and 4 are listed in Table A.1. Most data sources required re-formatting (covered in Appendix B) and spatial data needed to be re-projected to a common map projection and vertical and horizontal datums for:

1. precipitation bias analyses (Chapter 3); and
2. use in the NOAA/NWS RDHM for the hydrologic *monte carlo* experiments discussed in Chapter 3.

Chapter 5 utilized USGS river stage observation data and operational forecast data archived in the OHRFC PostgreSQL *archive database*, *vfypairs* and *pecrcep* tables, used by verification software to compare OHRFC operational forecast errors to MMEFS NAEFS ensemble median and mean forecast errors. MMEFS simulation output ensemble time-series data were written to custom binary files by the NWSRFS *Ensemble Streamflow Prediction (ESP)* system software, described by Day [51]. These data are available by request from the author.

Explanations of the individual datasets are given in the Chapters listed in Table A.1. The

Table A.1: Data sources.

Name	Type	Format	Source	Chapter	Contact
PRISM	precipitation	BIL	http://prism.oregonstate.edu	3	PRISM Climate Group, Oregon State Univ
MPE	precipitation	GRIB1	https://water.weather.gov/precip/data_availability.php	3	NOAA/NWS
MPE	precipitation	xmrg	https://www.weather.gov/abrfc/	3	NOAA/NWS Arkansas-Red Basin RFC
QPF	verification statistics	ascii	http://www.wpc.ncep.noaa.gov/qpf/qpf2.shtml	3	NOAA/NWS Weather Prediction Center (available upon request: robert.rausch@noaa.gov)
QPF	verification statistics	ascii	http://origin.wpc.ncep.noaa.gov/npvu_test/index.shtml	3	NOAA/NWS Weather Prediction Center, NPVU
Observation	stage	ascii	OHRFC *Archive database pecrcep* table	4	NOAA/NWS Ohio RFC (available upon request: james.noel@noaa.gov)
Experiments 1 & 2 simulations	stage	ascii	OHRFC *Archive database vfypairs* table	4	NOAA/NWS Ohio RFC (available upon request: james.noel@noaa.gov)
Observation - Alderson, WV (03183500)	stage	ascii	https://waterdata.usgs.gov/usa/nwis/uv?03183500	4	USGS
Observation	temperature	GRIB2	https://disc.sci.gsfc.nasa.gov/datasets?keywords=NLDAS	4	NASA GES DISC
Observation	precipitation	GRIB2	https://disc.sci.gsfc.nasa.gov/datasets?keywords=NLDAS	4	NASA GES DISC

NASA NLDAS dataset used, listed in Table A.1, was the NLDAS Secondary Forcing Data, identified as L4 Hourly, 0.125 x 0.125 degree V002 (NLDAS_FORB0125_H.002), covering 1979-01-01 to 2018-04-23, which includes both precipitation and temperature grids, as well as many additional hydrometeorological variables.

Appendix B

Data Analyses

B.1 Methodology

B.1.1 Chapter 3 analyses

Chapter 3 utilizes *GRASS GIS* [86] for analyses consisting of:

1. bias calculation (Equation 3.1), comparison of OHRFC MPE estimates to PRISM estimates on a monthly, seasonal, and annual basis;

2. spatial transposition of 1000 randomly selected geographical locations used as storm centers, using observed precipitation for the 24-h period, spanning the June 22-24, 2016 period;

3. *threat score (TS)* calculation (Equation 3.14) for the 24-h rainfall accumulation using the 1000 transposed storms;

4. run RDHM hydrologic *monte carlo* simulations for 88 transposed storms with TS$\geq$0.06,

including transposing each of the hourly precipitation grids for the 88 storms, and reformatting each file to xmrg format for use in the RDHM;

5. pair RDHM peak flow value, TS, and storm transposed distance from the reference location for the 88 simulated events;

6. the paired data is imported into the *R Language and Environment for Statistical Computing* [145] for graphical analysis.

B.1.2 Chapter 4 analyses

Chapter 4 includes two independent operational experiments. Experiment-1 compares, by lead-time, ME, MAE, and RMSE values for forecasts using WPC 24-h QPF with forecasts using zero-QPF. Experiment-2 uses WPC QPF for six QPF durations (6-, 12-, 24-, 36-, 48-,72-h) and the OHRFC NWSRFS modeling system for experimental simulations. Air temperature forcings are the same for all simulations.. Analyses consist simply of calculating ME, MAE, and RMSE for the six WPC QPF scenarios, by lead-time, from the simulated data stored in the OHRFC archive database using the *R Language and Environment for Statistical Computing* [145] and the contributed *verification* package [135].

B.1.3 Chapter 5 analyses

Analyses utilize OHRFC operational and MMEFS NAEFS ensemble river stage forecasts from 54 forecast point locations, representing different geographical location, watershed areas, and basin response times. Chapter 5 analyses include:

1. deterministic forecast verification using the *R Language and Environment for Statistical Computing* [145] and the contributed *verification* package [135] from forecast data

stored in the OHRFC PostgreSQL *archive database*, *vfypairs* table;

2. forecast verification of the MMEFS NAEFS ensemble river stage forecasts using the NOAA/NWS EVS software [29] and USGS observed river stage data stored in the OHRFC PostgreSQL *archive database*, *pecrcep* table. Ensemble mean and median river stage forecast ME, MAE, RMSE, and CRPSS values, by lead-time, are also generated for each forecast point location;

3. data re-formatting of the MMEFS NAEFS ensemble mean and median river stage forecast verification statistics from the EVS XML data format for import into R (for CRPSS: Listing B.22; for ME, MAE, RMSE: Listing B.23);

4. forecast comparisons and graphics generation using R, following data import.

B.2 Codes

The complete set of Perl, Bash shell, and R scripts used in the dissertation for the various experiments and analyses, discussed in Chapter B, Section B.1, are listed in the code Listings.

Listings

```
lid,response
ALDW2,1
ATHO1,2
BEAP1,1
BEDI3,2
BRKI3,1
BUCW2,1
CARI2,3
CDIO1,2
CLKW2,1
CNTI3,1
COLO1,2
CYCK2,1
DLYW2,1
ELRP1,1
ELZW2,2
EVVI3,3
FDYO1,1
FFTK2,2
FLRK2,2
FRAT1,1
FRKP1,2
FTWI3,2
GOLI2,3
GRTW2,1
HAMO1,1
HUFI3,3
INDI3,1
INDO1,1
```

```
JKNK2,1
KANW2,2
KILO1,1
LAFI3,2
MILO1,1
MLGO1,1
NATP1,2
NHRI3,3
NWBI3,3
OLNN6,1
OLPO1,1
PARP1,2
PKTO1,2
PSNW2,1
PTRI3,3
PTTP1,2
SERI3,2
SHLI3,2
SIDO1,1
SLMN6,1
SPNI3,2
STRO1,2
SWDP1,1
WILW2,1
WLBK2,2
WTVO1,3
```

Listing B.1: Listing of forecast point locations in the *paper3_exp_pts.txt* file from Chapter 5.

```
\o jknk2.obs.txt
select lid,obstime,z0000,z0600,z1200,z1800 from pecrsep where pe1='H' and pe2='G' and t='R' and s
    ='G' and lid='JKNK2' order by obstime asc;

/home/teaiii/scripts/perl/pecrsepObs2TS.pl jknk2.obs.txt
jknk2.obs.txt.out

/home/teaiii/scripts/perl/ts2datacard.pl jknk2.obs.txt.out --id=JKNK2
JKNK2.OBS.20010326-20120524.datacard
```

Listing B.2: Example PostgreSQL query from *pecrcep* table for river stage observations for use in MMEFS EVS verification software used in Chapter 5.

```
/home/teaiii/scripts/perl/mmefsXML2R.pl /home/teaiii/phd/paper3_exp_pts.txt

mmefs<-read.table("/home/teaiii/phd/mmefs_stats.txt",sep="\t", header=T)

mmefs_mean<-read.table("/home/teaiii/phd/mmefs_mean_stats.txt",sep="\t", header=T)
mmefs_median<-read.table("/home/teaiii/phd/mmefs_median_stats.txt",sep="\t", header=T)
```

Listing B.3: Example code to read EVS XML code from MMEFS ensemble forecast experiments used in Chapter 5.

```
#--------------------------------
# Define: summarySE
#--------------------------------
summarySE <- function(data=NULL, measurevar, groupvars=NULL, na.rm=FALSE,
                      conf.interval=.95, .drop=TRUE) {
```

```
    library(plyr)

    # New version of length which can handle NA's: if na.rm==T, don't count them
    length2 <- function (x, na.rm=FALSE) {
        if (na.rm) sum(!is.na(x))
        else       length(x)
    }

    # This does the summary. For each group's data frame, return a vector with
    # N, mean, and sd
    datac <- ddply(data, groupvars, .drop=.drop,
      .fun = function(xx, col) {
        c(N    = length2(xx[[col]], na.rm=na.rm),
          mean = mean   (xx[[col]], na.rm=na.rm),
          sd   = sd     (xx[[col]], na.rm=na.rm)
        )
      },
      measurevar
    )

    # Rename the "mean" column
    datac <- rename(datac, c("mean" = measurevar))

    datac$se <- datac$sd / sqrt(datac$N)  # Calculate standard error of the mean

    # Confidence interval multiplier for standard error
    # Calculate t-statistic for confidence interval:
    # e.g., if conf.interval is .95, use .975 (above/below), and use df=N-1
    ciMult <- qt(conf.interval/2 + .5, datac$N-1)
    datac$ci <- datac$se * ciMult

    return(datac)
}
```

Listing B.4: Definition of R *summarySE* function.

```
##############################################################################
# Stats 04/01/2018
##############################################################################

library(verification)
library(RPostgreSQL)
library(reshape)
library(reshape2)
library(ggplot2)
library("dplyr")

leadtime<-c(6,12,18,24,30,36,42,48,54,60,66,72,78,84,90,96,102,108,114,120,126,\
132,138,144,150,156,162,168)
numLeads<-length(leadtime)

resp_time<-read.table("/media/teaiii/development1/grass/paper2_ex2_pts_resp_clean.txt",sep=',',
    header=T)
numLIDs<-nrow(resp_time)

fcstType<-c("G","H","I","J","K","L")
numfcstType<-length(fcstType)

floodCat<-c("above","below")
numfloodCat<-length(floodCat)

#Create a new dataframe:
```

```
stats<-data.frame(lid=character(),type=character(),response=character(),cat=character(),leadtime=
    numeric(),me=numeric(),mae=numeric(),rmse=numeric())
names(stats) <- c("lid","type","response","cat","leadtime","me","mae","rmse")

drv<-dbDriver("PostgreSQL")

for (i in 1:numLIDs) {

datStratified<-data.frame(lid=character(),type=character(),response=character(),cat=character(),
    basistime=character(),fcstvalue=numeric(),obsvalue=numeric(),leadtime=numeric())
names(datStratified) <- c("lid","type","response","cat","basistime","fcstvalue","obsvalue","
    leadtime")

  lid<-as.character(resp_time$lid[i])
  resTime<-resp_time$response[i]

  #### Get Flood Stage, fs
  sql_fs<-paste(c("select fs from ohrfc_fs where id=","'",lid,"'"),collapse="")
  con<-dbConnect(drv,dbname="teaiii",user="teaiii")
  fs_df<-dbGetQuery(con,sql_fs)
  dbDisconnect(con)
  fs<-fs_df$fs

  for (j in 1:numfcstType) {

    typeStr<-paste(c("X",fcstType[j]),collapse="")

    #### Get data for lid,typeStr
    sql<-paste(c("select lid,basistime,validtime,obstime,fcstvalue,obsvalue,(EXTRACT(epoch FROM (
    SELECT (validtime-basistime)))/3600)::int as leadtime from vfypairs where lid=","'",lid,"'","
     and fcst_t='X' and fcst_s=","'",fcstType[j],"'"," order by lid,basistime,validtime asc"),
    collapse="")

    con<-dbConnect(drv,dbname="teaiii",user="teaiii")
    dat<-dbGetQuery(con,sql)
    dbDisconnect(con)

    ###################
    # Above FS
    ###################
    aboveFS<-subset(dat, fcstvalue >= fs)
    exceed_basis_times<-as.character(unique(aboveFS$basistime))
    numAboveFS<-length(exceed_basis_times)
    if (numAboveFS > 0) {
      for (m in 1:numAboveFS) {

        #### Get data for lid,typeStr,basistimes with a forecast value > fs
        sql<-paste(c("select lid,basistime,validtime,obstime,fcstvalue,obsvalue,(EXTRACT(epoch
    FROM (SELECT (validtime-basistime)))/3600)::int as leadtime from vfypairs where lid=","'",lid
    ,"'"," and fcst_t='X' and fcst_s=","'",fcstType[j],"'"," and basistime=","'",exceed_basis_
    times[m],"'"," order by lid,basistime,validtime asc"),collapse="")

        con<-dbConnect(drv,dbname="teaiii",user="teaiii")
        df<-dbGetQuery(con,sql)
        dbDisconnect(con)

        temp_df<-data.frame(lid,typeStr,resTime,"above",df$basistime,df$fcstvalue,df$obsvalue,df$
    leadtime)
        names(temp_df) <- c("lid","type","response","cat","basistime","fcstvalue","obsvalue","
    leadtime")
        datStratified <- rbind(datStratified, temp_df)
        #Kluge to handle TimeZone Changes
        datStratified$leadtime<-(as.integer((datStratified$leadtime)/6))*6
        rm(temp_df)
```

```
      }
    }

    ###################
    # Below FS
    ###################
    belowFS<-subset(dat, fcstvalue < fs)
    below_basis_times<-as.character(unique(belowFS$basistime))
    numBelowFS<-length(below_basis_times)
    for (n in 1:numBelowFS) {

      #### Get data for lid,typeStr,basistimes with a forecast value < fs
      sql<-paste(c("select lid,basistime,validtime,obstime,fcstvalue,obsvalue,(EXTRACT(epoch FROM
     (SELECT (validtime-basistime)))/3600)::int as leadtime from vfypairs where lid=","'",lid,"'"
    ," and fcst_t='X' and fcst_s=","'",fcstType[j],"'"," and basistime=","'",below_basis_times[n
    ],"'"," order by lid,basistime,validtime asc"),collapse="")

      con<-dbConnect(drv,dbname="teaiii",user="teaiii")
      df<-dbGetQuery(con,sql)
      dbDisconnect(con)

      temp_df<-data.frame(lid,typeStr,resTime,"below",df$basistime,df$fcstvalue,df$obsvalue,df$
    leadtime)
      names(temp_df) <- c("lid","type","response","cat","basistime","fcstvalue","obsvalue","
    leadtime")
      datStratified <- rbind(datStratified, temp_df)
      #Kluge to handle TimeZone Changes
      datStratified$leadtime<-(as.integer((datStratified$leadtime)/6))*6
      rm(temp_df)
    }
    #---------------------------------
    #LOOP over leadtimes:
    #---------------------------------

    for (k in 1:numLeads) {

      cat_incr<-k+(i-1)*numLeads

      newdata <- subset(datStratified, leadtime == leadtime[[k]],select=c(fcstvalue,obsvalue))
      A<- verify(newdata$obsvalue,newdata$fcstvalue,frcst.type = "cont", obs.type = "cont")

      temp_df<-data.frame(lid,typeStr,resTime,datStratified$cat[cat_incr],datStratified$leadtime[
    k],A$ME,A$MAE,sqrt(A$MSE))
      names(temp_df) <- c("lid","type","response","cat","leadtime","me","mae","rmse")
      stats <- rbind(stats, temp_df)

    }
  }
}
```

