IMPROVING ANALYSIS OF HEAVY TO EXTREME PRECIPITATION USING CONDITIONAL BIAS-PENALIZED OPTIMAL ESTIMATION

by

RIDWAN SIDDIQUE

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF CIVIL ENGINEERING

THE UNIVERSITY OF TEXAS AT ARLINGTON

May 2013

Acknowledgements

First and foremost, I would like to remember the Almighty and express my gratitude to Him for giving me an opportunity and enough patience and perseverance to finish this work.

I would like to express my sincere gratitude to my supervising professor Dr. DJ Seo for his constant guidance and support to accomplish this work. It was highly motivating to work with him and I consider myself extremely lucky to get the opportunity. This study would not have been possible without his valuable advices and constant feedback.

I am extremely thankful to Dr. Yu Zhang of National Weather Service (NWS) and Dr. Dongsoo Kim of National Climatic Data Center (NCDC) for their valuable contributions to this work. All their suggestions and advice played a big part in this work.

I would like to convey my sincere appreciation to Dr. Xinbao Yu, Dr. Hayek Choi and Dr. Yu Zhang for serving on my thesis committee. I am grateful to them for their valuable time and suggestions.

I have always enjoyed working with my lab mates as a research team. Throughout my MS study, they have always been very kind and supportive to me as a friend and colleague. My special thanks goes to Arezoo Rafieenasab, Manabendra Saharia, Seo Young Kim, Hamideh Riazi, Sanjib Sharma, Dr. Sunghee Kim, Amir Norouzi and Dr. Changmin Shin for all their love and support.

My parents have always been an important source of encouragement. This work would not have been possible without their unconditional support.

April 22, 2013

Abstract

IMPROVING ANALYSIS OF HEAVY TO EXTREME PRECIPITATION

USING CONDITIONAL BIAS-PENALIZED

OPTIMAL ESTIMATION

Ridwan Siddique, MS

The University of Texas at Arlington, 2013

Supervising Professor: Dong-Jun Seo

Precipitation estimation is a very important topic from the societal perspective as heavy precipitation can cause flooding from which loss of lives and damage to properties can occur. In current practice, a number of spatial interpolation techniques are used for precipitation estimation using rain gauge data. Most of them are based on minimizing error variance but none of them consider Type-II conditional bias. As such, the existing techniques work well in the mid ranges of the precipitation distribution but tend to under- and overestimate large and small precipitation amounts, respectively. Conditional bias-penalized kriging (CBPK) adds a penalty term for Type-II conditional bias in addition to the error variance to improve estimation of large precipitation amounts. CBPK, however, tends to produce negative estimates in areas of very small or no precipitation. This problem is addressed in this work by an extension of CBPK, referred to as extended conditional bias-penalized kriging (ECPBK). For comparative evaluation, several real world experiments have been carried out using hourly rain gauge data. Also, synthetic experiments have been carried out for MAP analysis using the Stage IV data as truth and

by creating synthetic gauge networks within the Stage IV precipitation field. The cross validation results of ECBPK are compared with those of the Single Optimal Estimation technique used in the NWS's Multisensor Precipitation Estimator.

Table Of Contents

List of Illustrations

List of Table

Chapter 1

Introduction

1.1 Background

For its obvious importance, quantitative precipitation estimation (QPE) has been a topic of active research for over a century (Thiessen 1911). Whether it is based on gauge-only or multisensor estimation, QPE generally involves spatial prediction using statistical or dynamical-statistical models. Statistical models, by far the more widely used of the two to date, use optimal (in some sense of the word) estimation, of which various types of linear and nonlinear techniques are available (see e.g. Creutin and Obled (1982), Tabios and Salas (1985) and references therein). For example, the algorithms used operationally in the National Weather Service (NWS) for gauge-only and radar-gauge analyses in their Multisensor Precipitation Estimator (MPE, Seo et al. 2010) are variants of kriging and cokriging, respectively (Seo 1998a,b). More recently, artificial neural networks (Bellerby et al. 2000, Grimes et al. 2003, Hsu et al. 2007, Chiang et al. 2007) and support vector regression (Chen et al. 2011) have been added to the list as potential techniques for QPE.

Real-time QPE demands accurate estimation particularly of large amounts as they represent greater hazards to lives and properties. In flood forecasting, what matters most for QPE is the ability to estimate large amounts of precipitation as accurately as possible over the range of spatiotemporal scales of aggregation associated with the size and response time of the basin. Kriging or its variants do produce, as theoretically expected, precipitation estimates that are unbiased and of minimum error variance in the unconditional sense. In the conditional sense, however, these so-called optimal estimation techniques very often severely underestimate heavy precipitation and overestimate light precipitation (Seo and Breidenbach 2002, Ciach et al. 2000, Habib et al. 2012). These results arise because, to achieve (unconditional) minimum error variance, it is necessary to reduce the error variance associated with light to moderate precipitation, which occurs frequently and over large areas, even if it may

increase the error variance associated with heavy precipitation, which occurs relatively rarely and generally over small areas. For accurate estimation of large amounts, however, it is more important to reduce conditional bias (CB), in particular Type-II CB, than to minimize unconditional error variance. QPE for flood forecasting is a prime example of that. In the above, Type-II CB is defined as $E[\hat{X} \mid X = x] - x$ where X, $\hat{X}$ and x denote the unknown truth, the estimate, and the realization of X, respectively (Joliffe and Stephenson 2003).