Listing B.5: R code for Experiment-2 Chapter 3.

```
#================================================
# Code 2018-03-29
#================================================
library(verification)
library(RPostgreSQL)
library(reshape)
library(reshape2)
library(ggplot2)

# basins
```

```
resp_time<-read.table("/home/teaiii/phd/paper3_exp_pts.txt",sep=',',header=T)
numLIDs<-nrow(resp_time)

leadtime<-c(6,12,18,24,30,36,42,48,54,60,66,72,78,84,90,96,102,108,114,120)
numLeads<-length(leadtime)

#---------------------------------
# Define: my.HoursDiff
#---------------------------------
my.HoursDiff <- function(x,y) {
    z<-as.numeric(d<-as.difftime(c(as.character(x,format = "%Y-%m-%d %H:%M:%S"),as.character(y,
    format = "%Y-%m-%d %H:%M:%S")),format = "%Y-%m-%d %H:%M:%S",units = "hours"))
    d<-as.numeric(z[1]-z[2])
    return(as.integer(d))
}

#---------------------------------
# Define: my.bastimeToSynoptic
#---------------------------------
my.bastimeToSynoptic <- function(x) {
  f<-unlist(strsplit(as.character(x), " "))
  hr<-unlist(strsplit(f[2], ":"))
  synoptic<-as.integer(as.numeric(hr[1])/6)*6
  tdate<-paste(c(f[1]," ",as.character(synoptic),":00:00"),collapse="")
  d<-as.POSIXct(tdate,format = "%Y-%m-%d %H:%M:%S", tz="EST")
    return(d)
}

#Create a new dataframe:
stats<-data.frame(lid=character(),type=character(),response=character(),leadtime=numeric(),me=
    numeric(),mae=numeric(),rmse=numeric())
names(stats) <- c("lid","type","response","leadtime","me","mae","rmse")

for (i in 1:numLIDs) {

  lid<-as.character(resp_time$lid[i])
  resTime<-resp_time$response[i]

  drv<-dbDriver("PostgreSQL")
  con<-dbConnect(drv,dbname="teaiii",user="teaiii")

  sql_ff<-paste(c("select lid,basistime,validtime,obstime,fcstvalue,obsvalue from vfypairs where
    lid=","'",lid,"'"," and fcst_t='F' and fcst_s='F' and basistime>='2010-11-29 12:00:00' and
    basistime<='2012-05-25 12:00:00' order by lid,basistime,validtime asc"),collapse="")

  dat_ff<-dbGetQuery(con,sql_ff)
  dbDisconnect(con)

dat_ff$basistime<-as.POSIXct(apply(dat_ff[,2,drop=F], MARGIN=1, my.bastimeToSynoptic),origin = "
    1970-01-01", tz = "EST")
dat_ff$leadtime<-mapply(my.HoursDiff, dat_ff$validtime,dat_ff$basistime)

  #---------------------------------
  #LOOP over leadtimes:
  #---------------------------------
  for (j in 1:numLeads) {

    newdata <- subset(dat_ff, leadtime == leadtime[[j]],select=c(fcstvalue,obsvalue))
    A<- verify(newdata$obsvalue,newdata$fcstvalue,frcst.type = "cont", obs.type = "cont")

    temp_df<-data.frame(lid,"FF",resTime,dat_ff$leadtime[j],A$ME,A$MAE,sqrt(A$MSE))
    names(temp_df) <- c("lid","type","response","leadtime","me","mae","rmse")
    stats <- rbind(stats, temp_df)

```

```
  }

}
```

Listing B.6: R code for the Experiment in Chapter 5 to query PostgreSQL *vfypairs* table and calculate ME MAE and RMSE for each of the 54 forecast point locations for OHRFC operational forecasts.

```
a#!/bin/bash

################################################################
# Program: run_extractMany
#
# Written by: Thomas Adams
# Date:   03/30/2018
# Updated:  03/30/2018
#
################################################################
#
# Run from directory: /media/teaiii/Seagate8/naefs
#
#
# Script location:  /home/teaiii/scripts/run_extractMany
#
# File location:  /home/teaiii/phd/add_evs_locs.txt
#
#
#
#
#
#
################################################################

list_PATH=$1  # PATH to file 'pt_cat_list' with GRASS vector point cats

# For each lid in the file
for lid in $(<$list_PATH/add_evs_locs.txt);
  do
    echo "processing ${lid}..."
    /home/teaiii/scripts/extractMany $lid
  done

exit
```

Listing B.7: run_extractMany Bash shell script code used to extract NWSRFS ESP binary CS file time-series data for MMEFS verification in the EVS.

```
#!/bin/bash

id=$1

#Move many files from the current directory to a different directory
#and echo the names of the files being copied; strip off '.tmp' extension.

dir_path_str=/media/teaiii/Seagate8/naefs/cs/${id}
if [ ! -d $dir_path_str ]; then

      mkdir -p $dir_path_str;
else
```

```
13
14   rm -R $dir_path_str/*
15
16 fi;
17
18 for file in $(find . -name "*.gz") ; do
19
20 tar -xzf $file ./${id}.${id}.SSTG.06.CS
21
22 base=$(basename $file .CS.tar.gz)
23 date_str=$(echo $file | cut -c20-29)
24
25 echo "$date_str"
26
27 mv ./${id}.${id}.SSTG.06.CS $dir_path_str/${id}.${date_str}.SSTG.06.CS
28 done
```

Listing B.8: extractMany Bash shell script called by the run_extractMany Bash shell script

```
1  #!/bin/bash
2
3  ################################################################
4  # Program: run_bil2grass
5  #
6  # Written by: Thomas Adams
7  # Date:   01/30/2018
8  # Updated:  01/30/2018
9  #
10 ################################################################
11
12 suffix=".bil"
13
14 currDir=$(pwd)
15
16 for file in $(find . -name "*.bil") ; do
17
18         #echo $file
19   fname=$(basename $file)
20         #echo $fname
21   map_name=${fname%$suffix}
22   dirname=${map_name}_bil
23         #echo $dirname
24   path=$currDir/$dirname
25
26   echo "input = $path/$fname"
27
28   /home/teaiii/scripts/bil2grass $path $fname $map_name
29
30 done
31
32 exit
```

Listing B.9: The run_bil2grass bash shell script is used to import PRISM precipitation data into GRASS GIS for MPE bias analyses.

```
1  #!/bin/bash
2
3  ################################################################
4  # Program: bil2grass
5  #
```

```
# Written by: Thomas Adams
# Date:   01/30/2018
# Updated:  01/30/2018
#
#
# /home/teaiii/.grass7/rc
# MAPSET: teaiii
# GISDBASE: /media/teaiii/development1/grass
# LOCATION_NAME: world_location
# GUI: wxpython
#
#
################################################################

################################################################
# Set GRASS Environment Variables
################################################################
echo "GISDBASE: /media/teaiii/development1/grass"  > $HOME/.grassrc6
echo "LOCATION_NAME: world_location"      >> $HOME/.grassrc6
echo "MAPSET: teaiii"         >> $HOME/.grassrc6
echo "GRASS_GUI: wxpython"        >> $HOME/.grassrc6
#
export USER=$USER
export GISBASE=/usr/local/grass-7.2.1
export GISDBASE=/media/teaiii/development1/grass
export MAPSET=teaiii
export LOCATION_NAME=world_location
export GISRC=$HOME/.grassrc6
export PATH=$GISBASE/bin:$PATH:$GISBASE/scripts
export LD_LIBRARY_PATH=$LD_LIBRARY_PATH:$GISBASE/lib
#
################################################################

dirname=$1
fname=$2
map_name=$3

echo "input = $dirname/$fname"

echo
echo "=========  Begin processing of bil file $fname ==========="
echo

r.in.gdal -o input=$dirname/$fname output=$map_name memory=300 offset=0 --overwrite

echo
echo "=========  Step 1: Completed importing bil file into GRASS Location ==========="
echo

echo
echo "==========  Clean-up GRASS DB $GISDBASE & LOCATION $LOCATION_NAME  ==========="
echo

# Cleanup -- Remove temporary GRASS files
$GISBASE/etc/clean_temp

exit
```

Listing B.10: bil2grass bash shell script is called by run_bil2grass.

```

#!/bin/bash
```

```
3
4  ################################################################
5  # *************           run_nldas2xmrg              *************
6  ################################################################
7  #
8  # Filename example: NLDAS_FORA0125_H.A20110101.1800.002.grb
9  #
10 # Example PATH: /media/teaiii/development1/NLDAS_grids/grib/2011/001
11 #
12 ################################################################
13 ################################################################
14
15 path=$1 # Without /2011/001 from Example PATH, above
16 year=$2
17
18 cd $path
19
20 if [ $((year % 4)) -ne 0 ] ; then
21   #  not a leap year
22   days=365
23 elif [ $((year % 400)) -eq 0 ] ; then
24   # is a leap year
25   days=366
26 elif [ $((year % 100)) -eq 0 ] ; then
27   # not a leap year
28   days=365
29 else
30   # is a leap year
31   days=366
32 fi
33
34 prefix="./NLDAS_FORA0125_H.A"
35 string=${file#$prefix}
36
37 for i in `seq 1 $days`;
38   do
39
40   if [ $i -lt 10 ] ; then
41     day_str=00${i}
42   elif [ $i -lt 100 ] ; then
43     day_str=0${i}
44   else
45     day_str=$i
46   fi
47
48   echo "Day of year: $day_str"
49
50   filePATH=${path}/${year}/${day_str}
51   cd $filePATH
52   for file in $(find . -name "*.grb") ; do
53
54     string=${file#$prefix}
55
56     year=$(echo $string | cut -c1-4)
57     month=$(echo $string | cut -c5-6)
58     day=$(echo $string | cut -c7-8)
59     hour=$(echo $string | cut -c10-11)
60
61     date_time=${year}${month}${day}${hour}
62     echo "Processing... $date_time" #Processing... 20121228.1000.002.grb
63
64     wgrib -s $file  | grep ":APCP:" | wgrib -i -grib $file -o NLDAS.apcp.grb
65     wgrib -s $file  | grep ":TMP:" | wgrib -i -grib $file -o NLDAS.tmp.grb
66
```

```
    /home/teaiii/scripts/nldas2xmrg $filePATH $date_time

    rm NLDAS.apcp.grb
    rm NLDAS.tmp.grb
  done
done
```

Listing B.11: run_nldas2xmrg bash shell script is used to import NASA NLDAS precipitation and temperature grids into GRASS GIS.

```
#!/bin/bash

###############################################################
# *************           nldas2xmrg              *************

###############################################################
###############################################################

data_path=$1
date_time=$2

GISDBASE=/media/teaiii/development1/grass
MAPSET=$USER
LOCATION_NAME=NLDAS2

echo "LOCATION_NAME: $LOCATION_NAME"   > $HOME/.grass7/rc_grib2xmrg
echo "MAPSET: $MAPSET"      >> $HOME/.grass7/rc_grib2xmrg
echo "DIGITIZER: none"      >> $HOME/.grass7/rc_grib2xmrg
echo "GISDBASE: $GISDBASE"    >> $HOME/.grass7/rc_grib2xmrg
echo "GRASS_GUI: text"      >> $HOME/.grass7/rc_grib2xmrg

export GISBASE=/usr/local/grass-7.2.0
export GISRC=$HOME/.grass7/rc_grib2xmrg
export PATH=$PATH:$GISBASE/bin:$PATH:$GISBASE/scripts
#
export GISBASE=/usr/local/grass-7.2.0
export GISDBASE=/awips/hydroapps/grass/data
export MAPSET=$USER
export LOCATION_NAME=NLDAS2
export GISRC=$HOME/.grass7/rc_grib2xmrg
export PATH=$PATH:$GISBASE/bin:$GISBASE/scripts
export LD_LIBRARY_PATH=$LD_LIBRARY_PATH:$GISBASE/lib

###############################################################
# Script does the following:
#
# (1) extracts all surface precipitation fields matching ":APCP:"
#   (2) extracts all surface temperature fields matching ":TMP:"
#   (3) passed 'day_of_year' value is actually day_of_year-1
#
#
###############################################################
#
# Files read are:
#
# NLDAS.apcp.grb
# NLDAS.tmp.grb
#
# from: $data_path
#
# XMRG-like temperature files (in degrees F) have file name formats: tair1231201015z.gz
#
```

```
# north:      -3705225    N: -3424237.5  S:   -7620000
# south:      -7620000    E:  3433762.5  W:   -1905000
# west:       -1905000
# east:       3148013
#
# g.region -dp e=3433762.5 w=-1905000 n=-3424237.5 s=-7620000 res=4762.5
#
#############################################################

echo "PATH:  $data_path..."
echo "Year and Day:  $date_time..."

today=$(date +%Y%m%d)

#if [ $day_of_year -eq 0 ] ; then
# fileDate=${year}0101
#else
# fileDate=$(date -d "$day_of_year days $year-01-01" +"%Y%m%d")
#fi
year=$(echo $date_time | cut -c1-4)
month=$(echo $date_time | cut -c5-6)
day=$(echo $date_time | cut -c7-8)
hour=$(echo $date_time | cut -c9-10)

# Output file date/time
output_DateTime_str=${month}${day}${year}${hour}

echo "Input File date:   $date_time..."
echo "Output File date:  $output_DateTime_str..."

#=========================================================
echo "Step-1 import GRIB grib files into GRASS NLDAS2 LOCATION"
#=========================================================

echo
echo "=========  Begin processing of grib files from $data_path ==========="
echo

cd $data_path

echo
echo "=========  Importing APCP and TMP grib files into GRASS ==========="
echo

echo "r.in.gdal input=NLDAS.apcp.grb  output=NLDAS.apcp"
r.in.gdal input=NLDAS.apcp.grb  output=NLDAS.apcp --overwrite

echo "r.in.gdal input=NLDAS.tmp.grb  output=NLDAS.tmp"
r.in.gdal input=NLDAS.tmp.grb  output=NLDAS.tmp --overwrite

#=========================================================
echo "Step-2 Re-project to Stereographic projection; export Arc ASCII grid; convert to xmrg"
#=========================================================

MAPSET=$USER
GISDBASE=/media/teaiii/development1/grass
LOCATION_NAME=newster

echo "LOCATION_NAME: $LOCATION_NAME"   > $HOME/.grass7/rc_nldas
echo "GISDBASE: $GISDBASE"    >> $HOME/.grass7/rc_nldas
echo "MAPSET: $MAPSET"      >> $HOME/.grass7/rc_nldas
echo "GUI: text"      >> $HOME/.grass7/rc_nldas

export GISRC=$HOME/.grass7/rc_nldas
```

```
export LOCATION_NAME=newster

# MARFC Region
#g.region -dp e=3048000 w=2185987.5 n=-4381500 s=-5419725 res=4762.5
# CONUS Region
g.region -dp e=3433762.5 w=-1905000 n=-3424237.5 s=-7620000 res=4762.5

r.proj input=NLDAS.apcp location=NLDAS2 output=NLDAS.apcp mapset=$MAPSET dbase=$GISDBASE --
    overwrite
r.out.gdal --overwrite input=NLDAS.apcp output=precip_grid.asc format=AAIGrid nodata=-9999
/usr/local/bin/asctoxmrg -i precip_grid.asc -o xmrg${output_DateTime_str}z -p ster
mv xmrg${output_DateTime_str}z.gz /media/teaiii/Seagate8/nldas_conus_xmrg/precip/

g.remove -f type=raster name=temp_grid
r.proj input=NLDAS.tmp location=NLDAS2 output=tmp_grd mapset=$MAPSET dbase=$GISDBASE --overwrite
r.mapcalc expression="temp_grid=9*tmp_grd/5 + 32" --overwrite
r.out.gdal --overwrite input=temp_grid output=temp_grid.asc format=AAIGrid nodata=-9999
/usr/local/bin/asctoxmrg -i temp_grid.asc -o tair${output_DateTime_str}z -p ster
mv tair${output_DateTime_str}z.gz /media/teaiii/Seagate8/nldas_conus_xmrg/temps/

echo
echo "================ Cleanup follows ================"
echo

# Cleanup -- Remove temporary GRASS files
$GISBASE/etc/clean_temp

# remove session tmp directory:
rm -rf /tmp/grass7-$USER-$GIS_LOCK

rm temp_grid.prj temp_grid.asc temp_grid.asc.aux.xml NLDAS.apcp.grb NLDAS.tmp.grb precip_grid.asc

echo
echo "================ Cleanup completed ================"
echo
```

Listing B.12: nldas2xmrg bash shell script is called by the run_nldas2xmrg bash shell script.

```
#!/bin/bash

############################################################
# *************       run_QPF_TS_experiment       *************

############################################################
# Set initial GRASS Environment Variables
############################################################
# Program: run_QPF_TS_experiment
#
# Written by: Thomas Adams
# Date:   09/15/2017
# Updated:  10/17/2017
#
# Program uses r.in.arc to import a text file containing XMRG files
#
# /home/teaiii/.grass7/rc
# MAPSET: adams
# GISDBASE: /media/teaiii/development1/grass
# LOCATION_NAME: newster
# GUI: wxpython
#
# Default Raster Map Bounds (CONUS)
# north:      -3705225
```

```
# south:      -7620000
# west:       -1905000
# east:       3148013
#
# For map = xmrg2016062406_24hr_ge2
# Number of map grid cells >= 2 inches (50.8 mm): 4317
#
# NOTE: Imported xmrg data has units of mm
#
############################################################
#export GIS_LOCK=$$

############################################################
# Set GRASS Environment Variables
############################################################
#echo "GISDBASE: /home/teaiii/grass/data"   > $HOME/.grassrc6
echo "GISDBASE: /media/teaiii/development1/grass"  > $HOME/.grassrc6
echo "LOCATION_NAME: ster"        >> $HOME/.grassrc6
echo "MAPSET: test_mapset"        >> $HOME/.grassrc6
echo "GRASS_GUI: wxpython"        >> $HOME/.grassrc6
#
export USER=$USER
export GISBASE=/usr/local/grass-7.2.1
#export GISDBASE=
export GISDBASE=/media/teaiii/development1/grass
export MAPSET=test_mapset
export LOCATION_NAME=ster
export GISRC=$HOME/.grassrc6
export PATH=$GISBASE/bin:$PATH:$GISBASE/scripts
export LD_LIBRARY_PATH=$LD_LIBRARY_PATH:$GISBASE/lib
#
############################################################

loc=1000
outputFilename="QPF_TS_vals.dat"

for i in `seq 1 $loc`;
  do
    echo "Processing Random Point $i ..."

    loc=$(v.db.select map=greenbrier_exp_random_pts@teaiii columns=x where=cat=$i)
    a_loc=($loc)
    export x_loc=$(echo ${a_loc[1]})

    loc=$(v.db.select map=greenbrier_exp_random_pts@teaiii columns=y where=cat=$i)
    a_loc=($loc)
    export y_loc=$(echo ${a_loc[1]})

    # Raster map xmrg2016062406_24hr was manually derived
    g.copy --overwrite raster=xmrg2016062406_24hr,xmrg_temp

    #r.region map=xmrg_temp n=-3779868 s=-7694643 e=3057837 w=-1995176

    perl -e '
      use POSIX;
      #              OFFSET     REF CENTER      SAMPLE POINT
      #  x -- ADD: -90176.17 = 2416752.01754  - 2326575.84362
      #  y -- ADD: -74643.28 = -5293585.88792 - (-5368229.16897)

      $x = $ENV{"x_loc"};
      $y = $ENV{"y_loc"};

      $x_offset = sprintf("%.0f",2416752.01754 - $x);
      $y_offset = sprintf("%.0f",-5293585.88792 - $y);
```

```
89
90      #  north:      -3779868 = -3705225 -74643
91      #  south:      -7694643 = -7620000 -74643
92      #  east:       3057837  = 3148013  -90176
93      #  west:       -1995176 = -1905000 -90176
94
95      $north = -3705225 - $y_offset;
96      $south = -7620000 - $y_offset;
97      $east  = 3148013  - $x_offset;
98      $west  = -1905000 - $x_offset;
99
100     $cmd = "r.region map=xmrg_temp n=".$north." s=".$south." e=".$east." w=".$west;
101     printf("Executing %s...\n", $cmd );
102     system($cmd);
103   '
104
105   # Raster map xmrg2016062406_24hr_ge2 was manually derived...
106   #xmrg2016062406_24hr_ge2=if(xmrg2016062406_24hr>=50.8,1,null())
107
108   r.mapcalc expression="xmrg_temp_ge2=if(xmrg_temp>=50.8,1,null())" --overwrite
109   #-------------------------------------------------------------------------
110   # hits are NOT calculated for null() cells, so only cells for the INTERSECTION
111   # of the two maps will show values of '2' -- we only need the number of such
112   # cells
113   #-------------------------------------------------------------------------
114   r.mapcalc expression="hits=xmrg_temp_ge2+xmrg2016062406_24hr_ge2" --overwrite
115
116   # Find the number of HITS (Overlapping map grid cells)
117   items=$(r.univar -g map=hits | grep "n=")
118   arr=($items)
119   echo ${arr[0]}
120   num=$(echo ${arr[0]} | cut -c3-)
121   echo $num
122
123   if ((i >= 1)) ; then
124
125     if [ $i -ge 1000 ] ; then
126       istr="Pt"${i}
127
128     elif [ $i -ge 100 ] ; then
129       istr="Pt0"${i}
130     elif [ $i -ge 10 ] ; then
131       istr="Pt00"${i}
132     else
133       istr="Pt000"${i}
134     fi
135
136     if ((num >= 1)) ; then
137       echo "${istr} $num" >> $HOME/$outputFilename
138     else
139       echo "${istr} 0" >> $HOME/$outputFilename
140     fi
141
142   else
143     if ((num >= 1)) ; then
144       echo "Pt0001 $num"  > $HOME/$outputFilename
145     else
146       echo "Pt0001  0" >> $HOME/$outputFilename
147     fi
148
149     istr="Pt0001"
150   fi
151
152   newMap="xmrg_"${istr}
```

```
153     newMapGE2="xmrg_ge2"${istr}
154     g.copy --overwrite raster=xmrg_temp,$newMap
155     g.copy --overwrite raster=xmrg_temp_ge2,$newMapGE2
156     r.to.vect --overwrite input=$newMapGE2 output=$newMapGE2 type=area
157
158   done
```

Listing B.13: run_QPF_TS_experiment bash shell script calculates *threat score* values from the RDHM *monte carlo* hydrologic simulations for the Chapter 5 experiment.

```
1  #!/bin/bash
2
3  ################################################################
4  # *************       renameCONUS2xmrg            *************
5  #
6  # rxmrg_1hr_1231200711.gz to xmrg1231200711.gz
7  #
8  ################################################################
9  ################################################################
10
11 #year=$1
12
13 prefix="./rxmrg_1hr_"
14
15 for file in $(find . -name "rxmrg_1hr_*") ; do
16
17 #echo "${file}"
18
19 string=${file#$prefix}
20 newName=xmrg${string}
21
22 #echo "${newName}"
23 body=$(echo $newName | cut -c1-14)
24 #echo "${body}"
25
26 mv $file ${body}z.gz
27
28 #exit
29 done
```

Listing B.14: renameCONUS2xmrg shell script is used to rename CONUS scale MPE *xmrg* format file used in RDHM simulations.

```
1  #!/bin/bash
2
3  ################################################################
4  # Program: setupRDHMwarm_states
5  #
6  # Written by: Thomas Adams
7  # Date:   11/16/2015
8  # Updated:  11/16/2015
9  #
10 # Requires:
11 #
12 # (1) Location of RDHM model states
13 # (2) Start Date-time -- YYYYMMDDHH format
14 # (3) number of hours from $start_date_time
15 #
16 ################################################################
```

```
17
18 dir_path=$1 # Location of RDHM model states
19 cd $dir_path
20 echo $dir_path
21
22 start_date_time=$2  # format is YYYYMMDDHH
23 periods=$3    # number of hours from $start_date_time
24
25 year=$(echo $start_date_time | cut -c1-4)
26 month=$(echo $start_date_time | cut -c5-6)
27 day=$(echo $start_date_time | cut -c7-8)
28 hour=$(echo $start_date_time | cut -c9-10)
29
30 warm_states_dir_path=/media/teaiii/development1/RDHM/warm_states
31
32 cd $dir_path
33 for dir_name in $(ls -d */);
34   do
35     trimmed_dir_name=$(echo $dir_name | sed 's:/*$::')
36     echo $trimmed_dir_name
37
38     dir_path_str=$warm_states_dir_path/$trimmed_dir_name
39
40     if [ ! -d $dir_path_str ]; then
41
42           mkdir -p $dir_path_str;
43     else
44
45       rm -R $dir_path_str/*
46
47     fi;
48
49     # filename format: discharge0904201106z.gz --
50
51     for i in `seq 1 $periods`;
52       do
53         next_date=$(date +"%m%d%Y%H" --date "$year-$month-$day 00:00:00 $i hours" -u)
54         filename=${trimmed_dir_name}${next_date}z.gz
55                     echo $filename
56
57         state_file=${dir_path}/${trimmed_dir_name}/${filename}
58         if [ -f "$state_file" ]; then
59           cp $state_file $dir_path_str
60         fi
61
62       done
63   done
```

Listing B.15: setupRDHMwarm_states is a shell script used to create a directory structure and move RDHM model initialization files for the RDHM *monte carlo* simulation experiments.

```
1 #!/bin/bash
2
3 ############################################################
4 # *************      grassTranslateXMRG        *************
5
6 ############################################################
7 # Set initial GRASS Environment Variables
8 ############################################################
9 # Program: grassTranslateXMRG
```

```
#
# Written by: Thomas Adams
# Date:   08/09/2017
# Updated:  08/09/2017
#
# Program uses r.in.arc to import a text file containing XMRG files
#
# /home/teaiii/.grass7/rc
# MAPSET: adams
# GISDBASE: /media/teaiii/development1/grass
# LOCATION_NAME: ster
# GUI: wxpython
#
# IMPORTANT!
#
# The Map reference point and bounds MUST be changed to
# match the problem at hand
#
# Default Raster Map Bounds (CONUS)
# north:      -3705225
# south:      -7620000
# west:       -1905000
# east:       3148013
#
# For map = xmrg2016062406_24hr_ge2
# Number of map grid cells >= 2 inches (50.8 mm): 4317
#
# NOTE: Imported xmrg data has units of mm
#
################################################################

################################################################
# Set GRASS Environment Variables
################################################################
#echo "GISDBASE: /home/teaiii/grass/data"   > $HOME/.grassrc6
echo "GISDBASE: /media/teaiii/development1/grass"  > $HOME/.grassrc6
echo "LOCATION_NAME: ster"        >> $HOME/.grassrc6
echo "MAPSET: test_mapset"        >> $HOME/.grassrc6
echo "GRASS_GUI: wxpython"        >> $HOME/.grassrc6
#
export USER=$USER
export GISBASE=/usr/local/grass-7.2.1
#export GISDBASE=
export GISDBASE=/media/teaiii/development1/grass
export MAPSET=test_mapset
export LOCATION_NAME=ster
export GISRC=$HOME/.grassrc6
export PATH=$GISBASE/bin:$PATH:$GISBASE/scripts
export LD_LIBRARY_PATH=$LD_LIBRARY_PATH:$GISBASE/lib
#
################################################################

raster_map_name=$1
pt_cat=$2
vector_pt_map=$3
outputPATH=$4

#for i in `seq 1 $loc`;
# do
    echo "Processing Random Point GRASS cat: $pt_cat ..."

    loc=$(v.db.select map=$vector_pt_map columns=x where=cat=$pt_cat)
    a_loc=($loc)
    export x_loc=$(echo ${a_loc[1]})
```

```
74
75      loc=$(v.db.select map=$vector_pt_map columns=y where=cat=$pt_cat)
76      a_loc=($loc)
77      export y_loc=$(echo ${a_loc[1]})
78
79      # Copy Raster map to temp Map
80      g.copy --overwrite raster=$raster_map_name,xmrg_trans
81
82      #r.region map=xmrg_trans n=-3779868 s=-7694643 e=3057837 w=-1995176
83
84      perl -e '
85        use POSIX;
86        #              OFFSET     REF CENTER      SAMPLE POINT
87        #  x -- ADD: -90176.17 = 2416752.01754  - 2326575.84362
88        #  y -- ADD: -74643.28 = -5293585.88792 - (-5368229.16897)
89
90        $x = $ENV{"x_loc"};
91        $y = $ENV{"y_loc"};
92
93        $x_offset = sprintf("%.0f",2416752.01754 - $x);
94        $y_offset = sprintf("%.0f",-5293585.88792 - $y);
95
96        #  north:      -3779868 = -3705225 -74643
97        #  south:      -7694643 = -7620000 -74643
98        #  east:       3057837  = 3148013  -90176
99        #  west:       -1995176 = -1905000 -90176
100
101       $north = -3705225 - $y_offset;
102       $south = -7620000 - $y_offset;
103       $east  = 3148013  - $x_offset;
104       $west  = -1905000 - $x_offset;
105
106       # GRASS map 'xmrg_trans' is the translated xmrg file, which is ouput below
107       $cmd = "r.region map=xmrg_trans n=".$north." s=".$south." e=".$east." w=".$west;
108       printf("Executing %s...\n", $cmd );
109       system($cmd);
110     '
111
112 ##################################################
113 # Write out translated map and convert to xmrg raster_map_name
114 #
115 # xmrg raster name formated: xmrgMMDDYYYYHHz
116 #
117 ##################################################
118 year=$(echo $raster_map_name | cut -c9-12)
119 month=$(echo $raster_map_name | cut -c5-6)
120 day=$(echo $raster_map_name | cut -c7-8)
121 hour=$(echo $raster_map_name | cut -c13-14)
122
123 output_DateTime_str=${month}${day}${year}${hour}
124 echo "Output date-time string: $output_DateTime_str"
125
126 # CONUS Region
127 g.region -dp e=3433762.5 w=-1905000 n=-3424237.5 s=-7620000 res=4762.5
128
129 r.out.gdal --overwrite input=xmrg_trans output=precip_grid.asc format=AAIGrid nodata=-9999
130 /usr/local/bin/asctoxmrg -i precip_grid.asc -o xmrg${output_DateTime_str}z -p ster
131 mv xmrg${output_DateTime_str}z.gz $outputPATH
132
133 #   g.copy --overwrite raster=xmrg_trans,$newMap
134 #   g.copy --overwrite raster=xmrg_trans_ge2,$newMapGE2
135 #   r.to.vect --overwrite input=$newMapGE2 output=$newMapGE2 type=area
136
137 # done
```

```

echo
echo "================ Cleanup follows ================"
echo

# Cleanup -- Remove temporary GRASS files
$GISBASE/etc/clean_temp

# remove session tmp directory:
rm -rf /tmp/grass7-$USER-$GIS_LOCK

#rm temp_grid.prj temp_grid.asc temp_grid.asc.aux.xml NLDAS.apcp.grb NLDAS.tmp.grb precip_grid.
    asc

echo
echo "================ Cleanup completed ================"
echo
```

Listing B.16: grassTranslateXMRG shell script is called by run_QPF_TS_RDHM used to translate xmrg storm maps in GRASS GIS.

```
#!/bin/bash

#################################################################
# *************       importxmrg2grass            *************
# xmrg0625200613z.gz
# discharge0625200613z.gz
#
#################################################################

xmrg_type=$1
#year=$2
list=$2
file_path=$3

prefix=${file_path}/

#for file in $(find $file_path -name "$xmrg_type*$year*.gz") ; do
#for file in $(find $file_path -name "$xmrg_type*$year*") ; do
for file in $(cat $list) ; do

echo "${file}"

string=${file#$prefix}

#map_name=$(echo $string | cut -c1-20)
map_name=$(echo $string | cut -c1-15)
#file_name=${map_name}.gz
file_name=$map_name
echo "${map_name}"

ERRORFILE=script.errors

/home/teaiii/scripts/xmrg2newster $file_path $file_name $map_name 2>>$ERRORFILE

done
```

Listing B.17: importxmrg2grass shell script is used to import MPE xmrg files into GRASS GIS.

```
#!/bin/bash

############################################################
# *************          xmrg2newster           *************
############################################################
# Set initial GRASS Environment Variables
############################################################
# Program: xmrg2newster
#
# Written by: Thomas Adams
# Date:   10/18/2014
# Updated:  01/17/2015
#
# Program uses r.in.arc to import a text file containing XMRG files
#
# /home/teaiii/.grass7/rc
# MAPSET: adams
# GISDBASE: /media/teaiii/development1/grass
# LOCATION_NAME: newster
# GUI: wxpython
#
#
############################################################
#export GIS_LOCK=$$

############################################################
# Set GRASS Environment Variables
############################################################
#echo "GISDBASE: /home/teaiii/grass/data"   > $HOME/.grassrc6
echo "GISDBASE: /media/teaiii/development1/grass"  > $HOME/.grassrc6
echo "LOCATION_NAME: ster"        >> $HOME/.grassrc6
echo "MAPSET: teaiii"         >> $HOME/.grassrc6
echo "GRASS_GUI: wxpython"        >> $HOME/.grassrc6
#
export USER=$USER
export GISBASE=/usr/local/grass-7.2.1
#export GISDBASE=
export GISDBASE=/media/teaiii/development1/grass
export MAPSET=teaiii
export LOCATION_NAME=ster
export GISRC=$HOME/.grassrc6
export PATH=$GISBASE/bin:$PATH:$GISBASE/scripts
export LD_LIBRARY_PATH=$LD_LIBRARY_PATH:$GISBASE/lib
#
############################################################

file_path=$1
file_name=$2
map_name=$3

# OHRFC Region
#g.region -dp e=2614612.5 w=1423987.5 n=-4624387.5 s=-5862637.5
# MARFC Region
#g.region -dp e=3048000 w=2185987.5 n=-4381500 s=-5419725

echo "${file_path}"
echo "${file_name}"

  today=$(date +%Y%m%d)

  echo
  echo "=========  Begin processing of XMRG file $map_name ==========="
  echo

```

```
  xmrgtograss $file_path $file_name $map_name ster

  echo
  echo "=========  Step 1: Completed importing XMRG file into newster GRASS Location ==========="
  echo

echo
echo "==========  Clean-up GRASS DB $GISDBASE & LOCATION $LOCATION_NAME  =========="
echo

  # Cleanup -- Remove temporary GRASS files
  $GISBASE/etc/clean_temp
```