Recently, Seo (2012) has demonstrated in the context of gauge-only QPE that the potential impact of reducing Type-II CB (hereafter referred to as CB for short) at sub-daily and mean areal precipitation (MAP) scales (O(10^0)~O(10^3) km^2) is substantial. The synthetic experiments suggest that the margin of improvement for estimating heavy precipitation from reducing CB may exceed that from greatly increasing the density of the rain gauge network or, equivalently in multisensor estimation, that from significantly improving the quality of the remotely sensed data or scale-compatible numerical weather prediction (NWP) precipitation analysis. Often, lack of performance by linear estimators has been attributed to their linear (as opposed to nonlinear) nature. Experimental and empirical evidences suggest, however, that the marginal improvement by nonlinear estimation is rather small (see e.g. Azimi-Zonooz et al. 1989, Seo 1996), and that CB may be a much more important limiting factor than linearity in estimation of heavy-to-extreme precipitation.

Seo (2012) extended classic optimal linear estimation theory in which, in addition to error variance, CB is explicitly minimized. The resulting Fisher-like solution may also serve as an alternative or complementary observation equation for a range of Fisher solution-based static or dynamic filters, such as Kalman filter and its variants. When cast in the form of well-known kriging or its variants used in the Multisensor Precipitation Estimator (MPE) (Seo et al. 2010) of the NWS, the proposed methodology yields a new kriging estimator, referred to as CB-penalized kriging (CBPK). CBPK, however, yields estimates that may be significantly negative in areas of light precipitation. To address this, an extension of CBPK, referred to herein as

extended CB-penalized kriging (ECBPK), has been developed. In this work, ECBPK is described and comparatively evaluated with ordinary kriging (OK) (Journel and Huijbregts 1978), a variant of which is used in MPE for gauge-only precipitation estimation. The evaluation is carried out for estimation of point and MAP through real-world and synthetic experiments, respectively, for a number of heavy-to-extreme precipitation events in Oklahoma, the Southeast and Texas.

1.2 Objective

The main objective of this study is to improve the estimation procedure for heavy to extreme precipitation. Conventional estimation techniques including currently popular OK do not consider Type-II CB and focus primarily on minimizing error variance. As a consequence, these techniques tend to work better in estimating light to moderate precipitation than in estimating large precipitation amounts. CBPK which minimizes Type-II CB in addition to error variance shows substantial improvement in estimating heavy precipitation but produces inferior estimates than OK for light to moderate precipitation. This study aims at combining these two techniques to obtain best possible estimates in all ranges of the distribution for both point and MAP. The second goal of this study is to carry out a comparative evaluation between the newly proposed technique with OK using both synthetic experiments and real world case studies.

1.3 Outline of the Thesis

Chapter 1 presents the background of the study and the statement of the problem, specific objectives of the study and thesis organization.

Chapter 2 reviews the literature on the geostatistical and non-geostatistical spatial interpolation techniques, CB and key findings from relevant previous studies in precipitation estimation.

Chapter 3 provides mathematical representation and description of the proposed new technique ECBPK.

Chapter 4 discusses the data used for evaluation and also different quality control measures used in this study.

Chapter 5 describes the real world case studies and synthetic experiments carried out for evaluation of the newly proposed technique.

Chapter 6 presents the results of the real world and synthetic experiments.

Chapter 7 presents the conclusion and suggests recommendations for further improvement of the proposed technique.

Chapter 2

Literature Review

2.1 Spatial Interpolation Techniques

Many types of spatial interpolation techniques have been developed and used in different fields. One of the common problems in spatial interpolation is to estimate the variable of interest at an ungauged location or where the data is missing. Sometimes, these spatial interpolation techniques are also modified and tailored to serve specific purposes in different fields. In hydrology, precipitation estimation is an age-old problem for which numerous studies have been carried out in different times for improvement. Because precipitation is a highly-variable stochastic phenomenon, estimating precipitation at an ungauged location using available gauge data is a challenge. Many researchers have applied a wide range of different interpolation techniques with varying degrees of complexity to improve spatial estimation of precipitation. Many of these techniques may be categorized as geostatistical or non-geostatistical. Some of them are briefly discussed here to provide a background to the newly proposed technique.

2.1.1 Non-Geostatistical Methods

Non-geostatistical methods do not use any geostatistical information in the estimation procedure. The Thiessen polygon (Thiessen, 1911), isohyetal (Linsley et al. 1949), arithmetic mean (Paulhus and Kohler, 1952), normal ratio (Paulhus and Kohler, 1952) and inverse distance weight method (Wei and Mcguinness, 1973) can be categorized as non-geostatistical techniques. The Thiessen polygon method is not flexible for the users as a new Thiessen network needs to be constructed each time for any change in the gauge network although it can produce more accurate estimates than many other non-geostatistical methods mentioned above (Chow et al. 2004). The isohyetal method is flexible but requires a dense gauge network for accurate construction of the isohyets for complex storms (Chow et al. 2004). The arithmatic

mean and normal ratio methods are not effective if the gauge network is not uniformly distributed. All these non-geostatistical methods may give a comparable result for long-term accumulations (e.g. at annual scale) but their estimates may vary by a large margin when applied for a smaller time scale i.e. hourly or daily precipitation (Chow et al. 2004).

2.1.2 Geostatistical Method

Geostatistical approaches may be used to estimate the spatial patterns or mean areal quantifies of precipitation or any other random phenomena and to model the uncertainty associated with them.

Kriging is a popular optimal geostatistical estimation technique first introduced by Danie G. Krige (1951) and further developed by the French mathematician George Matheron (1963). Many researchers have utilized and compared various kriging techniques for precipitation estimation: e.g. Chua and Bras (1982); Cruetin and Obled (1982); Bastin et al.(1984); Tabios and Salas (1985); Seo et al.(1990); Barancourt and Cruetin (1992). Cruetin and Obled (1982) noted that kriging and objective analysis may give better estimates when the precipitation is highly variable. Bastin et al.(1984) presented a simple technique to estimate in real time mean areal precipitation (MAP) by calculating both seasonal and time-invariant variograms. Tabios and Salas (1985) compared the applicability of different spatial interpolation techniques for estimating annual precipitation at ungauged locations in the north-central United States. Their analysis included the Thiessen polygon method, polynomial interpolation, the inverse distance method, multiquadratic interpolation and kriging. The comparison was made based on different error statistics. It was found that kriging and optimal interpolators provided the best estimates among all techniques considered.