Listing B.18: xmrg2newster is a shell script called by importxmrg2grass.

```
#!/bin/bash

#############################################################
# Program: run_QPF_TS_RDHM
#
# Written by: Thomas Adams
# Date:   07/22/2017
# Updated:  07/25/2017
#
#############################################################
#
# Program runs RDHM simulations for QPF scenarios based on
# re-positioning NEXRAD radar based fields to reflect predetermined
# Threat Score values
#
# Process:
#
# (0) setup RDHM warm states from previous historical RDHM
#     model run -- done one time as all simulations start
#     from the same date-time
# (1) generate 1-hourly re-positioned xmrg files from GRASS
#     xmrg maps (imported previously)
# (2) generate RDHM control file
# (3) run RDHM using control file from (2)
#
#
#############################################################

xmrg_map_list=$1  # List of xmrg maps in GRASS GIS
qpf_list_PATH=$2  # PATH to file 'pt_cat_list' with GRASS vector point cats
vector_pt_map=$3  # Name of GRASS vector point map
startDate=$4    # YYYMMDDHH format

cd $xmrg_input_path
echo $xmrg_input_path

prefix="qpfts"
output_dir="/media/teaiii/Seagate8/rdhm/ohrfc/output"
rdhm_grids_dir="${output_dir}/grids/coldstart/nldas_grids"

echo $startDate
qpfts_year=$(echo $startDate | cut -c1-4)
qpfts_month=$(echo $startDate | cut -c5-6)
qpfts_day=$(echo $startDate | cut -c7-8)
qpfts_hr=$(echo $startDate | cut -c9-10)

endDate=$(date +"%Y%m%d" -d "$startDate 7 days" -u)
```

```
48
49 # Setup RDHM warm states
50 /home/teaiii/scripts/setupRDHMwarm_states $rdhm_grids_dir $startDate 6
51
52 #for filename in $(find . -name "qpfts.*.dat");
53
54 # For each pre-determined GRASS GIS vector point location (category) -- need xmrg map list
55 for pt_cat in $(<$qpf_list_PATH/pt_cat_list);
56   do
57
58     #echo $filename
59     #file="${filename##./}"
60
61 #   for i in `seq 1 11`;
62 #
63 #     do
64 #                   echo $i
65 #       if [ $i -lt 10 ]; then
66 #         ens_num="0"${i}
67 #       else
68 #         ens_num=$i
69 #       fi;
70 #     done
71
72     # Make Output subdir if it does not exist
73     xmrgOutputPATH=$output_dir/${startDate}/${pt_cat}/qpf
74     rdhmOutputPATH=$output_dir/${startDate}/${pt_cat}/grids
75     echo
76
77     if [ ! -d $xmrgOutputPATH ]; then
78
79           mkdir -p $xmrgOutputPATH;
80     fi;
81
82     if [ ! -d $rdhmOutputPATH ]; then
83
84           mkdir -p $rdhmOutputPATH;
85     fi;
86
87     # Reposition xmrg files as QPF for each hour of QPF found in ster GRASS GIS LOCATION
88     # for each GRASS raster map in list
89     for map_name in $(<$xmrg_map_list);
90       do
91       /home/teaiii/scripts/grassTranslateXMRG $map_name $pt_cat $vector_pt_map $xmrgOutputPATH
92     done
93
94     # Generate RDHM control file...
95     # inputPATH is the location of QPF files; outputPATH is where RDHM output should be written
       to
96     # which is different for each QPF scenario
97     /home/teaiii/scripts/perl/makeQPFTS_RDHMcontrol.pl $startDate $endDate $xmrgOutputPATH
       $rdhmOutputPATH
98
99     # run RDHM using control file
100    cd /home/teaiii/RDHM/greenbrier/input
101    rdhm rdhm_qpfts_control
102
103  done
104
105 exit
```

Listing B.19: run_QPF_TS_RDHM is the main shell script used to run the RDHM *monte carlo* hydrologic simulation experiment in Chapter 3.

```
#!/bin/bash

#################################################################
# Program: run_RDHMthreatScoreOutput2TS.sh
#
# Written by: Thomas Adams
# Date:   08/14/2017
# Updated:  09/01/2017
#
# From GRASS GIS analysis...
# ptsTS_filePATH: /home/teaiii/greenbrier_exp_random_pts.cs -- example...
#
# 1,1785704.09195794,-4725999.78752539,0,0
# 2,2012502.48292766,-5462366.76223028,51,0.00594197832925551
# 3,1956830.85439216,-4832299.18443406,0,0
# 4,2325365.72357129,-5416092.173973,639,0.0799249530956848
# 5,3243210.05695664,-5530243.64979288,5,0.000579441418472592
# 6,2055676.65000313,-4694853.41755759,0,0
# 7,2298779.17685,-5357818.10506215,1376,0.18958390741251
# 8,2991071.11891698,-5710947.34429017,0,0
# 9,3301670.73026136,-5184457.83943911,28,0.00325354403904253
# 10,1777929.3101368,-4930867.31850804,0,0
# 11,2361071.31951339,-5814232.53012011,1,0.000115834588208039
# 12,2920682.88454495,-4859494.26362881,157,0.0185207030789194
# 13,3229154.4837265,-5636408.06350202,0,0
# 14,1820600.45859242,-5778990.33301007,1,0.000115834588208039
# *
# *
# *
#
#################################################################

USGS_ID=$1
rdhm_output_path=$2 # Location of RDHM output subdirectories
ptsTS_filePATH=$3

cd $rdhm_output_path
echo $rdhm_output_path

rm $HOME/TS_maxval.out

for dir in $(ls) ;
  do

    echo "Entering ${rdhm_output_path}/${dir}/grids"
    cd ${rdhm_output_path}/${dir}/grids

    outputPATH=${rdhm_output_path}

    /home/teaiii/scripts/perl/rdhmOutletQ2tsCats.pl ${USGS_ID}_discharge_outlet.ts
    $ptsTS_filePATH $dir --ymd --cms
    mv ${USGS_ID}_discharge_outlet.ts.out $HOME/${USGS_ID}_${dir}_discharge_outlet.ts.dat
done
```

Listing B.20: run_RDHMthreatScoreOutput2TS.sh shell script used to calculate *threat score* from the RDHM *monte carlo* hydrologic simulation experiment in Chapter 3

```
#!/bin/bash

#################################################################
# Program: run_obsText2Datacard
#
```

```
# Written by: Thomas Adams
# Date:   03/30/2018
# Updated:  03/31/2018
#
#################################################################
#
# Run from directory: /media/teaiii/Seagate8/naefs
#
#
# Script location:  /home/teaiii/scripts/run_obsText2Datacard
#
# File location:  /home/teaiii/phd/add_evs_locs.txt
#
# Data file names:  ${lid}.obs.txt from psql...
#
# \o PRDK2.obs.txt
# select lid,obstime,z0000,z0600,z1200,z1800 from pecrsep \
# where pe1='H' and pe2='G' and t='R' and s='G' and lid='PRDK2' \
# order by obstime asc;
#
#################################################################

list_PATH=$1  # PATH to file 'add_evs_locs.txt'

# For each lid in the file
for lid in $(<$list_PATH/add_evs_locs.txt);
  do
    echo "processing ${lid}..."
    /home/teaiii/scripts/perl/pecrsepObs2TS.pl ${lid}.obs.txt
    /home/teaiii/scripts/perl/ts2datacard.pl ${lid}.obs.txt.out --id=${lid} --delta_t=6
  done

exit
```

Listing B.21: run_obsText2Datacard shell script is used to run the pecrsepObs2TS.pl and ts2datacard.pl Perl scripts.

```
#!/usr/bin/perl

use Getopt::Long;

#################################################################
# evsCRPSSXML2R.pl
#
# File Names have format: ELRP1.SSTG.Mean_continuous_ranked_probability_skill_score.xml
#
# Location: /home/teaiii/evs_analyses/ELRP1/ (for example)
#
# types: "MMEFS" "MMEFS_MEAN" "MMEFS_MEDIAN"
#
#
#
#sed -e '/All data/,/values/!d' ELRP1.SSTG.Mean_continuous_ranked_probability_skill_score.xml |
    grep values
#sed -e '/Pr=0.9/,/values/!d' ELRP1.SSTG.Mean_continuous_ranked_probability_skill_score.xml |
    grep values
#
#################################################################
my $TRUE  = 1;
my $FALSE = 0;
my $cms   = $FALSE;
my $tabs  = $TRUE;
```

```
my $type  = "MMEFS";
my $evs_dir = "/home/teaiii/evs_analyses";

my ($ymd,$verbose);
GetOptions('evs_dir=s'     => \$evs_dir,
     'cms!' => \$cms,
     'tabs!' => \$tabs,
     'type=s' => \$type,
          'verbose!' => \$verbose );

$lid_list   = $ARGV[0];

print "Processing file ",$lid_list,"\n";
# Open LID list file
open (FH, $lid_list);

@a = <FH>;
close(FH);

select STDOUT;
$nLID = scalar(@a);
printf("Number of lines = %d\n", $nLID);

if($type eq "MMEFS_MEAN") {
  $outFileName = "mmefs_mean_rpss.txt";
  open (outFH, ">$outFileName");
}
elsif($type eq "MMEFS_MEDIAN") {
  $outFileName = "mmefs_median_rpss.txt";
  open (outFH, ">$outFileName");
}
else{
  $outFileName = "mmefs_rpss.txt";
  open (outFH, ">$outFileName");
}

select(outFH);
printf("lid\ttype\tresponse\tleadtime\trpss_all\trpss_90\n");

$i = 0;
foreach $lid (@a) {

  if($i > 0){
    chomp($lid);
    ($id,$response) = split(/\,/, $lid);

    select STDOUT;
    printf("Processing Location: %s...\n",$id);

    # ALL
    $filename = $id.".SSTG.Mean_continuous_ranked_probability_skill_score.xml";
    $path = $evs_dir."/".$id."/".$filename;

    $cmd = "sed -e '/All data/,/values/!d' ".$path." | grep values";
    $str = `$cmd`;

    my @lines = split /\n/, $str;
    $nlines = scalar(@lines);

    for($j=0;$j<$nlines;$j++) {
      chomp($lines[$j]);
      @b = split /[<>]+/, $lines[$j];
      $rpss_all[$j] = sprintf("%6.4f\n", $b[2]);
      chomp($rpss_all[$j]);
```

```
    }

    # Pr=0.9
    $filename = $id.".SSTG.Mean_continuous_ranked_probability_skill_score.xml";
    $path = $evs_dir."/".$id."/".$filename;

    $cmd = "sed -e '/Pr=0.9/,/values/!d' ".$path." | grep values";
    $str = `$cmd`;

    my @lines = split /\n/, $str;
    $nlines = scalar(@lines);

    for($j=0;$j<$nlines;$j++) {
      chomp($lines[$j]);
      @b = split /[<>]+/, $lines[$j];
      $rpss_90[$j] = sprintf("%6.4f\n", $b[2]);
      chomp($rpss_90[$j]);
    }

    select(outFH);
    for($j=0;$j<$nlines;$j++) {
      $leadtime = ($j+1)*6;
      printf("%s\t%s\t%d\t%d\t%s\t%s\n",$id,$type,$response,$leadtime,$rpss_all[$j],$rpss_90[$j])
    ;
    }
  }
  $i++;
}


```

Listing B.22: evsCRPSSXML2R.pl Perl script reads EVS generated XML files for CRPSS values from MMEFS ensemble simulations for each forecast point location for import into R.

```
#!/usr/bin/perl

use Getopt::Long;

##############################################################
# mmefsXML2R.pl
#
# File Names have format: ATHO1.SSTG.Mean_absolute_error.xml
#       ATHO1.SSTG.Mean_error.xml
#       ATHO1.SSTG.Root_mean_square_error.xml
#
# Location: /home/teaiii/evs_analyses/ATHO1/ (for example)
#
# types: "MMEFS" "MMEFS_MEAN" "MMEFS_MEDIAN"
#
##############################################################
my $TRUE  = 1;
my $FALSE = 0;
my $cms   = $FALSE;
my $tabs  = $TRUE;
my $type  = "MMEFS";
my $evs_dir = "/home/teaiii/evs_analyses";

my ($ymd,$verbose);
GetOptions('evs_dir=s'     => \$evs_dir,
     'cms!' => \$cms,
```

```
27      'tabs!' => \$tabs,
28      'type=s' => \$type,
29            'verbose!' => \$verbose );
30
31 $lid_list   = $ARGV[0];
32
33 print "Processing file ",$lid_list,"\n";
34 # Open LID list file
35 open (FH, $lid_list);
36
37 @a = <FH>;
38 close(FH);
39
40 select STDOUT;
41 $nLID = scalar(@a);
42 printf("Number of lines = %d\n", $nLID);
43
44 if($type eq "MMEFS_MEAN") {
45   $outFileName = "mmefs_mean_stats.txt";
46   open (outFH, ">$outFileName");
47 }
48 elsif($type eq "MMEFS_MEDIAN") {
49   $outFileName = "mmefs_median_stats.txt";
50   open (outFH, ">$outFileName");
51 }
52 else{
53   $outFileName = "mmefs_stats.txt";
54   open (outFH, ">$outFileName");
55 }
56
57 select(outFH);
58 printf("lid\ttype\tresponse\tleadtime\tme\tmae\trmse\n");
59
60 $i = 0;
61 foreach $lid (@a) {
62
63   if($i > 0){
64     chomp($lid);
65     ($id,$respone) = split(/\,/, $lid);
66
67     select STDOUT;
68     printf("Processing Location: %s...\n",$id);
69
70     # ME
71     $filename = $id.".SSTG.Mean_error.xml";
72     $path = $evs_dir."/".$id."/".$filename;
73
74     $cmd = "sed -e '/All data/,/values/!d' ".$path." | grep values";
75     $str = `$cmd`;
76
77     my @lines = split /\n/, $str;
78     $nlines = scalar(@lines);
79
80     for($j=0;$j<$nlines;$j++) {
81       chomp($lines[$j]);
82       @b = split /[<>]+/, $lines[$j];
83       $me[$j] = sprintf("%6.4f\n", $b[2]);
84       chomp($me[$j]);
85     }
86
87     # MAE
88     $filename = $id.".SSTG.Mean_absolute_error.xml";
89     $path = $evs_dir."/".$id."/".$filename;
90
```

```
    $cmd = "sed -e '/All data/,/values/!d' ".$path." | grep values";
    $str = `$cmd`;

    my @lines = split /\n/, $str;
    $nlines = scalar(@lines);

    for($j=0;$j<$nlines;$j++) {
      chomp($lines[$j]);
      @b = split /[<>]+/, $lines[$j];
      $mae[$j] = sprintf("%6.4f\n", $b[2]);
      chomp($mae[$j]);
    }

    # RMSE
    $filename = $id.".SSTG.Root_mean_square_error.xml";
    $path = $evs_dir."/".$id."/".$filename;

    $cmd = "sed -e '/All data/,/values/!d' ".$path." | grep values";
    $str = `$cmd`;

    my @lines = split /\n/, $str;
    $nlines = scalar(@lines);

    for($j=0;$j<$nlines;$j++) {
      chomp($lines[$j]);
      @b = split /[<>]+/, $lines[$j];
      $rmse[$j] = sprintf("%6.4f\n", $b[2]);
      chomp($rmse[$j]);
    }

    select(outFH);
    for($j=0;$j<$nlines;$j++) {
      $leadtime = ($j+1)*6;
      printf("%s\t%s\t%d\t%d\t%s\t%s\t%s\n",$id,$type,$respone,$leadtime,$me[$j],$mae[$j],$rmse[
    $j]);
    }
  }
  $i++;
}