Numerous studies have been carried out for improving precipitation estimation using kriging. But very few works have been done regarding heavy precipitation estimation. Accurate estimation of heavy to extreme precipitation is very important in the sense that heavy

precipitation may produce large runoff and cause inundation of low lands and floods. By improving the analysis of heavy to extreme precipitation, streamflow prediction and flood forecasting can be improved significantly. In this work, a new kriging technique will be presented to improve the estimation of heavy precipitation by combining the estimates of two kriging techniques.

Kriging can be described as an interpolation method that linearly weights the observations. In general, the linear kriging estimators aim to achieve unbiasedness and minimum error variance. If linear, such an estimator is referred to as the best linear unbiased estimator (BLUE). Unbiasedness is achieved when the expected value of the estimate is equal to the expected value of the unknown truth. The weights are a function of the correlation structure of the precipitation field, the spatial geometry of the rain gauges used in the estimation and the point at which an estimate is desired.

There are many different types of kriging. Based on the mathematical form of the estimator, a kriging method can be described as linear or non-linear. Also, kriging can be used to estimate the variable of interest over a specific area or at a particular point. Some of the kriging methods are listed below:

1) Simple kriging
2) Ordinary kriging
3) Indicator kriging
4) Universal kriging
5) Block kriging
6) Disjunctive kriging

When multiple variables are involved, kriging is referred to as cokriging (Journel and Huijbregts, 1978). Among all these techniques, OK, simple kriging (SK) and indicator kriging (IK) are of particular relevance to this work. The new technique proposed in this work is an extension of conditional bias-penalized kriging (CBPK) (Seo, 2012) which is illustrated using

SK. Currently, OK is one of the most widely used kriging techniques, including precipitation estimation. As such, in this study, a comparative evaluation is carried out between OK and the newly proposed technique. Indicator variables have been used in this work to examine the spatial variability in intermittency of precipitation. IK may be used to estimate probabilistic quantiles such as the probability of precipitation (PoP) by evaluation the expectation of an indicator variable. These three kriging techniques are briefly discussed below:

2.1.1.1 Ordinary Kriging (OK)

OK is used when no priori information is available about the mean of the unknown truth. It is one of the most widely used kriging techniques. In OK, the weights given to the observations are obtained by minimizing the error variance under the constraint that the sum of the weights equal to unity. By applying this constraint, the unbiasedness of the OK estimates is ensured. The error variance (σ^2) is defined as:

$$\sigma^2 = Var[Z_0^* - Z_0] \tag{2.1}$$

where Z_0^* and Z_0 denote the estimate and the truth, respectively at location x_0 . The OK estimator can be expressed by the following equations:

$$Z_0^* = \sum_{i=1}^{n} \lambda_i Z_i \tag{2.3}$$

$$\sum_{i=1}^{n} \lambda_i = 1 \tag{2.4}$$

where Z_0^* denotes the estimate at the ungauged location x_0, Z_i denotes the observation at location x_i, *n* denotes the number of observations, and λ_i denotes the weight to the *ith* observation. In OK, to obtain the weights, one minimizes:

$$F = \sigma^2 + 2\mu \left(\sum_{i=1}^{n} \lambda_i - 1\right) \tag{2.5}$$

where σ^2 denotes the error variance (see Eq.(2.1)) and μ denotes the Lagrange multiplier (Isaak and Srivastava, 1989) which converts the constrained minimization problem to an unconstrained minimization problem.

2.1.1.2 Simple Kriging (SK)

In SK, it is assumed that the mean of the variable to be estimated is known. The weights assigned to the observations are obtained by minimizing the error variance (Eq.(2.1)). The SK estimator, which is equivalent to the Bayesian optimal linear estimator with perfect observations can be expressed as:

$$Z_0^* = m + \sum_{i=1}^{n} \lambda_i [Z_i - m] \tag{2.2}$$

where Z_0^* denotes the estimate at the ungauged location x_0, m denotes the known mean assumed constant over the entire domain in this example, Z_i denotes the observation at location x_i, *n* denotes the number of observations and λ_i denotes the weight given to the *i*th observation.

SK does not use an unbiasedness constraint as the information about mean is already known. In comparison to OK, SK assumes more information and hence, gives a smaller estimation variance than OK.

2.1.1.3 Indicator Kriging (IK)

IK is a non-linear kriging method. In IK, a set of thresholds for the variable of interest is chosen and each observation is rendered as a binary variable of 0 and 1, referred to as indicator variable, depending on whether it is above or below each threshold. The indicator variable can be expressed as:

$$I(x_0; z_k) = \begin{cases} 0, & if\ Z_0 \leq z_k \\ 1, & if\ Z_0 > z_k \end{cases} \tag{2.6}$$

where z_k denotes the fixed threshold and $I()$ denotes the indicator variable. IK gives an estimate of the conditional cumulative distribution function (ccdf) of the variable of interest at location x_0. The IK estimator is given by:

$$I(x_0; z_k)^* = E\big(I(x_{0,} z_k | p\big)$$

$$= Prob(Z_0 > z_k | p) \tag{2.7}$$

where $I(x_0; z_k)^*$ denotes the IK estimate at location x_0 at threshold z_k and p denotes the conditional events in the neighborhood of location x_0 .

2.2 Conditional Bias (CB)

Most of the conventional precipitation estimation techniques produce unbiased estimates having minimum error variance in the unconditional sense. But in the conditional (on the magnitude of the truth being estimated) sense, these techniques tend to underestimate heavy precipitation and overestimate light to moderate precipitation (Seo, 2012). There are two types of CB.