```

Listing B.23: mmefsXML2R.pl Perl script Perl script reads EVS generated XML files of ME MAE and RMSE from MMEFS forecast ensemble simulations for each forecast point location for import into R

```

#!/usr/bin/perl

use Getopt::Long;

##############################################################################
# ts2datacard.pl
#
# Perl script that takes a command line argument:
#
# Expected file format:
#
# 2011-03-01 05:00:00 UTC 18200
# 2011-03-01 05:30:00 UTC 18400
# 2011-03-01 06:00:00 UTC 18600
# 2011-03-01 06:30:00 UTC 18800
```

```
# 2011-03-01 07:00:00 UTC 19000
# 2011-03-01 07:30:00 UTC 19200
# 2011-03-01 08:00:00 UTC 19300
# 2011-03-01 08:30:00 UTC 19600
# 2011-03-01 09:00:00 UTC 19800
# 2011-03-01 09:30:00 UTC 19900
# 2011-03-01 10:00:00 UTC 20100
# 2011-03-01 10:30:00 UTC 20400
# 2011-03-01 11:00:00 UTC 20500
# 2011-03-01 11:30:00 UTC 20700
# 2011-03-01 12:00:00 UTC 20800
#
# Example usage:
#
#   ts2datacard.pl psnw2.obs.txt.out --id=PSNW2 --delta_t=6
#
# Default:
#
#   No TZ Code (e.g., UTC); with TZ Code, use --tz_code
#
##########################################################################

$TRUE  = 1;
$FALSE = 0;

#-------------------------------------------
# TS type definitions
#-------------------------------------------
$infw = "INFW";
$map  = "MAP";
$mapx = "MAPX";
$mat  = "MAT";
$pele = "PELE";
$pelv = "PELV";
$qin  = "QIN";
$qine = "QINE";
$raim = "RAIM";
$rocl = "ROCL";
$rqot = "RQOT";
$sasc = "SASC";
$smzc = "SMZC";
$spel = "SPEL";
$sqin = "SQIN";
$sqme = "SQME";
$sstg = "SSTG";
$stg  = "STG";
$stw  = "STW";
$swe  = "SWE";
$twel = "TWEL";

my $verify  = $FALSE;
my $SI    = $FALSE;
my $read_usgs = $FALSE;
my $delta_t = 1;
my $tz_code = $FALSE;
my $id    = "NONE";
my $fgid  = "NONE";
my $FLOW  = $FALSE;

my$result = GetOptions("si!" => \$SI,
           "verify!" => \$verify,
           "read_usgs!" => \$read_usgs,
           "flow!" => \$FLOW,
           "delta_t=f" => \$delta_t,
```

```
            "tz_code!" => \$TZ_CODE,
            "type=s" => \$data_type,
            "id=s" => \$id,
            "map_basin=s" => \$map_id);

#-------------------------------------------------------------
# Read input file
#-------------------------------------------------------------
$inFileName = $ARGV[0];

print $inFileName,"\n";
open (FH, $inFileName);

@tsdata = <FH>;
close(FH);

$i = 0;
foreach $one (@tsdata) {

  chomp($one);
  ($datetime_str,$value) = split(/\t/, $one);
  if($TZ_CODE) {
    ($date_str,$time_str,$time_zone_code) = split(/ +/, $datetime_str);
  }
  else {
    ($date_str,$time_str) = split(/ +/, $datetime_str);
  }
  ($year,$mo,$day) = split(/-/, $date_str);

  if($i == 0) {
    $start_year  = $year;
    $start_month = $mo;
    $start_day   = $day;
  }

  $yr2 = substr($year,2,2);

  if(index($one,":00:00")>=0){

    if($value <= -999) {
      $value = -999;
    }
    $outputData[$i] = sprintf("%-12s%2s%2s%4d%10.3f\n", $id,$mo,$yr2,$day,$value);

  }

  $i++;
}

$dataLines = $i;

$end_year  = $year;
$end_month = $mo;
$end_day   = $day;

  #-----------------------------------------------------------------
  # Prepare for output
  #-----------------------------------------------------------------

printf("%d missing hours found in USGS data...\n", $m);
&write_output;

sub write_output
  {
```

```
145   if($SI) {
146     if($FLOW) {
147       $dim    = "L3/T";
148       $units    = "CMS";
149     }
150     else {
151       $dim    = "L";
152       $units    = "M";
153     }
154   }
155   else {
156     if($FLOW) {
157       $dim    = "L3/T";
158       $units    = "CFS";
159     }
160     else {
161       $dim    = "L";
162       $units    = "FT";
163     }
164   }
165   $name   = $id;
166   $type   = "OBS";
167   $deltat   = $delta_t;
168   $description  = "OBSERVED USGS";
169   $period   = 24/$deltat;
170
171   $outFileName = $id.".".$type.".".$start_year.$start_month.$start_day."-".$end_year.$end_month.
        $end_day.".datacard";
172
173   select STDOUT;
174   print "Output file name: ",$outFileName,"\n";
175   open (outFH, ">$outFileName");
176
177   $start_year_str = substr($start_year,2,2);
178
179   #---------------------------------------------------------------
180   # Write header line for current time series
181   #---------------------------------------------------------------
182   $datacard = $id;
183   $fileNameLen = length($datacard);
184   if($fileNameLen > 12) {
185     $name = substr($datacard,0,12);
186   }
187   else {
188     $name = $datacard . ' ' x (12-$fileNameLen);
189   }
190
191   select outFH;
192   printf("\$ OFS DATACARD OUTPUT FROM ts2datacard.pl\n", );
193   printf("\$  IDENTIFIER=%-12s DESCRIPTION=%-20s\n", $id,$description);
194   printf("\$  PERIOD OF RECORD=%2s/%4s THRU %2s/%4s\n", $start_month,$start_year,$end_month,
        $end_year);
195   printf("\$  SYMBOL FOR MISSING DATA=-999.00   SYMBOL FOR ACCUMULATED DATA=-998.00\n", );
196   printf("\$  TYPE=%-4s   UNITS=%-2s     DIMENSIONS=%-3s    DATA TIME INTERVAL=%2d HOURS\n",
        $type,$units,$dim,$deltat);
197   printf("\$  OUTPUT FORMAT=(3A4,2I2,I4,1F10.3)\n");
198   printf("%-12s  %-4s %-4s %-4s %2d   %-12s   %-20s\n", $name,$type,$dim,$units,$deltat,$id,
        $description);
199   printf("%2s  %4s %2s   %4s %2d  1F10.3\n", $start_month,$start_year,$end_month,$end_year,1);
200
201   #---------------------------------------------------------------
202   # Pad TS data with -999.00 as needed, beginning with the end of
203   # 1st time period following 00Z (depending on time step)
204   #---------------------------------------------------------------
```

```
  $day = 1;
  $num_periods = 2;
  $FIRST_TIME = $TRUE;
  $start = 1;
  #$num = (($start_day - 1)*$period)+$first_period;
  $num = $start_day*$period - 1 - ($period - $first_period);
  for($j=$start;$j<=$num;$j++) {
    $output = sprintf("%-12s%2s%2s%4d%10.3f\n", $id,$start_month,$start_year_str,$day,-999.00);

    printf("%s", $output);

    if($num_periods >= $period) {
      $day++;
      $num_periods = 0;
      if($FIRST_TIME){
        $FIRST_TIME = $FALSE;
        $start = 0;
      }
    }
    $num_periods++;
  }

  #-------------------------------------------------------------
  # Write the data
  #-------------------------------------------------------------

  select outFH;
  for($j=0;$j<$dataLines;$j++) {
    printf("%s", $outputData[$j]);
  }
  close outFH;
}
```

Listing B.24: ts2datacard.pl Perl script reformats ascii format time-series data into NWSRFS DATACARD format files for use in the EVS for MMEFS ensemble forecast verification.

```

#!/usr/bin/perl

use Getopt::Long;

#------------------------------------------------------------------------
# prdutil2ts.pl
#
# Reads NWSRFS PRDUTIL TSDATA Datacard format files and outputs
# data to SHEF format. ONLY observed data (for all IDs) are reformatted
# unless a single MAP basin is specified -OR- if future data are
# specified (either for all IDs or just one)
#------------------------------------------------------------------------
# Written by: Thomas Adams, NOAA/NWS/OHRFC, 07/13/2010
#------------------------------------------------------------------------
# The Perl script takes a command line argument:
#
# (1) name of output file from the NWSRFS OFS PRDUTIL TSDATA command
#
#
# Optionally (if present):
#
#
# --future  to set that future data is desired as well; otherwise, -ONLY-
#     OBSERVED data will be retrieved.
#
```

```
27 # --map_basin to process only a single specified MAP_BASIN_ID
28 #
29 # --type    to identify the TS type if a single MAP basin is specified
30 #
31 # -h or --help  to get help on running 'ts2hec.pl' (NOT IMPLEMENTED YET)
32 #
33 # -f or --file  to identify a file containing a list of Time (NOT IMPLEMENTED YET)
34 #     Series IDs & types to extract from the input
35 #     file
36 #
37 # Usage:
38 #   ts2hec.pl tsdata.20040817 (all time series for all IDs -- DATACARD format)
39 #
40 #   ts2hec.pl tsdata.20040817 --map_basin=PSNW2 (all time series for ID = PSNW2 -- DATACARD
    format)
41 #
42 #   ts2hec.pl tsdata.20040817 --list=some_file_name (all time series for TSIDs in the list)
43 #
44 #
45 # Note: Output consists of individual files named:
46 #
47 #
48 # Also: *ONLY* the observed data period is written out as the default
49 #
50 #-------------------------------------------------------------------------

52 $TRUE   = 1;
53 $FALSE    = 0;
54 $USE_LIST = $FALSE;
55
56 #-------------------------------------------
57 # TS type definitions
58 #-------------------------------------------
59 $infw = "INFW";
60 $map  = "MAP";
61 $mapx = "MAPX";
62 $mat  = "MAT";
63 $pele = "PELE";
64 $pelv = "PELV";
65 $qin  = "QIN";
66 $qine = "QINE";
67 $raim = "RAIM";
68 $rocl = "ROCL";
69 $rqot = "RQOT";
70 $sasc = "SASC";
71 $smzc = "SMZC";
72 $spel = "SPEL";
73 $sqin = "SQIN";
74 $sqme = "SQME";
75 $sstg = "SSTG";
76 $stge = "STGE";
77 $stg  = "STG";
78 $stw  = "STW";
79 $swe  = "SWE";
80 $twel = "TWEL";
81
82 $datacard   = "datacard";
83 $no_future_data   = "FUTURE DATA   FUTURE TIME SERIES ID=  NONE";
84 $future_data_str  = "FUTURE DATA";
85 $NEXT_ID    = "0TIME SERIES ID=";
86 $dashed_line    = "0------------------------------------------------------------";
87 $none_str   = "    NONE";
88 #-------------------------------------------
89
```

```
my $list_file     = "NONE";
my $map_id    = "ALL_IDS";
my $map_ts_type   = "NONE";
my $id      = "NONE";
my $FUTURE_DATA   = $FALSE;

my$result = GetOptions("list=s" => \$list_file,
           "type=s" => \$map_ts_type,
           "future!" => \$FUTURE_DATA,
           "map_basin=s" => \$map_id);

#---------------------------------------------------------------
# Datum file path
#---------------------------------------------------------------
#$datum_file_path_name = "/awips/rep/lx/rfc/nwsrfs/ofs/input/oper/prdutil/DATUM.OHWTMP";
#$datum_file_path_name = "/awips/hydroapps/lx/local_apps/ofsshef/input/ts2shef.input";

#---------------------------------------------------------------
# Read input PRDUTIL TSDATA file
#---------------------------------------------------------------
$inFileName = $ARGV[0];

print $inFileName,"\n";
open (FH, $inFileName);

@tsdata = <FH>;
close(FH);

$len_tsdata = scalar(@tsdata);

$today = `date +"%Y%m%d%H%M" -u`;
chomp($today);
printf("Today's date & time = %s...\n", $today);

#-------------------------------------------------------
# MAIN DECISION POINT -- use list Y/N?
#-------------------------------------------------------
# If we're processing all data or just one MAP area;
#
# We check FIRST to see if we're retrieving FUTURE data,
#-------------------------------------------------------

if($map_id eq "ALL_IDS") {

  $outFileName = $inFileName.".shef";
  select STDOUT;
  print "Output file name: ",$outFileName,"\n";
  open (outFH, ">$outFileName");
  $FIRST_OPEN = $FALSE;

  printf("*****************************************\n");
  printf("****** Open %s ******\n", $outFileName);
  printf("*****************************************\n");

}
else {
  if($map_ts_type eq "NONE") {
    printf("The Time Series TYPE must be provided by, e.g. --type=PELV...EXITING...\n");
    exit;
  }

  $outFileName = $map_id.".shef";
  select STDOUT;
  print "Output file name: ",$outFileName,"\n";
```

```
154   open (outFH, ">$outFileName");
155 }
156
157 &main;
158 &write_data;
159
160 #------------------------------------------------------------
161 # Define 'Main' subroutine
162 #------------------------------------------------------------
163
164 sub main
165   {
166
167   $FOUND_FUTURE_DATA = $FALSE;
168   $FIRST_PASS = $TRUE;
169   $i = 0;
170   $tstype = 0;
171
172   foreach $one (@tsdata) {
173
174     if(index($one, "0TIME SERIES ID=") >= 0){
175
176      $prev_id = $id;
177
178      ($junk,$idStr,$typeStr,$unitsStr,$deltatStr,$vstepStr,$descriptionStr,$qpfStr) = split(/=/,
     $one);
179        ($id)           = split(/ +/, $idStr);
180        ($type_str)        = split(/ +/, $typeStr);
181        ($units)         = split(/ +/, $unitsStr);
182        ($deltat[$tstype])     = split(/ +/, $deltatStr);
183        ($vstep)         = split(/ +/, $vstepStr);
184
185        $deltat[$tstype] = substr($deltatStr,0,2);
186        if(index($deltat[$tstype], " ") >= 0) {
187          $deltat[$tstype] = substr($deltatStr,1,1);
188        }
189
190      $dt_str[$tstype] = $deltat[$tstype]."HOUR";
191      $period = 24/$deltat[$tstype];
192
193      $description = substr($descriptionStr,0,20);
194
195      select STDOUT;
196      printf("%s %s %s %s %s\n", $id,$type_str,$units,$deltat[$tstype],$description);
197      $nextID[$tstype] = $id;
198      $nextType[$tstype] = $type_str;
199     }
200
201    $map_id = $id;
202
203    #------------------------------------------------------------
204    # 'REGULAR DATA' line marks the beginning of observed
205    # data, so start reading the observed data; otherwise
206    # read 'FUTURE DATA'
207    #------------------------------------------------------------
208    if(index($one, "REGULAR DATA") >= 0){
209
210      printf("Reading 6-hr REGULAR DATA...\n");
211
212      $type[$tstype] = $type_str;
213      $k = 1;
214      $start = $i+1;
215      $LOOP = $TRUE;
216      #------------------------------------------------------
```

```
217       # 1-hour data
218       #---------------------------------------------------------
219       if($deltat[$tstype] == 1) {
220         while($LOOP) {
221           #-------------------------------------------------------------------------
222           # Test if the current & previous lines have '/'
223           #-------------------------------------------------------------------------
224           if(index($tsdata[$i+$k], "\/") < 0 && index($tsdata[$i+$k-1], "\/") < 0) {
225             if(index($tsdata[$i+$k], $future_data_str) >= 0) {
226
227               $FOUND_FUTURE_DATA = $TRUE;
228               select STDOUT;
229               printf("Found future data... %d %d\n", $k,$end_obs_line);
230
231               $future_lines = $i;
232               printf("Future lines = %d...\n", $future_lines);
233
234               if (index($tsdata[$i+$k+1], $none_str) >= 0) {
235                 $k++;
236                 printf("No  future data found... %d\n", $k);
237                 $LOOP = $FALSE;
238               }
239               #elsif(index($tsdata[$i+$k-1], "NWSRFS FORECAST SYSTEM") < 0) {
240               # if (index($tsdata[$i+$k], $future_data_str) < 0) {
241               #   $k++;
242               #   printf("End of future data... %d\n", $k);
243               #   $LOOP = $FALSE;
244               # }
245               #}
246               $k++;
247               $end_obs_line = $i+$k;
248             }
249             elsif(index($tsdata[$i+$k], $no_future_data) >= 0) {
250               printf("No future data found... %d\n", $k);
251               $LOOP = $FALSE;
252             }
253             elsif(index($tsdata[$i+$k], $dashed_line) >= 0) {
254               printf("Dashed line found... %d\n", $i+$k);
255               $LOOP = $FALSE;
256             }
257             elsif(index($tsdata[$i+$k], ">>>>>>>> END") >= 0) {
258               printf("File line end found... %d\n", $i+$k);
259               $LOOP = $FALSE;
260             }
261             #elsif(index($tsdata[$i+$k-1], "NWSRFS FORECAST SYSTEM") < 0) {
262             # if(length($tsdata[$i+$k-1]) > 1) {
263             #   $LOOP = $FALSE;
264             # }
265             #}
266             else {
267               $k++;
268             }
269           }
270           select STDOUT;
271           printf("k = %d\n", $k);
272           #printf("%s",$tsdata[$i+$k]);
273           $k++;
274         } # end 1-hr loop
275       }
276       #---------------------------------------------------------
277       # 6-hour data (assumed)
278       #---------------------------------------------------------
279       else {
280         printf("Reading 6-hr FUTURE data...\n");
```

```
        $FUTURE_DATA_LINE = $i+$k;
        while($LOOP) {

          if(index($tsdata[$i+$k], "\/") < 0) {

            if(index($tsdata[$i+$k], $future_data_str) >= 0) {

              $FOUND_FUTURE_DATA = $TRUE;
              select STDOUT;
              printf("Found future data... %d %d\n", $k,$end_obs_line);

              $future_lines = $i - 3;
              printf("Future lines = %d...\n", $future_lines);

              if (index($tsdata[$i+$k+1], $none_str) >= 0) {
                printf("No  future data found... %d\n", $k);
                $LOOP = $FALSE;
              }
              elsif(index($tsdata[$i+$k-1], "NWSRFS FORECAST SYSTEM") < 0) {
                if (index($tsdata[$i+$k], $future_data_str) < 0) {
                  printf("End of future data... %d\n", $k);
                  $LOOP = $FALSE;
                }
              }
              $k++;
              $end_obs_line = $i+$k;
            }
            elsif(index($tsdata[$i+$k], $no_future_data) >= 0) {
              $k++;
              printf("No  future data found... %d\n", $k);
              $LOOP = $FALSE;
            }
            elsif(index($tsdata[$i+$k-1], "NWSRFS FORECAST SYSTEM") < 0) {
              if(length($tsdata[$i+$k-1]) > 1) {
                $LOOP = $FALSE;
              }
            }
            else {
              $k++;
            }
          }
          select STDOUT;
          printf("k = %d\n", $k);
          #printf("%s",$tsdata[$i+$k]);
          $k++;
        } # end 6-hr loop
      }

      $end = $i+$k-1;
      #if($deltat[$tstype] == 1) {
      #   $end = $i+$k;
      #}

      if($FUTURE_DATA == $FALSE) {
        $end = $end_obs_line;
      }
      printf("start = %d, end = %d, DT = %d\n", $start,$end, $deltat[$tstype]);
      &extract_data;
      $tstype++;
    }
    #-----------------------------------------------------------
    # End 'REGULAR DATA'
    #-----------------------------------------------------------
    $i++;
```

```
345
346    } # End foreach
347 } # End sub
348
349 #------------------------------------------------------------
350 # Define 'Extract Data' subroutine
351 #------------------------------------------------------------
352
353 sub extract_data {
354
355   $dataLines = 0;
356   $less[$tstype] = 0;
357
358   select STDOUT;
359   printf("extract_data(): number of lines, i = %d %d\n", $nlines,$i);
360
361   $REDUCE_LINES = $FALSE;
362   $lines = 0;
363   $lessLines = 0;
364   for($j=$start;$j<=$end;$j++) {
365
366     #printf("%s", $tsdata[$j]);
367
368     if(index($tsdata[$j], "NWSRFS FORECAST SYSTEM") >= 0) {
369       shift @tsdata;
370       $REDUCE_LINES = $TRUE;
371       $lessLines++;
372     }
373     elsif(index($tsdata[$j], $future_data_str) >= 0) {
374       $REDUCE_LINES = $TRUE;
375       $lessLines++;
376     }
377     else {
378       $data_line[$lines] = $tsdata[$j];
379       #printf("%d %s", $lines,$data_line[$lines]);
380       $lines++;
381     }
382   }
383
384   #if($deltat[$tstype] == 1) {
385   # $lessLines = 0;
386   #}
387
388   if($REDUCE_LINES == $TRUE) {
389     $lines = $lines-$lessLines;
390   }
391
392   printf("%d lines will be processed; lessLines = %d...\n", $lines,$lessLines);
393   $k = 0;
394   $first_time = $TRUE;
395   $first_line = $TRUE;
396   $num_counter = 0;
397
398   #------------------------------------------------------------
399   # Read the data lines
400   #------------------------------------------------------------
401   for($j=0;$j<$lines;$j++) {
402
403     printf("Inside extract_data() %s",$data_line[$j]);
404
405     $data[0] = substr($data_line[$j],1,13);
406     for($m=1;$m<=12;$m++) {
407       $data[$m] = substr($data_line[$j],15+($m-1)*9,8);
408       printf("%d %s\n", $m,$data[$m]);
```

```
409
410       #-------------------------------------------------------
411       # If we have trailing white space on the last line, it'll
412       # get converted to ZERO values & get written out. So, count
413       # the occurences and NOT write it out in 'sub write_data'
414       #-------------------------------------------------------
415       $_ = $data[$m];
416       s/\s+$//;
417       $data[$m] =$_;
418       if(length($data[$m]) == 0) {
419         $less[$tstype]++;
420         #printf("%d *%s*\n", $m,$data[$m]);
421       }
422     }
423
424     $len = scalar(@data);
425     #printf("%s\n", $len);
426     $num = $len;
427
428     #-------------------------------------------------------------
429     # Get the date for the current line; for 1-hour data the 2nd
430     # line does not have the date, which is the NEXT day
431     #-------------------------------------------------------------
432     if($deltat[$tstype] == 1 && $first_line == $FALSE) {
433       $dateStr = sprintf("%s/%s/%s", $month,$day,$year);
434       printf("DateStr: %s\n", $dateStr);
435       $mo_day_yr = `date +"%m %d %Y" -d "$dateStr"`;
436
437       ($month,$day,$yr) = split(/ +/, $mo_day_yr);
438       chomp $yr;
439       $first_line == $TRUE;
440       #printf("1-hr, 2nd line...\n");
441     }
442     else {
443       $date[$k] = $data[0];
444       ($month,$day,$year,$time_str) = split(/\//, $date[$k]);
445       $time = substr($time_str,0,2);
446       $time_len = length($time);
447
448       if($time_len == 2) {
449         $start_time = $time;
450         $timeStr = $time."00";
451       }
452       else {
453         $start_time = substr($time_str,1,1);
454         $timeStr = "0".$time."00";
455       }
456
457       if($deltat[$tstype] == 1) {
458         $first_line = $FALSE;
459       }
460     }
461
462     if($first_time == $TRUE) {
463       $first_time = $FALSE;
464       $counter = $start_time/$deltat[$tstype];
465
466       $first_period = $counter;
467       $start_month = $month;
468       $start_day   = $day;
469       $start_year_str = $year;
470
471       if($year < 40) {
472         $start_year = 2000+$year;
```

```
473       }
474       else {
475         $start_year = 1900+$year;
476       }
477       $yr = $start_year;
478       $dateStr = sprintf("%s/%s/%s", $month,$day,$year);
479       printf("Date String: %s\n", $dateStr);
480     }
481
482     for($m=1;$m<$num;$m++) {
483
484       #-------------------------------------------------------------
485       # Skip zero-length data values
486       #-------------------------------------------------------------
487       if(length($data[$m])!=0) {
488
489         if(index($data[$m], "/") >= 0) {
490           ($new,$trash) = split(/\//, $data[$m]);
491           $data[$m] = $new;
492         }
493
494         if($month == 1) {
495           $mo_str = "01";
496         }
497         elsif($month == 2) {
498           $mo_str = "02";
499         }
500         elsif($month == 3) {
501           $mo_str = "03";
502         }
503         elsif($month == 4) {
504           $mo_str = "04";
505         }
506         elsif($month == 5) {
507           $mo_str = "05";
508         }
509         elsif($month == 6) {
510           $mo_str = "06";
511         }
512         elsif($month == 7) {
513           $mo_str = "07";
514         }
515         elsif($month == 8) {
516           $mo_str = "08";
517         }
518         elsif($month == 9) {
519           $mo_str = "09";
520         }
521         elsif($month == 10) {
522           $mo_str = "10";
523         }
524         elsif($month == 11) {
525           $mo_str = "11";
526         }
527         elsif($month == 12) {
528           $mo_str = "12";
529         }
530
531         $dateStr = $mo_str."/".$day."/".$yr;
532         printf("%s...\n", $dateStr);
533
534         #-------------------------------------------------------------
535         # Handle '**NONE**' in future data (different from NULL or missing data)
536         #-------------------------------------------------------------
```

```
537         if(index("**NONE**",$data[$m]) == 0) {
538           $data[$m] = -999.;
539         }
540 
541           $outputData[$tstype][$dataLines] = sprintf("%s,%s,%s,%s,", $fgid,$map_id,$dateStr,
    $timeStr);
542           chomp($data[$m]);
543           $data_t[$num_counter] = $data[$m];
544 
545           printf("%s %s %f %s...\n", $dateStr,$timeStr,$data_t[$num_counter],$type[$tstype]);
546 
547           #-------------------------------------------------------------
548           # For Time Series types:
549           #
550           # STG,SSTG,STGE,TWEL,PELV,SPEL,PELE,STW
551           #
552           # add the gauge datum
553           #-------------------------------------------------------------
554           if(index("STG",$type[$tstype]) == 0 ||
555              index("SSTG",$type[$tstype]) == 0 ||
556              index("STGE",$type[$tstype]) == 0 ||
557                index("TWEL",$type[$tstype]) == 0 ||
558                index("PELV",$type[$tstype]) == 0 ||
559                index("PELE",$type[$tstype]) == 0 ||
560                index("STW",$type[$tstype]) == 0 ||
561                index("SPEL",$type[$tstype]) == 0) {
562 
563             #printf("Processing Time Series types: STG,SSTG,STGE,TWEL,PELV,SPEL,PELE,STW\n");
564 
565             if($data_t[$num_counter] != -999) {
566               if($id_datum[$tstype] != -999) {
567                 if(length($data_t[$num_counter])!=0) {
568 
569                   $dataVal[$tstype][$dataLines] = $data_t[$num_counter];
570                   $idType[$tstype][$dataLines] = $datum_type[$tstype];
571                   printf("Adding datum for: %s %s %s %s %s %s...\n",
572                     $map_id,
573                     $datum_type[$tstype],
574                     $data_t[$num_counter],
575                     $id_datum[$tstype],
576                     $dataVal[$tstype][$dataLines],
577                     $tstype);
578                 }
579               }
580               else {
581                 $dataVal[$tstype][$dataLines] = $data_t[$num_counter];
582               }
583             }
584             else {
585               $dataVal[$tstype][$dataLines] = $data_t[$num_counter];
586             }
587           }
588           else {
589             $dataVal[$tstype][$dataLines] = $data_t[$num_counter];
590           }
591 
592           $num_counter++;
593           $dataLines++;
594 
595       }
596 
597       if($counter == $period) {
598         $dateStr = sprintf("%s/%s/%s", $month,$day,$year);
599         $mo_day_yr = `date +"%m %d %Y" -d "$dateStr +1 day"`;
```

```
600         printf("Month-Day-Year: %s\n", $mo_day_yr);
601         ($month,$day,$yr) = split(/ +/, $mo_day_yr);
602         chomp $yr;
603 
604         $end_month = $month;
605         $end_day   = $day;
606         $end_year  = $yr;
607 
608         $year = substr($yr,2,2);
609         $counter = 0;
610       }
611 
612       $counter++;
613 
614       $curr_time = $counter*$deltat[$tstype];
615       $curr_time_len = length($curr_time);
616       if($curr_time_len == 2) {
617         if($curr_time == 24) {
618           #$timeStr = "0000";
619           $timeStr = "2400";
620         }
621         else {
622           $timeStr = $curr_time."00";
623         }
624       }
625       else {
626         $timeStr = "0".$curr_time."00";
627       }
628       $k++;
629     }
630   }
631   $last_data_value[$tstype] = $dataLines;
632   $numDataLines[$tstype] = $dataLines;
633   printf("%d lines for Type = %d...\n", $dataLines,$tstype);
634 }
635 
636 #-------------------------------------------------------------
637 # Define 'Write Data' subroutine
638 #-------------------------------------------------------------
639 
640 sub write_data
641   {
642 
643   $dt = $deltat[0];
644   $numLines = $numDataLines[0];
645   $tsIndex = 0;
646 
647   #-----------------------------------------------------------------
648   # Write header line for current time series
649   #-----------------------------------------------------------------
650   select STDOUT;
651   $ts = "";
652   for($j=0;$j<$tstype;$j++) {
653     $ts = $ts.",".$type[$j];
654 
655     if($deltat[$j] > $dt) {
656       $dt = $deltat[$j];
657       if($numDataLines[$j] >= $numLines) {
658         $numLines = $numDataLines[$j];
659       }
660       $tsIndex = $j;
661     }
662 
663     #-----------------------------------------------------------------
```

```
664     # Find the time series start date & time
665     #-----------------------------------------------------------------
666     ($fg_str,$id_str,$date_str[$j],$time_str[$j]) = split(/\,/, $outputData[$j][0]);
667     $hour_str[$j] = substr($time_str[$j],0,2);
668     $min_str[$j] = substr($time_str[$j],2,2);
669     $date_time_str[$j] = $date_str[$j]." ".$hour_str[$j];
670     printf("Time series start date & time -- date_time_str: %s\n", $date_time_str[$j]);
671     $start_date_time_str[$j] = $date_time_str[$j];
672   }
673
674   #-----------------------------------------------------------------
675   # Write the data
676   #-----------------------------------------------------------------
677
678   $j = 0;
679   for($k=0;$k<$tstype;$k++) {
680     for($i=0;$i<$numDataLines[$k];$i++) {
681       $hrStr = $deltat[$k]*$i;
682       $next_date_string = $start_date_time_str[$k]." +".$hrStr." hours";
683       $next_date[$i] = `date +"%Y%m%d" -u -d "$next_date_string"`;
684       chomp($next_date[$i]);
685       $next_hour[$i] = `date +"%H" -u -d "$next_date_string"`;
686       chomp($next_hour[$i]);
687
688       if($next_hour[$i] == 00) {
689         $next_hour[$i] = 24;
690         $next_date_string = $prevDate;
691         $next_date[$i] = `date +"%Y%m%d" -u -d "$next_date_string"`;
692         chomp($next_date[$i]);
693       }
694
695       #printf("%s\n",$next_date[$i]);
696       if($deltat[$k] == 1) {
697         $output_Data[$j] = sprintf(".A %s %s Z DH%s/DC%s/DUE/XXHRZ ", $nextID[$k],$next_date[$i],
    $next_hour[$i],$today);
698       }
699       else {
700         $output_Data[$j] = sprintf(".A %s %s Z DH%s/DC%s/DUE/XXQRZ ", $nextID[$k],$next_date[$i],
    $next_hour[$i],$today);
701       }
702
703       $datVal[$j] = $dataVal[$k][$i];
704       #printf("Data Line output: %s %d %d\n", $datVal[$j],$i,$k);
705       $dstr[$j] = sprintf("%7.2f",$datVal[$j]);
706       if ("TWEL" eq $nextType[$k]) {
707         $shef_type[$j] = HT;
708       }
709       elsif("PELV" eq $nextType[$k]) {
710         $shef_type[$j] = HP;
711       }
712       else {
713         $shef_type[$j] = HG;
714       }
715
716       $prevDate = $next_date[$i];
717       $j++;
718     }
719   }
720   $nperiods = $j;
721
722   select outFH;
723   for($j=0;$j<$nperiods;$j++) {
724     $_ = $output_Data[$j];
725     s/XX/$shef_type[$j]/g;
```

```
726     $output_Data[$j] = $_;
727 
728     if(index($dstr[$j], "-999") >= 0) {
729       $_ = $dstr[$j];
730       s/-999./-9999./g;
731       $dstr[$j] = $_;
732     }
733     printf("%s%s\n", $output_Data[$j],$dstr[$j]);
734   }
735 }
736 close outFH;
```

Listing B.25: prdutil2ts.pl Perl script reformats ascii format files from NWSRFS OFS PRDUTIL output into ascii format time-series files.

```
1 
2 #!/usr/bin/perl
3 
4 use Getopt::Long;
5 
6 ################################################################################
7 # pecrsepObs2TS.pl
8 #
9 ################################################################################
10 my $TRUE   = 1;
11 my $FALSE  = 0;
12 my $cms    = $FALSE;
13 my $tabs   = $TRUE;
14 my $cumsum = $FALSE;
15 
16 my ($ymd,$verbose);
17 GetOptions('ymd!'     => \$ymd,
18      'cms!' => \$cms,
19      'tabs!' => \$tabs,
20      'cumsum!' => \$cumsum,
21            'verbose!' => \$verbose );
22 
23 $inFileName = $ARGV[0];
24 print "Processing file ",$inFileName,"\n";
25 # Open TS file
26 open (FH, $inFileName);
27 
28 @a = <FH>;
29 close(FH);
30 
31 select STDOUT;
32 $numLines = scalar(@a);
33 printf("Number of lines = %d\n", $numLines);
34 
35 $outFileName = $inFileName . ".out";
36 # Open file for writing reformatted data
37 open (outFH, ">$outFileName");
38 
39 sub  trim { my $s = shift; $s =~ s/^\s+|\s+$//g; return $s };
40 
41 $i = 0;
42 #-------------------------------------------
43 # Loop through each line of file - main loop
44 #-------------------------------------------
45 foreach $line (@a) {
46 
47   chomp($line);
```

```

  if($i > 1){
    if(index($line, "|") > 0){
      #($lid,$obstime,$z0000,$z0600,$z1200,$z1800) = split(/\|/, $line);
      ($lidz,$obstimez,$z0000z,$z0600z,$z1200z,$z1800z) = split(/\|/, $line);

      $lid = trim($lidz);
      $obstime = trim($obstimez);
      $z0000 = trim($z0000z);
      $z0600 = trim($z0600z);
      $z1200 = trim($z1200z);
      $z1800 = trim($z1800z);

      #select STDOUT;
      #printf("%s %s %s %s %s %s\n",$lid,$obstime,$z0000,$z0600,$z1200,$z1800);

      select outFH;
      printf("%s 00:00:00\t%s\n",$obstime,$z0000);
      printf("%s 06:00:00\t%s\n",$obstime,$z0600);
      printf("%s 12:00:00\t%s\n",$obstime,$z1200);
      printf("%s 18:00:00\t%s\n",$obstime,$z1800);
    }
  }
  else {
    select outFH;
    #printf("date\tstage\n");
  }
  $i++;
  #close(outFH);
}
```