1) Type-I conditional bias (Type-I CB)
2) Type-II conditional bias (Type-II CB)

There is a widespread confusion about CB in the literature and among the practitioners of geostatistics (Mclennan and Duestch, 2003). Isaak (2004) stated that CB is poorly understood even though it is a well-recognized problem in geostatistics. Lack of distinction between Type-I and Type-II CB in the literature may be one of the main sources of such confusion (Seo, 2012).

2.2.1 Type-I CB

Type-I CB is defined as:

$$E[X|X^*] - X^*$$

where X^* and X denote the estimate and the truth, respectively. Type-I CB occurs when the estimate is biased against the expected value of the true precipitation conditional on the estimate.

2.2.2 Type-II CB

Type-II CB is defined as:

$$E[X^*|X] - X$$

Type-II CB exists when the expected value of the estimate given the truth differs from the truth.

This study is concerned with reducing Type-II CB. Different scientists and geostatisticians discuss the problem of Type-II CB in different contexts e.g. in those of mining and radar rainfall estimation.

Brown and Seo (2012) proposed a new non-parametric technique to minimize Type-II CB in streamflow prediction. This technique is analogous to indicator cokriging (ICK) and is called conditional bias-penalized indicator cokriging (CBP-ICK). It is found that CBP-ICK successfully reduce Type-II CB and produce estimates that are more skillful than the estimates from other post processors used in hydrologic prediction.

Seo (2012) proposed and described a new estimation technique, CBPK, which is an extension of SK. CBPK adds a penalty term for Type-II CB in addition to error variance. Seo (2012) evaluated CBPK using normal and lognormal synthetic experiments and found that CBPK successfully reduces Type-II CB for large precipitation amounts. Seo (2012) also described a Fisher-like solution of CBPK.

Chapter 3

Methodology

In this section, the proposed procedure, extended CBPK or ECBPK is described. It builds on CBPK (Seo, 2012) and uses as its theoretical basis a procedure that combines estimates from two different estimators.

3.1 Conditional Bias-Penalized Kriging

CBPK may be considered as an extension of SK in which the objective function is made of not only error variance but also CB. The SK estimator (see e.g. Journel and Huijbregts 1978) is given by:

$$Z_0^* = m_0 + \sum_{i=1}^{n} \lambda_i (Z_i - m_i) \tag{3.1}$$

where Z_0^* denotes the SK estimate for Z_0, m_0 denotes the mean of Z_0, λ_i denotes the weight assigned to Z_i, m_i denotes the mean of Z_i and n denotes the number of neighbors used in estimation. In SK, the weights are obtained by minimizing the error variance of the estimate, J_{SK}:

$$J_{SK} = E_{Z_0^*, Z_0}[(Z_0^* - Z_0)^2] = E_{Z_i, Z_0}[\{\sum_{i=1}^{n} \lambda_i (Z_i - m_i) - (Z_o - m_0)\}^2] \tag{3.2}$$

where the expectation operations are with respect to the variables subscripted. In CBPK, the CB penalty term, or the unconditional expectation of CB squared, is added to the objective function as follows:

$$\begin{aligned} J_{CBPK} &= E_{Z_0^*, Z_0}[(Z_0^* - Z_0)^2] + E_{Z_0}[\{E_{Z_0^*}[Z_0^* \mid Z_o] - Z_o\}^2] \\ &= E_{Z_i, Z_0}[\{\sum_{i=1}^{n} \lambda_i (Z_i - m_i) - (Z_o - m_0)\}^2] \\ &+ \int \{E_{Z_i}[\sum_{i=1}^{n} \lambda_i (Z_i - m_i) \mid Z_o = z_o] - (z_o - m_0)\}^2 f_{Z_o}(z_o) dz_0 \end{aligned} \tag{3.3}$$

where z_0 denotes the experimental values of Z_0 and $f_{Z_0}(z_0)$ denotes the marginal probability density function (pdf) of Z_0. The CBPK system results from minimizing J_{CBPK}:

$$\sum_{j=1}^{n} \lambda_j (\rho_{ij} + \rho_{i0}\rho_{j0}) \sigma_i \sigma_j = 2\rho_{i0}\sigma_i \sigma_0, \qquad i = 1, ..., n \tag{3.4}$$

where ρ_{ij} denotes the correlation between Z_i and Z_j , and σ_i denotes the standard deviation of Z_i. When minimizing J_{CBPK}, we may constrain the weights to sum to unity and arrive at the OK analogue of CBPK. Throughout this paper, by CBPK, we mean the OK analogue. The above formulation has also been used successfully in the form of CBP-ICK for post-processing of streamflow predictions (Brown and Seo, 2012). For nonnegative variables such as precipitation, however, the raw CBPK estimates may be significantly negative, particularly in areas of no to light precipitation. To address this, we formulate a procedure to combine estimates from two different estimators. Below, we describe the procedure in the context of combining OK and CBPK estimates which provides the theoretical basis for ECBPK.

3.2 Combining Estimates from Two Different Estimators

While CBPK is superior to OK over the tail ends of the distribution, it is inferior over the mid-ranges (Seo, 2012). One may combine the OK and CBPK estimates in such a way that the combined estimate is close to the more accurate of the two, depending on the magnitude of the precipitation amount being estimated. Toward that end, we write the estimate, $E[Z_0 \mid Z_1 = z_1,...,Z_n = z_n]$, as a combination of OK and CBPK estimates:

$$E[Z_0 \mid \bullet] = \sum_{k=1}^{K} E[Z_0 \mid \bullet, Z_0 \in A_k] \Pr[Z_0 \in A_k \mid \bullet] \tag{3.7}$$

where the event $\{\bullet\}$ denotes the event $\{Z_1 = z_1,...,Z_n = z_n\}$, A_k's denote the sub-ranges of the truth the union of which encompasses the entire range of the truth, and Pr[ ] denotes the probability of occurrence of the event bracketed. We rewrite the conditional expectation in Eq.(3.7) as:

$$\begin{aligned} E[Z_0 \mid \bullet, Z_0 \in A_k] &= E[Z_0 \mid \bullet, Z_0 \in A_k, OK^*] \Pr[OK^* \mid \bullet, Z_0 \in A_k] \\ &+ E[Z_0 \mid \bullet, Z_0 \in A_k, CBPK^*] \Pr[CBPK^* \mid \bullet, Z_0 \in A_k] \end{aligned} \tag{3.8}$$

where OK* or CBPK* denotes the event that the OK or CBPK estimate is more accurate than the CBPK or OK estimate, respectively. Because we do not know which sub-range may enclose the truth, we approximate the conditional expectations in Eq.(3.8) as $E[Z_0 \mid \bullet, Z_0 \in A_k, OK^*] \approx E[Z_0 \mid \bullet, OK^*]$ and $E[Z_0 \mid \bullet, Z_0 \in A_k, CBPK^*] \approx E[Z_0 \mid \bullet, CBPK^*]$. For evaluation of the conditional probabilities in Eq.(3.8), one may consider $\{Z_0 \in A_k\}$ to be more informative than $\{\bullet\}$ and approximate them as $\Pr[OK^* \mid \bullet, Z_0 \in A_k] \approx \Pr[OK^* \mid Z_0 \in A_k]$ and $\Pr[CBPK^* \mid \bullet, Z_0 \in A_k] \approx \Pr[CBPK^* \mid Z_0 \in A_k]$. With these approximations, Eq.(3.7) may be written as:

$$E[Z_0 \mid \bullet] \approx w_{OK} E[Z_0 \mid \bullet, OK^*] + (1 - w_{OK}) E[Z_0 \mid \bullet, CBPK^*] \tag{3.9}$$

where the weight for the OK estimate, w_{OK}, is given by:

$$w_{OK} = \sum_{k=1}^{K} \Pr[OK^* \mid Z_0 \in A_k] \Pr[Z_0 \in A_k \mid \bullet] \tag{3.10}$$

In Eq.(3.10), $\Pr[OK^* \mid Z_0 \in A_k]$ may be estimated from cross validation and $\Pr[Z_0 \in A_k \mid \bullet]$ may be estimated via indicator kriging or cokriging (Journel 1983, Seo 1996). It can be easily shown that the estimation variance associated with $E[Z_0 \mid \bullet]$ in Eq.(3.9) is given by:

$$\begin{aligned} Var[Z_0 \mid \bullet] &= w_{ok} Var[Z_0 \mid \bullet, OK^*] + (1 - w_{ok}) Var[Z_0 \mid \bullet, CBPK^*] \\ &+ w_{ok}(1 - w_{ok}) \{E[Z_0 \mid \bullet, OK^*] - E[Z_0 \mid \bullet, CBPK^*]\}^2 \end{aligned} \tag{3.11}$$

3.3 Extended Conditional Bias-Penalized Kriging (ECBPK)

To address negative CBPK estimates, we apply the above results but employ only two conditioning events, $\{Z_0 \in A_1\} = \{Z_0 = 0\}$ and $\{Z_0 \in A_2\} = \{Z_0 > 0\}$, and replace OK with a trivial estimator, ZERO, whose estimate is always zero. We then have $E[Z_0 \mid \bullet, ZERO^*] = 0$ where ZERO* denotes the event that an estimate of zero is more accurate than the CBPK estimate. With the above simplification, one may rewrite Eq.(3.9) as:

$$E[Z_0 \mid \bullet] = \gamma \; E[Z_0 \mid \bullet, CBPK^*] \tag{3.12}$$

In the above, CBPK* denotes the event that the CBPK estimate is more accurate than ZERO* (i.e. an estimate of zero) and the scaling coefficient γ is given by:

$$\gamma = \Pr[CBPK^* \mid \bullet] \approx \Pr[Z_0 > 0 \mid \bullet] \Pr[CBPK^* \mid Z_0 > 0] \tag{3.13}$$

In arriving at the approximation for γ above, we used $\Pr[CBPK^* \mid Z_0 = 0] = 0$. In practice, it is difficult to use Eq. (3.7) directly because it is difficult to estimate the two terms in the above approximation for γ with accuracy. Instead, one may estimate γ empirically by noting that the approximation for γ is a function of the conditional PoP, and that the final estimate, $E[Z_0 \mid \bullet]$, must be unbiased in the unconditional mean sense:

$$\gamma_e = \frac{\overline{CBPK_{raw}}}{\overline{CBPK_{trunc}}}, \quad for \quad different \quad levels \quad of \quad conditional \quad PoP \tag{3.14}$$

In the above, the overbar denotes the sample mean over some spatiotemporal scale (see below), $CBPK_{raw}$ denotes the raw CBPK estimate and $CBPK_{trunc}$ is defined as:

$$CBPK_{trunc} = \begin{cases} 0 & if \quad CBPK_{raw} < 0 \\ CBPK_{raw} & otherwise \end{cases} \tag{3.15}$$

With the above, the ECBPK estimate is then given by:

$$ECBPK = \gamma_e CBPK_{trunc} \tag{3.16}$$

where γ_e is a function of the conditional PoP. The estimation variance is obtained by multiplying γ_e^2 to the CBPK estimation variance.