Listing B.26: pecrsepObs2TS.pl Perl script reformats ascii format files from PostgreSQL queries of the NWS verification database *pecrsep* observation data tables for conversion to time-series format.

```
#!/usr/bin/perl

use Getopt::Long;

##############################################################################
# rdhmOutletQ2ts.pl
#
##############################################################################
my $TRUE  = 1;
my $FALSE = 0;
my $cms   = $FALSE;
my $tabs  = $TRUE;

my ($ymd,$verbose,$cms,$tabs);
GetOptions('ymd!' => \$ymd,
     'cms!'   => \$cms,
     'tabs!'  => \$tabs,
           'verbose!' => \$verbose );

$inFileName = $ARGV[0];
print "Processing file ",$inFileName,"\n";
# Open TS file
open (FH, $inFileName);

@a = <FH>;
close(FH);
```

```
27
28 $outputFileName = $inFileName.".out";
29 open(outputFH, ">$outputFileName");
30
31 #-------------------------------------------
32 # Keyword definitions
33 #-------------------------------------------
34 $datacard = "DATACARD";
35 #-------------------------------------------
36
37 $present_year = `date +%Y`;
38 chomp($present_year);
39 $next_year  = `date +'%Y' -d 'next year'`;
40 chomp($next_year);
41 $today  = `date +%Y%m%d`;
42 chomp($today);
43
44 printf("Present Year = %s, Next Year = %s, Today = %s\n",$present_year,$next_year,$today);
45
46 select STDOUT;
47 $numLines = scalar(@a);
48 printf("Number of lines = %d\n", $numLines);
49
50 $i=0; # Counter for each line read
51 $k=0; # Counter for each data value read
52 $lines = 0; # Total lines in file
53 $total_vals = 0;
54 $period = 1;
55 $FIRST_TIME = $TRUE;
56 $NO_SKIP = $TRUE;
57 $LAST_DATE = $FALSE;
58 $PROCESS_TS = $FALSE;
59
60 #-------------------------------------------
61 # Loop through each line of file - main loop
62 #-------------------------------------------
63 foreach $one (@a) {
64
65   #-----------------------------------------
66   # Test line for beginning of TS data
67   # Set Flag to process data only
68   #-----------------------------------------
69   if (index($one, "\$") != 0) {
70     if($NO_SKIP == $TRUE) {
71
72       chomp($one);
73       printf("%s\n",$one);
74       #RDHM OUTPUTS  SQIN L3/T CMS  1    CRCO1
75       ($model,$output_str,$ts_type,$dimensions,$units,$deltat,$id) = split(/ +/, $one);
76       #01  2011 10   2011  1   F13.4
77       ($start_mo_str,$start_yr,$end_mo_str,$end_yr,$num_cols,$fmt) = split(/ +/, $a[$lines+1]);
78
79       chomp($fmt);
80       $ncols = substr($fmt,0,1);
81       printf("%s...\n", $fmt);
82       @width_str = split(/\./, $fmt_str);
83       $width = $width_str[0];
84       $width =  substr($fmt,2,4);
85       printf("FMT = %s, width = %d, num columns = %d\n", $fmt,$width,$ncols);
86
87       # Advance the array to the data
88       shift @a;
89
90       $numSteps = 24/$deltat;
```

```
91      printf("Number of time steps = %d\n",$numSteps);
92
93      $PROCESS_TS = $TRUE;
94      $NO_SKIP = $FALSE;
95    }
96
97    #-----------------------------------------------
98    # Process TS definitions, ignoring all other lines
99    #-----------------------------------------------
100   elsif ($PROCESS_TS == $TRUE){
101
102     #select STDOUT;
103     #printf("Inside PROCESS TS...\n");
104
105     chomp $one;
106     ($id,$dateTime,$hr,$val[$k]) = split(/ +/, $one);
107
108     $strlen_hr = length($hr);
109     if($strlen_hr == 1) {
110       $hour = "0".$hr;
111     }
112     else {
113       $hour = $hr;
114     }
115
116     $day = substr($dateTime,0,2);
117     $mo  = substr($dateTime,2,2);
118     $yr  = substr($dateTime,4,2);
119
120     #select STDOUT;
121     #printf("%s\n",$val[$k]);
122
123     $strlen_yr = length($yr);
124
125     #-----------------------------------------------
126     # "Y2K" kluge
127     #-----------------------------------------------
128     if($strlen_yr == 1) {
129       $year = "200".$yr;
130     }
131     elsif($yr < 50) {
132       $year = "20".$yr;
133     }
134     else {
135       $year = "19".$yr;
136     }
137     #-----------------------------------------------
138
139     if($ymd) {
140       $date_time_str = sprintf("%s-%s-%s %s:00:00 UTC",$year,$mo,$day,$hour);
141       $date_str = sprintf("%s-%s-%s",$year,$mo,$day);
142
143       if($hr==24) {
144         $next_day = `date --date="next day $date_str" +"%Y-%m-%d"`;
145         chomp($next_day);
146         $date_time_str = $next_day." 00:00:00 UTC";
147       }
148     }
149     else {
150       $date_time_str = sprintf("%s/%s/%s %s:00:00",$mo,$day,$year,$hour);
151       $date_str = sprintf("%s/%s/%s",$mo,$day,$year);
152
153       if($hr==24) {
154         $next_day = `date --date="next day $date_str" +"%m/%d/%Y"`;
```

```
155           chomp($next_day);
156           $date_time_str = $next_day." 00:00:00";
157         }
158       }
159       # 1.0 m^3 = 35.3147 ft^3
160       $convert = 35.3147;
161
162       # Set the field separator...
163       if($tabs){
164         $sep = "\t";
165         $sep_str = "TAB";
166       }
167       else {
168         $sep = "|";
169         $sep_str = "|";
170       }
171
172       #select STDOUT;
173       #printf("SEPARATOR %s\n", $sep_str);
174
175       select outputFH;
176       if($val[$k]>0) {
177         if($cms){
178           printf("%s%s%.2f\n",$date_time_str,$sep,$val[$k]);
179         }
180         else {
181           $value = $val[$k]*$convert;
182           printf("%s%s%.2f\n",$date_time_str,$sep,$value);
183
184         }
185       }
186       $k++;
187     }
188   }
189   #-------------------------------------------
190   # End processing of TS data
191   #-------------------------------------------
192   else {
193     select STDOUT;
194     print $one;
195   }
196 $lines++;
197 close(outFH);
198 }
```

Listing B.27: rdhmOutletQ2ts.pl Perl script that reformats RDHM discharge simulation output files to ascii time-series format.

```
1  #!/usr/bin/perl
2
3  use Getopt::Long;
4
5  #-------------------------------------------------------------------------
6  # makeQPFTS_RDHMcontrol.pl
7  #-------------------------------------------------------------------------
8  # USAGE: (on one line)
9  #
10 #-------------------------------------------------------------------------
11
12 $TRUE  = 1;
13 $FALSE = 0;
14
```

```
#-------------------------------------------------------------
# Read command line arguments.test
#-------------------------------------------------------------
$start_date = $ARGV[0];   # 20160315
$end_date = $ARGV[1];   # 20160322
$input_path = $ARGV[2];   # /home/teaiii/Desktop/GMEFP_schaake_shuffle/2010/0313/09
$output_path = $ARGV[3];  # /home/teaiii/Desktop/GMEFP_schaake_shuffle/output/20100313/09

#-------------------------------------------------------------
# Open file for output
#-------------------------------------------------------------
$outputDir = "/home/teaiii/RDHM/greenbrier/input/";
printf("Output directory: %s\n", $outputDir);
$outFileName = $outputDir."rdhm_qpfts_control";
open (outFH, ">$outFileName");

#-------------------------------------------------------------
# Write out the RDHM control File
#-------------------------------------------------------------
select(outFH);

printf("#simulation time period\n");
printf("time-period                          = %sT06 %sT06\n", $start_date, $end_date);
printf("\n");
printf("ignore-1d-xmrg = true\n");
printf("\n");
printf("#\n");
printf("#simulation time step in the format of HH:MM:SS.XXXX\n");
printf("time-step                    =  1\n");
printf("#\n");
printf("#OHRFC connectivity file\n");
printf("connectivity = ./ohrfc_adj_bsn.con\n");
printf("\n");
printf("#\n");
printf("pixel-size-hrap = 1.00\n");
printf("#\n");
printf("output-path        = %s\n",$output_path);
printf("#\n");
printf("#input paths\n");
printf("#\n");
printf("input-path         = ../parameters\n");
printf("input-path         = %s\n", $input_path);
printf("\n");
printf("input-path         = /media/teaiii/Seagate8/nldas_conus_xmrg/temps\n");
printf("\n");
printf("input-path         = /media/teaiii/development1/RDHM/warm_states\n");
printf("\n");
printf("#------------------------#-----------------------------------------------------------\
    n");
printf("# 6-hourly uniform xmrg disaggregation\n");
printf("#-------------------------------------------------------------\n");
printf("#disaggregation = xmrg interpolation=uniform freq=6:00:00 start=0:00:00\n");
printf("\n");
printf("#select operations\n");
printf("#available snow17, snow17_mcp3, sac, frz, api, rutpix7, rutpix9, musk\n");
printf("operations = snow17 sac frz rutpix9\n");
printf("\n");
printf("\n");
printf("#-------------------------------------------------------------\n");
printf("#-----------------Information for SAC-HT----------------------\n");
printf("DSINT  = 0.10 0.30 0.60 1.00 1.50\n");
printf("DSINTW = 0.10 0.30 0.60 1.00 1.50\n");
printf("\n");
printf("normalize-soil-moisture = true\n");
```

```
78 printf("\n");
79 printf("#-------------------------------------------------------------\n");
80 printf("#output-grid-after-timeloop          =\n");
81 printf("#-------------------------------------------------------------\n");
82 printf("output-grid-inside-timeloop   = tsint1 tsint2 tsint3 tsint4 tsint5\n");
83 printf("output-grid-inside-timeloop   = discharge\n");
84 printf("output-grid-inside-timeloop   = areac areac1 areac2 areac3 areac4 depth\n");
85 printf("output-grid-inside-timeloop   = swint1 swint2 swint3 swint4 swint5\n");
86 printf("output-grid-inside-timeloop   = smc0 smc1 smc2 smc3 smc4 smc5\n");
87 printf("output-grid-inside-timeloop   = sh2o0 sh2o1 sh2o2 sh2o3 sh2o4 sh2o5\n");
88 printf("output-grid-inside-timeloop   = uztwc uzfwc lztwc lzfsc lzfpc adimpc\n");
89 printf("output-grid-inside-timeloop   = upperSwint real_uztwc real_lztwc\n");
90 printf("output-grid-inside-timeloop   = surfaceFlow subsurfaceFlow\n");
91 printf("output-grid-inside-timeloop   = we neghs liqw tindex accmax sndpt sntmp cover\n");
92 printf("\n");
93 printf("#\n");
94 printf("#In number of timestep, example, 2 means output every 2 timestep\n");
95 printf("#\n");
96 printf("output-grid-step                    = 1\n");
97 printf("\n");
98 printf("#\n");
99 printf("#data to be output at the last 23hr, such as states\n");
100 printf("output-grid-last-step        = uztwc uzfwc lztwc lzfsc lzfpc adimpc \n");
101 printf("output-grid-after-timeloop   = discharge surfaceFlow subsurfaceFlow\n");
102 printf("\n");
103 printf("output-grid-last-step        = swint1 swint2 swint3 swint4 swint5\n");
104 printf("output-grid-last-step        = we neghs liqw tindex accmax sndpt sntmp cover\n");
105 printf("output-grid-last-step        = uztwc uzfwc lztwc lzfsc lzfpc adimpc\n");
106 printf("\n");
107 printf("# VK ---- you need to output these states too -------------------------\n");
108 printf("output-grid-last-step        = uztwh uzfwh lztwh lzfsh lzfph  \n");
109 printf("output-grid-last-step        = smc0 smc1 smc2 smc3 smc4 smc5\n");
110 printf("output-grid-last-step        = sh2o0 sh2o1 sh2o2 sh2o3 sh2o4 sh2o5\n");
111 printf("output-grid-last-step        = ts0 ts1 ts2 ts3 ts4\n");
112 printf("output-grid-last-step        = uztwc_prv uzfwc_prv lztwc_prv lzfsc_prv \n");
113 printf("output-grid-last-step        = lzfpc_prv adimpc_prv\n");
114 printf("# VK ---------------------------------------------------------------\n");
115 printf("\n");
116 printf("#\n");
117 printf("#Time series to be averaged over every basin\n");
118 printf("#\n");
119 printf("output-timeseries-basin-average      = xmrg\n");
120 printf("#output-timeseries-basin-average      = surfaceFlow\n");
121 printf("\n");
122 printf("#Time series at the outlet\n");
123 printf("output-timeseries-basin-outlet        = discharge\n");
124 printf("\n");
125 printf("#basin id and factors for input data\n");
126 printf("#------------------\n");
127 printf("#----- 03182700 ------\n");
128 printf("#------------------\n");
129 printf("input-data = 03182700\n");
130 printf("#=========================================================\n");
131 printf("# calibration adjustments\n");
132 printf("#=========================================================\n");
133 printf("#SAC Parameters, includes EFC/PCTIM\n");
134 printf("input-data = sac_UZTWM=-0.40\n");
135 printf("input-data = sac_UZFWM=-0.60\n");
136 printf("input-data = sac_UZK=-1.60\n");
137 printf("#input-data = sac_RIVA=0.04\n");
138 printf("#input-data = sac_ZPERC=-1.00\n");
139 printf("#input-data = sac_REXP=-1.0\n");
140 printf("#input-data = sac_LZTWM=-0.776\n");
141 printf("#input-data = sac_LZPK=-0.273\n");
```

```
printf("input-data = sac_LZFPM=-1.00\n");
printf("#input-data = sac_PFREE=-1.00\n");
printf("input-data = sac_LZSK=-01.00\n");
printf("#input-data = sac_LZFSM=-2.04\n");
printf("#input-data = sac_RSERV=0.3\n");
printf("#input-data = sac_PCTIM=-0.513\n");
printf("#input-data = sac_ADIMP=0\n");
printf("#input-data = sac_SIDE=0\n");
printf("#rutpix states\n");
printf("#input-data = rutpix_Q0CHN=-1.2\n");
printf("input-data = rutpix_QMCHN=-0.94\n");
printf("#=========================================================\n");
printf("#SAC parameters\n");
printf("input-data =      sac_ADIMP=0.0 sac_RIVA=0.001 sac_EFC=0.5\n");
printf("input-data =      sac_SIDE=0.0  sac_RSERV=0.3\n");
printf("input-data =      surf_water=-9999\n");
printf("\n");
printf("#rutpix states\n");
printf("#input-data =      rutpix_QMCHN=1.35\n");
printf("#snow states\n");
printf("input-data =      snow_PXTMP=1.0 snow_PLWHC=0.05 snow_TIPM=0.1 snow_SCF=1.0\n");
printf("input-data =      snow_NMF=0.15 snow_MBASE=0.0 snow_PGM=0.0 snow_LAEC=0.0\n");
printf("\n");
printf("input-data =      snow_SI=0.0 snow_ADC1=0.05 snow_ADC2=0.15 snow_ADC3=0.29\n");
printf("input-data =      snow_ADC4=0.41 snow_ADC5=0.51 snow_ADC6=0.60 snow_ADC7=0.65\n");
printf("input-data =      snow_ADC8=0.68 snow_ADC9=0.72 snow_ADC10=0.76 snow_ADC11=1.0\n");
printf("\n");
printf("input-data =      frz_RSMAX=0.58 frz_CKSL=8 frz_ZBOT=2.5\n");
printf("\n");
printf("input-data =      psfrac=9\n");
printf("input-data =      snow_RDCO=1\n");
printf("#---------------------------------\n");
printf("input-data = 03184000\n");
printf("#---------------------------------\n");
printf("#=========================================================\n");
printf("# calibration adjustments\n");
printf("#=========================================================\n");
printf("#SAC Parameters, includes EFC/PCTIM\n");
printf("input-data = sac_UZTWM=-0.40\n");
printf("input-data = sac_UZFWM=-0.60\n");
printf("input-data = sac_UZK=-1.60\n");
printf("#input-data = sac_RIVA=0.04\n");
printf("#input-data = sac_ZPERC=-1.00\n");
printf("#input-data = sac_REXP=-1.0\n");
printf("#input-data = sac_LZTWM=-0.776\n");
printf("#input-data = sac_LZPK=-0.273\n");
printf("input-data = sac_LZFPM=-1.00\n");
printf("#input-data = sac_PFREE=-1.00\n");
printf("input-data = sac_LZSK=-01.00\n");
printf("#input-data = sac_LZFSM=-2.04\n");
printf("#input-data = sac_RSERV=0.3\n");
printf("#input-data = sac_PCTIM=-0.513\n");
printf("#input-data = sac_ADIMP=0\n");
printf("#input-data = sac_SIDE=0\n");
printf("#rutpix states\n");
printf("#input-data = rutpix_Q0CHN=-1.2\n");
printf("input-data = rutpix_QMCHN=-0.94\n");
printf("#=========================================================\n");
printf("#SAC parameters\n");
printf("input-data =      sac_ADIMP=0.0 sac_RIVA=0.001 sac_EFC=0.5\n");
printf("input-data =      sac_SIDE=0.0  sac_RSERV=0.3\n");
printf("input-data =      surf_water=-9999\n");
printf("\n");
printf("#rutpix states\n");
```

```
printf("#input-data =      rutpix_QMCHN=1.35\n");
printf("#snow states\n");
printf("input-data =      snow_PXTMP=1.0 snow_PLWHC=0.05 snow_TIPM=0.1 snow_SCF=1.0\n");
printf("input-data =      snow_NMF=0.15 snow_MBASE=0.0 snow_PGM=0.0 snow_LAEC=0.0\n");
printf("\n");
printf("input-data =      snow_SI=0.0 snow_ADC1=0.05 snow_ADC2=0.15 snow_ADC3=0.29\n");
printf("input-data =      snow_ADC4=0.41 snow_ADC5=0.51 snow_ADC6=0.60 snow_ADC7=0.65\n");
printf("input-data =      snow_ADC8=0.68 snow_ADC9=0.72 snow_ADC10=0.76 snow_ADC11=1.0\n");
printf("\n");
printf("input-data =      frz_RSMAX=0.58 frz_CKSL=8 frz_ZBOT=2.5\n");
printf("\n");
printf("input-data =      psfrac=9\n");
printf("input-data =      snow_RDCO=1\n");
printf("#---------------------------------\n");
printf("input-data = 03180500\n");
printf("#---------------------------------\n");
printf("#=========================================================\n");
printf("# calibration adjustments\n");
printf("#=========================================================\n");
printf("#SAC Parameters, includes EFC/PCTIM\n");
printf("input-data = sac_UZTWM=-0.40\n");
printf("input-data = sac_UZFWM=-0.60\n");
printf("input-data = sac_UZK=-1.60\n");
printf("#input-data = sac_RIVA=0.04\n");
printf("#input-data = sac_ZPERC=-1.00\n");
printf("#input-data = sac_REXP=-1.0\n");
printf("#input-data = sac_LZTWM=-0.776\n");
printf("#input-data = sac_LZPK=-0.273\n");
printf("input-data = sac_LZFPM=-1.00\n");
printf("#input-data = sac_PFREE=-1.00\n");
printf("input-data = sac_LZSK=-01.00\n");
printf("#input-data = sac_LZFSM=-2.04\n");
printf("#input-data = sac_RSERV=0.3\n");
printf("#input-data = sac_PCTIM=-0.513\n");
printf("#input-data = sac_ADIMP=0\n");
printf("#input-data = sac_SIDE=0\n");
printf("#rutpix states\n");
printf("#input-data = rutpix_Q0CHN=-1.2\n");
printf("input-data = rutpix_QMCHN=-0.94\n");
printf("#=========================================================\n");
printf("#SAC parameters\n");
printf("input-data =      sac_ADIMP=0.0 sac_RIVA=0.001 sac_EFC=0.5\n");
printf("input-data =      sac_SIDE=0.0  sac_RSERV=0.3\n");
printf("input-data =      surf_water=-9999\n");
printf("\n");
printf("#rutpix states\n");
printf("#input-data =      rutpix_QMCHN=1.35\n");
printf("#snow states\n");
printf("input-data =      snow_PXTMP=1.0 snow_PLWHC=0.05 snow_TIPM=0.1 snow_SCF=1.0\n");
printf("input-data =      snow_NMF=0.15 snow_MBASE=0.0 snow_PGM=0.0 snow_LAEC=0.0\n");
printf("\n");
printf("input-data =      snow_SI=0.0 snow_ADC1=0.05 snow_ADC2=0.15 snow_ADC3=0.29\n");
printf("input-data =      snow_ADC4=0.41 snow_ADC5=0.51 snow_ADC6=0.60 snow_ADC7=0.65\n");
printf("input-data =      snow_ADC8=0.68 snow_ADC9=0.72 snow_ADC10=0.76 snow_ADC11=1.0\n");
printf("\n");
printf("input-data =      frz_RSMAX=0.58 frz_CKSL=8 frz_ZBOT=2.5\n");
printf("\n");
printf("input-data =      psfrac=9\n");
printf("input-data =      snow_RDCO=1\n");
printf("#---------------------------------\n");
printf("input-data = 03182500\n");
printf("#---------------------------------\n");
printf("#=========================================================\n");
printf("# calibration adjustments\n");
```

```
printf("#=========================================================\n");
printf("#SAC Parameters, includes EFC/PCTIM\n");
printf("input-data = sac_UZTWM=-0.40\n");
printf("input-data = sac_UZFWM=-0.60\n");
printf("input-data = sac_UZK=-1.60\n");
printf("#input-data = sac_RIVA=0.04\n");
printf("#input-data = sac_ZPERC=-1.00\n");
printf("#input-data = sac_REXP=-1.0\n");
printf("#input-data = sac_LZTWM=-0.776\n");
printf("#input-data = sac_LZPK=-0.273\n");
printf("input-data = sac_LZFPM=-1.00\n");
printf("#input-data = sac_PFREE=-1.00\n");
printf("input-data = sac_LZSK=-01.00\n");
printf("#input-data = sac_LZFSM=-2.04\n");
printf("#input-data = sac_RSERV=0.3\n");
printf("#input-data = sac_PCTIM=-0.513\n");
printf("#input-data = sac_ADIMP=0\n");
printf("#input-data = sac_SIDE=0\n");
printf("#rutpix states\n");
printf("#input-data = rutpix_Q0CHN=-1.2\n");
printf("input-data = rutpix_QMCHN=-0.94\n");
printf("#=========================================================\n");
printf("#SAC parameters\n");
printf("input-data =      sac_ADIMP=0.0 sac_RIVA=0.001 sac_EFC=0.5\n");
printf("input-data =      sac_SIDE=0.0  sac_RSERV=0.3\n");
printf("input-data =      surf_water=-9999\n");
printf("\n");
printf("#rutpix states\n");
printf("#input-data =      rutpix_QMCHN=1.35\n");
printf("#snow states\n");
printf("input-data =      snow_PXTMP=1.0 snow_PLWHC=0.05 snow_TIPM=0.1 snow_SCF=1.0\n");
printf("input-data =      snow_NMF=0.15 snow_MBASE=0.0 snow_PGM=0.0 snow_LAEC=0.0\n");
printf("\n");
printf("input-data =      snow_SI=0.0 snow_ADC1=0.05 snow_ADC2=0.15 snow_ADC3=0.29\n");
printf("input-data =      snow_ADC4=0.41 snow_ADC5=0.51 snow_ADC6=0.60 snow_ADC7=0.65\n");
printf("input-data =      snow_ADC8=0.68 snow_ADC9=0.72 snow_ADC10=0.76 snow_ADC11=1.0\n");
printf("\n");
printf("input-data =      frz_RSMAX=0.58 frz_CKSL=8 frz_ZBOT=2.5\n");
printf("\n");
printf("input-data =      psfrac=9\n");
printf("input-data =      snow_RDCO=1\n");
printf("#---------------------------------\n");
printf("input-data = 03183500\n");
printf("#---------------------------------\n");
printf("#=========================================================\n");
printf("# calibration adjustments\n");
printf("#=========================================================\n");
printf("#SAC Parameters, includes EFC/PCTIM\n");
printf("input-data = sac_UZTWM=-0.40\n");
printf("input-data = sac_UZFWM=-0.60\n");
printf("input-data = sac_UZK=-1.60\n");
printf("#input-data = sac_RIVA=0.04\n");
printf("#input-data = sac_ZPERC=-1.00\n");
printf("#input-data = sac_REXP=-1.0\n");
printf("#input-data = sac_LZTWM=-0.776\n");
printf("#input-data = sac_LZPK=-0.273\n");
printf("input-data = sac_LZFPM=-1.00\n");
printf("#input-data = sac_PFREE=-1.00\n");
printf("input-data = sac_LZSK=-01.00\n");
printf("#input-data = sac_LZFSM=-2.04\n");
printf("#input-data = sac_RSERV=0.3\n");
printf("#input-data = sac_PCTIM=-0.513\n");
printf("#input-data = sac_ADIMP=0\n");
printf("#input-data = sac_SIDE=0\n");
```

```
printf("#rutpix states\n");
printf("#input-data = rutpix_Q0CHN=-1.2\n");
printf("input-data = rutpix_QMCHN=-0.94\n");
printf("#=========================================================\n");
printf("#SAC parameters\n");
printf("input-data =      sac_ADIMP=0.0 sac_RIVA=0.001 sac_EFC=0.5\n");
printf("input-data =      sac_SIDE=0.0  sac_RSERV=0.3\n");
printf("input-data =      surf_water=-9999\n");
printf("\n");
printf("#rutpix states\n");
printf("#input-data =      rutpix_QMCHN=1.35\n");
printf("#snow states\n");
printf("input-data =      snow_PXTMP=1.0 snow_PLWHC=0.05 snow_TIPM=0.1 snow_SCF=1.0\n");
printf("input-data =      snow_NMF=0.15 snow_MBASE=0.0 snow_PGM=0.0 snow_LAEC=0.0\n");
printf("\n");
printf("input-data =      snow_SI=0.0 snow_ADC1=0.05 snow_ADC2=0.15 snow_ADC3=0.29\n");
printf("input-data =      snow_ADC4=0.41 snow_ADC5=0.51 snow_ADC6=0.60 snow_ADC7=0.65\n");
printf("input-data =      snow_ADC8=0.68 snow_ADC9=0.72 snow_ADC10=0.76 snow_ADC11=1.0\n");
printf("\n");
printf("input-data =      frz_RSMAX=0.58 frz_CKSL=8 frz_ZBOT=2.5\n");
printf("\n");
printf("input-data =      psfrac=9\n");
printf("input-data =      snow_RDCO=1\n");
printf("#---------------------------------\n");

close(outFH);
```