In the above, the conditional PoP may be estimated, e.g., from indicator kriging or cokriging (Journel 1983, Seo et al. 1996). To avoid additional computational burden, however, here we use the fractional coverage defined as the number of precipitation-reporting neighboring gauge observations divided by the total number of neighboring gauge observations. The fractional coverage is already calculated in the Multisensor Precipitation

Estimator (MPE) used by the National Weather Service (NWS) (Seo et al. 2010) by the Single Optimal Estimator (Seo 1998) against which ECBPK is evaluated. In practice, the sample size may not be large enough to estimate γ_e in a time-varying and location-specific manner. In this work, we assume that γ_e is constant in space and time for simplicity. It is possible to relax this assumption and estimate spatially-varying γ_e over some time scale (e.g. daily, weekly, monthly, seasonal, etc.) following Seo et al. (1999, 2000). Fig 3-1 shows an example of the scaling coefficient, γ_e, as a function of the fractional coverage as estimated via Eq.(3.13) for estimation of hourly precipitation over Oklahoma (see Fig 3-1). Note that, as suggested by the approximation for γ above, the smaller the fractional coverage (i.e. the conditional PoP) is, the larger the downward adjustment of the CBPK estimate is. As described above, the ECBPK algorithm is computationally no more expensive than an OK algorithm and requires only an extremely simple addition of updating the multiplicative adjustment factor, γ_e, in real time.

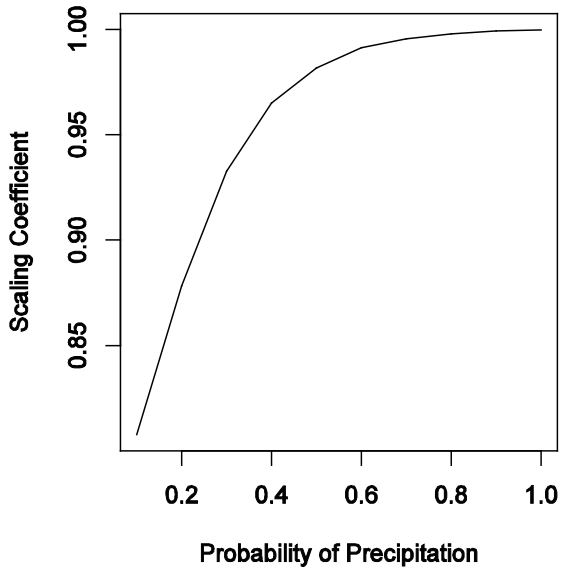

Fig 3-1 Scaling coefficient vs. probability of precipitation.

Chapter 4

Data Used

4.1 Hourly Gauge Precipitation Data

For evaluation of point precipitation estimation, hourly rain gauge data from three different regions in the United States are used. The selected three regions are the Lower Colorado River Basin Authority's (LCRA) service area in Texas, Southeastern part of the United States and the Arkansas-Red Basin River Forecast Center's (ABRFC) service area in Oklahoma and vicinity. The data was obtained from the LCRA, the National Climatic Data Center and ABRFC.

4.1.1 Arkansas-Red Basin River Forecast Center (ABRFC) Service Area

ABRFC is one of the 13 River Forecast Centers (RFC) in the United States situated in Tulsa, Oklahoma. The ABRFC service area covers almost 208,000 square miles including entire state of Oklahoma and parts of six other states. The gauge network under the ABRFC service area used for this study consists of approximately 400 gauges (based on the collected data). The gauge network includes tipping buckets and weighing gauges. ABRFC routinely quality-controls the gauge data for their real-time operations. In this study, the rain gauge data are quality-controlled using the radar precipitation data and the spatial consistency check algorithm (Kondragunta et al. 2005). The hourly data used in this work covers the years 1995 through 2000.

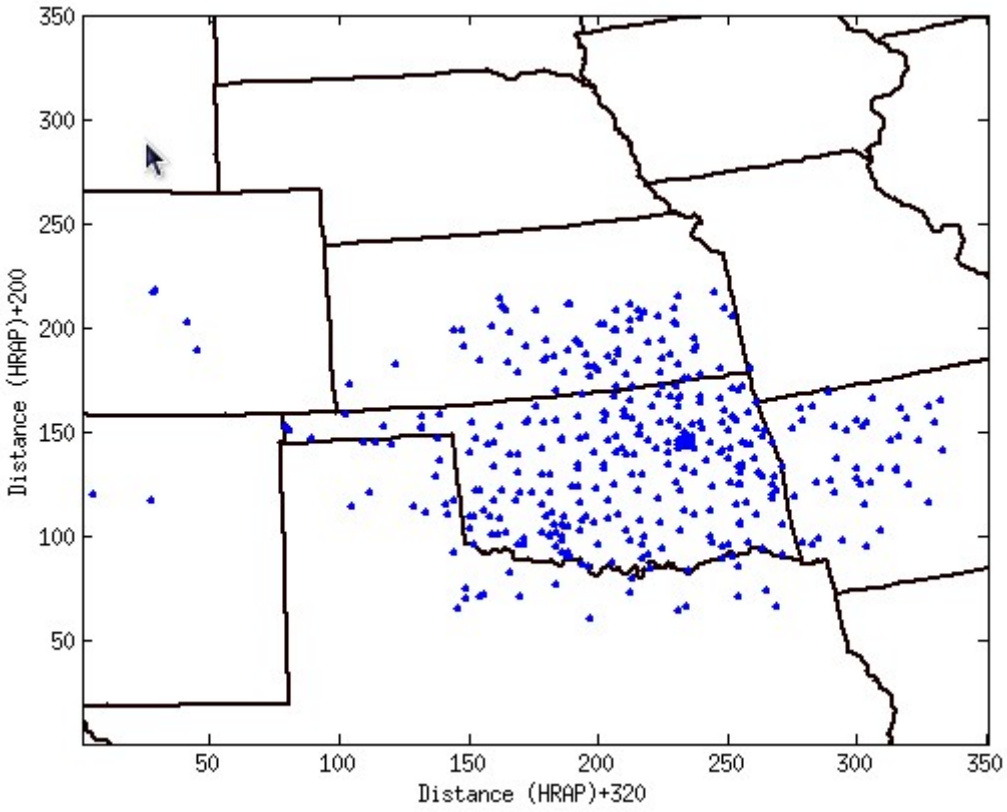

Fig 4-1 ABRFC service area gauge network.