Listing B.28: makeQPFTS_RDHMcontrol.pl Perl script used to generate RDHM simulation *control* files for RDHM *monte carlo* experiments called by run_QPF_TS_RDHM

```
#!/usr/bin/perl

use Getopt::Long;

##############################################################################
# rdhmOutletQ2tsCats.pl
#
##############################################################################
my $TRUE  = 1;
my $FALSE = 0;
my $cms   = $FALSE;
my $tabs  = $TRUE;

my ($ymd,$verbose,$cms,$tabs);
GetOptions('ymd!' => \$ymd,
     'cms!'   => \$cms,
     'tabs!'  => \$tabs,
          'verbose!' => \$verbose );

$inFileName1 = $ARGV[0];
$inFileName2 = $ARGV[1];
$cat = $ARGV[2];    # Category from GRASS GIS analysis, used to set factor
print "Processing file ",$inFileName1,"\n"; # Data file
print "Processing file ",$inFileName2,"\n"; # File used to set factor from Category
# Open TS file
open (FH, $inFileName1);

@a = <FH>;
close(FH);

open (FH2, $inFileName2);
@b = <FH2>;
```

```
close(FH2);

$num_lines = scalar(@b);
printf("Number of lines in Random Points file: %s\n", $num_lines);
for($i=0;$i<$num_lines;$i++) {
      chomp($b[$i]);
      ($which_cat,$x,$y,$num,$ts) = split(/\,/, $b[$i]);
      chomp($b[$i]);
      select STDOUT;
      #printf("%s %s %s %s %s\n",$which_cat,$x,$y,$num,$ts);

      if($cat == $which_cat){
        if($ts > 0.30){
          $factor = "A";
        }
        elsif ($ts <= 0.30 && $ts >0.25){
          $factor = "B";
        }
        elsif ($ts <= 0.25 && $ts >0.15){
          $factor = "C";
        }
        elsif ($ts <= 0.15 && $ts >0.06){
          $factor = "D";
        }
        else{
          $factor = "E";
        }
        last;
      }
}

printf("Factor is %s %s\n", $factor, $cat);

$outputFileName = $inFileName1.".out";
open(outputFH, ">$outputFileName");

$outputFileName2 = "/home/teaiii/TS_maxval.out";
open(outFH, ">>$outputFileName2");

#-----------------------------------------------
# Keyword definitions
#-----------------------------------------------
$datacard = "DATACARD";
#-----------------------------------------------

$present_year = `date +%Y`;
chomp($present_year);
$next_year  = `date +'%Y' -d 'next year'`;
chomp($next_year);
$today  = `date +%Y%m%d`;
chomp($today);

printf("Present Year = %s, Next Year = %s, Today = %s\n",$present_year,$next_year,$today);

select STDOUT;
$numLines = scalar(@a);
printf("Number of lines = %d\n", $numLines);

$i=0; # Counter for each line read
$k=0; # Counter for each data value read
$lines = 0; # Total lines in file
$total_vals = 0;
$period = 1;
$FIRST_TIME = $TRUE;
```

```
$NO_SKIP = $TRUE;
$LAST_DATE = $FALSE;
$PROCESS_TS = $FALSE;

#-------------------------------------------
# Loop through each line of file - main loop
#-------------------------------------------
foreach $one (@a) {

  #-------------------------------------------
  # Test line for beginning of TS data
  # Set Flag to process data only
  #-------------------------------------------
  if (index($one, "\$") != 0) {
    if($NO_SKIP == $TRUE) {

      chomp($one);
      printf("%s\n",$one);
      #RDHM OUTPUTS  SQIN L3/T CMS  1    CRCO1
      ($model,$output_str,$ts_type,$dimensions,$units,$deltat,$id) = split(/ +/, $one);
      #01  2011 10   2011  1   F13.4
      ($start_mo_str,$start_yr,$end_mo_str,$end_yr,$num_cols,$fmt) = split(/ +/, $a[$lines+1]);

      chomp($fmt);
      $ncols = substr($fmt,0,1);
      printf("%s...\n", $fmt);
      @width_str = split(/\./, $fmt_str);
      $width = $width_str[0];
      $width =  substr($fmt,2,4);
      printf("FMT = %s, width = %d, num columns = %d\n", $fmt,$width,$ncols);

      # Advance the array to the data
      shift @a;

      $numSteps = 24/$deltat;
      printf("Number of time steps = %d\n",$numSteps);

      $PROCESS_TS = $TRUE;
      $NO_SKIP = $FALSE;
    }

    #-------------------------------------------
    # Process TS definitions, ignoring all other lines
    #-------------------------------------------
    elsif ($PROCESS_TS == $TRUE){

      #select STDOUT;
      #printf("Inside PROCESS TS...\n");

      chomp $one;
      ($id,$dateTime,$hr,$val[$k]) = split(/ +/, $one);

      $strlen_hr = length($hr);
      if($strlen_hr == 1) {
        $hour = "0".$hr;
      }
      else {
        $hour = $hr;
      }

      $day = substr($dateTime,0,2);
      $mo  = substr($dateTime,2,2);
      $yr  = substr($dateTime,4,2);

```

```
161       #select STDOUT;
162       #printf("%s\n",$val[$k]);
163
164       $strlen_yr = length($yr);
165
166       #-------------------------------------------
167       # "Y2K" kluge
168       #-------------------------------------------
169       if($strlen_yr == 1) {
170         $year = "200".$yr;
171       }
172       elsif($yr < 50) {
173         $year = "20".$yr;
174       }
175       else {
176         $year = "19".$yr;
177       }
178       #-------------------------------------------
179
180       if($ymd) {
181         $date_time_str = sprintf("%s-%s-%s %s:00:00 UTC",$year,$mo,$day,$hour);
182         $date_str = sprintf("%s-%s-%s",$year,$mo,$day);
183
184         if($hr==24) {
185           $next_day = `date --date="next day $date_str" +"%Y-%m-%d"`;
186           chomp($next_day);
187           $date_time_str = $next_day." 00:00:00 UTC";
188         }
189       }
190       else {
191         $date_time_str = sprintf("%s/%s/%s %s:00:00",$mo,$day,$year,$hour);
192         $date_str = sprintf("%s/%s/%s",$mo,$day,$year);
193
194         if($hr==24) {
195           $next_day = `date --date="next day $date_str" +"%m/%d/%Y"`;
196           chomp($next_day);
197           $date_time_str = $next_day." 00:00:00";
198         }
199       }
200       # 1.0 m^3 = 35.3147 ft^3
201       $convert = 35.3147;
202
203       # Set the field separator...
204       if($tabs){
205         $sep = "\t";
206         $sep_str = "TAB";
207       }
208       else {
209         $sep = "|";
210         $sep_str = "|";
211       }
212
213       #select STDOUT;
214       #printf("SEPARATOR %s\n", $sep_str);
215
216       if($val[$k]>0) {
217
218         if($k == 1){
219           $maxal = $val[$k];
220         }
221         else {
222           if($val[$k] > $maxal){
223             $maxal = $val[$k];
224           }
```

```
225         }
226
227         select outputFH;
228         if($cms){
229           printf("%s%s%.2f%s%s\n",$date_time_str,$sep,$val[$k],$sep,$factor);
230         }
231         else {
232           $value = $val[$k]*$convert;
233           printf("%s%s%.2f%s%s\n",$date_time_str,$sep,$value,$sep,$factor);
234
235         }
236       }
237       $k++;
238     }
239   }
240   #-------------------------------------------
241   # End processing of TS data
242   #-------------------------------------------
243   else {
244     select STDOUT;
245     print $one;
246   }
247 $lines++;
248 #close(outputFH);
249 }
250
251 open (FH3, "/media/teaiii/development1/grass/Linear_dist_TS.txt");
252 @dist_array = <FH3>;
253 close(FH3);
254
255 $numDlines = scalar(@dist_array);
256 printf("Number of lines in cat-distance file: %s\n", $numDlines);
257 for($j=0;$j<$numDlines;$j++) {
258       chomp($dist_array[$j]);
259       ($from_cat,$to_cat,$distance[$j]) = split(/\|/, $dist_array[$j]);
260       chomp($distance[$j]);
261 }
262
263 select outFH;
264 if($cms){
265   printf("%s\t%s\t%s\t%s\t%s\n", $cat,$ts,$maxal,$factor,$distance[$cat-1]);
266 }
267 else{
268   $value = $maxval*$convert;
269   printf("%s\t%s\t%s\t%s\t%s\n", $cat,$ts,$value,$factor,$distance[$cat-1]);
270 }
271 #close(outFH);
```

Listing B.29: rdhmOutletQ2tsCats.pl Perl script used in RDHM *monte carlo* experiments for *threat score* category calculations called by run_RDHMthreatScoreOutput2TS.sh.