4.1.2 Southeastern United States

A catastrophic precipitation event occurred in Southeastern part of US in September, 2009. Parts of five different states, Georgia, Alabama, Tennessee and South and North Carolinas were affected by devastating flooding. Eleven people died and property damage and economic loss ware enormous. The total precipitation in that month was almost 300 to 600 percent of normal precipitation for that region (NOAA, 2010). An area of 600 km x 600 km (360000 km^2) in the Southeastern United States has been selected as the analysis domain and hourly precipitation gauge data within this domain for the month of September in 2009 were collected and processed. This gauge network consists of nearly 500 gauges (based on the

collected data). The rain gauge data collected for this event were quality- controlled by two procedures (See section 4.3).

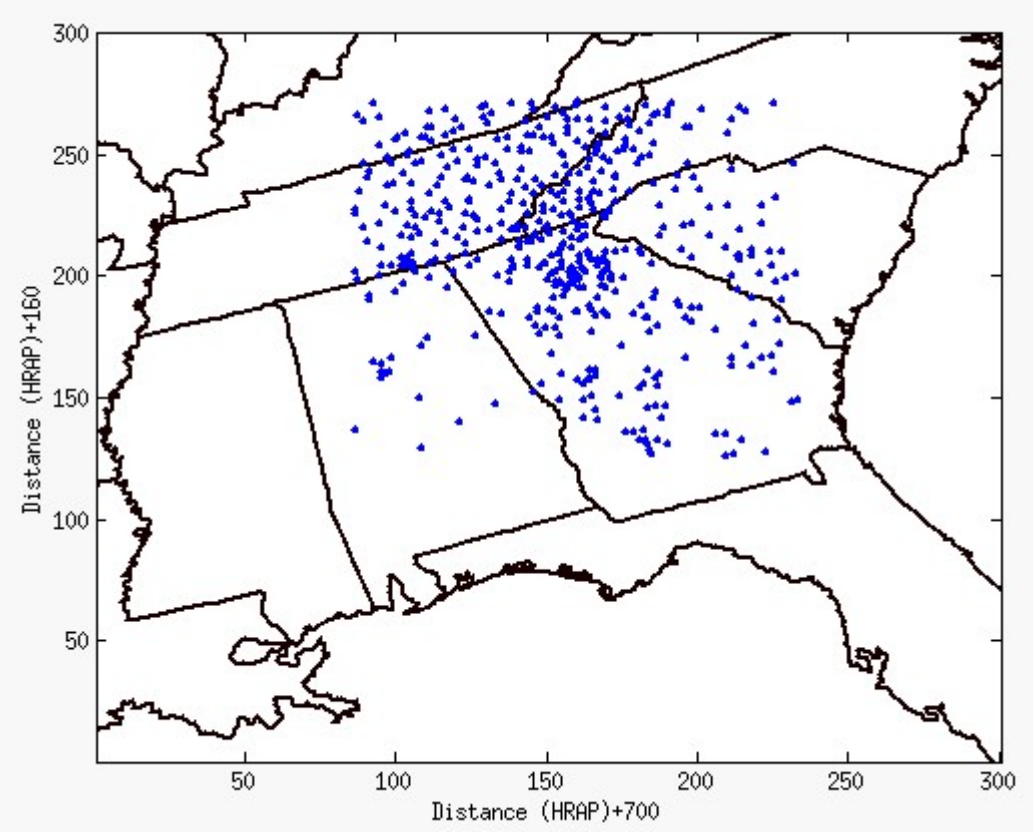

Fig 4-2 Gauge network of Southeastern US.

4.1.3 Lower Colorado River Authority (LCRA) Service Area

The Lower Colorado River Authority maintains a rain gauge network of nearly 240 gauges (LCRA, 2007). Hourly precipitation data from this network are used for the period of 2001-2007.

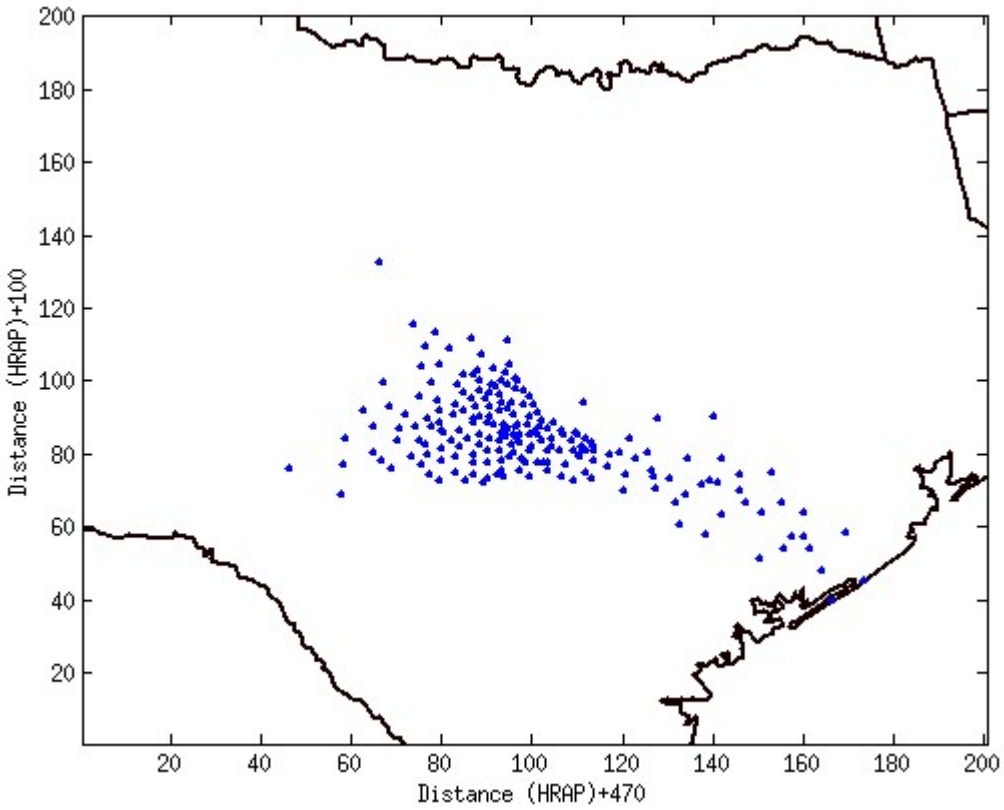

Fig 4-3 Gauge network of LCRA service area.

4.2 Hourly Stage IV Data

Stage IV is the final mosaicked version over the continental US (CONUS) of the regional MPE data which is hourly radar-rain gauge estimates of precipitation produced by 12 RFCs in CONUS (Lin and Mitchell, 2005). The MPE and hence Stage IV data are on a 4km x 4km grid known as the Hydrologic Rainfall Analysis Project (HRAP) grid (Greene et al. 1982). The mosaicking for the Stage IV data is performed at the National Centers for Environmental Protection (NCEP).

In this study, the hourly Stage IV data are used in two ways. Firstly, it is used for analysis of MAP and for generation of random gauge networks in the synthetic experiments (See section 5.2). Secondly, spatial correlation scales of precipitation are estimated using this

data. The entire CONUS is covered by a 1121 x 881 HRAP grid. An analysis domain of a 150 x 150 HRAP grid was used for the Southeast extreme event.

4.3 Quality Control of Rain Gauge Data

In this study, the rain gauge data used for Southeastern US were quality controlled by the following two procedures. First, the rain gauge data were compared with the Stage IV data. Then spatial continuity check (Kondragunta et al. 2005) was performed to eliminate the suspect rain gauge data.

4.3.1 Comparison with Stage IV

The Stage IV data is considered to be the closest representation of truth. The following procedure was applied to screen out suspect rain gauge observations.

1) If the rain gauge data reports positive precipitation but the Stage IV data shows no precipitation at that location, the rain gauge data is thrown out.
2) If the rain gauge data reports no precipitation but the Stage IV data at the same location reports positive precipitation, the rain gauge was considered stuck and thrown out.

4.3.2 Spatial Consistency Check

After the quality control using the Stage IV data, a spatial consistency check was performed for rain gauge data using an algorithm developed in NWS (Kondragunta et al. 2005). This algorithm identifies suspected outliers that show large differences than the other rain gauges in the vicinity. In the spatial consistency check algorithm, rain gauge observations in the screening area are selected and sorted for the median, 25^{th} percentile and 75^{th} percentile. Then the differences between the observation and the three percentiles for each observation are calculated and a quality control (QC) index is calculated by using Eq.(4.1) or (4.2).

If the 25th and 75th percentiles are not equal:

$$QC\ Index = \frac{|Median - Observation|}{75th\ percentile - 25th\ percentile} \quad (4.1)$$

If the 25th and 75th percentiles are equal:

$$QC\ Index = \frac{|Median - Observation|}{Average\ difference\ between\ median\ and\ observatons} \quad (4.2)$$

The QC index is compared to a threshold index of the user's choice and any observation showing a higher index than the threshold is identified as an outlier and thrown out. In this study, the threshold index used is 1.5. This means, any rain gauge observation showing an index higher than 1.5 is thrown out for inconsistency.

Chapter 5

Evaluation

In this study, a comparative evaluation is carried out between the newly proposed technique ECBPK and the currently widely-used technique OK. For evaluation of point precipitation estimation, cross validation is performed for both OK and ECBPK using the hourly rain gauge data for all three cases mentioned in the previous section. Synthetic experiments are carried out for comparatively evaluating MAP estimates for OK and ECBPK. Spatial correlation functions are estimated from the Stage IV data and used in both the real-world cases and the synthetic experiments.

5.1 Real World Cases

5.1.1 Spatial Correlation Structure

The real world cases include multiple heavy precipitation events over the ABRFC service area in Oklahoma, the 2009 Southeast extreme event, and three tropical storm events over the LCRA service area in Texas. For each case, the spatial conditional and indicator correlation scales (Seo and Smith, 1996) were calculated using the Stage IV data (Seo et al. 2010). Conditional correlograms characterize the spatial variability of precipitation within the precipitation area. Indicator correlograms characterize the spatial variability in intermittency of precipitation. The conditional (on occurrence of precipitation) and indicator correlograms were calculated in eight different directions of 0^0, 26.6^0, 45^0, 63.4^0, 90^0, 116.6^0, 135^0 and 153.4^0 (counterclockwise from due east) which were then fitted to the exponential, gaussian and spherical models (Journel and Huijbregts 1978). For all cases, exponential model provides the best fit for both conditional and indicator correlograms.

If the correlation scale is similar in all directions, the correlation structure is said to be isotropic. If it varies depending on the direction, the correlation structure is anisotropic. The directional correlograms (Figs 5-1 & 5-2), do not show any significant directional preferences.

As such, their correlation structures are considered isotropic and an average of all directional spatial correlation scales is used for each case as the representative isotropic correlation scale. Table 5-1 shows all conditional and indicator correlation scales used in this study.

Table 5-1 Conditional and indicator correlation scales used for different cases

Case	Conditional Correlation Scale (km)	Indicator Correlation Scale (km)
Oklahoma	14	82.5
Texas	22	92
Southeast extreme event	19.8	77.5

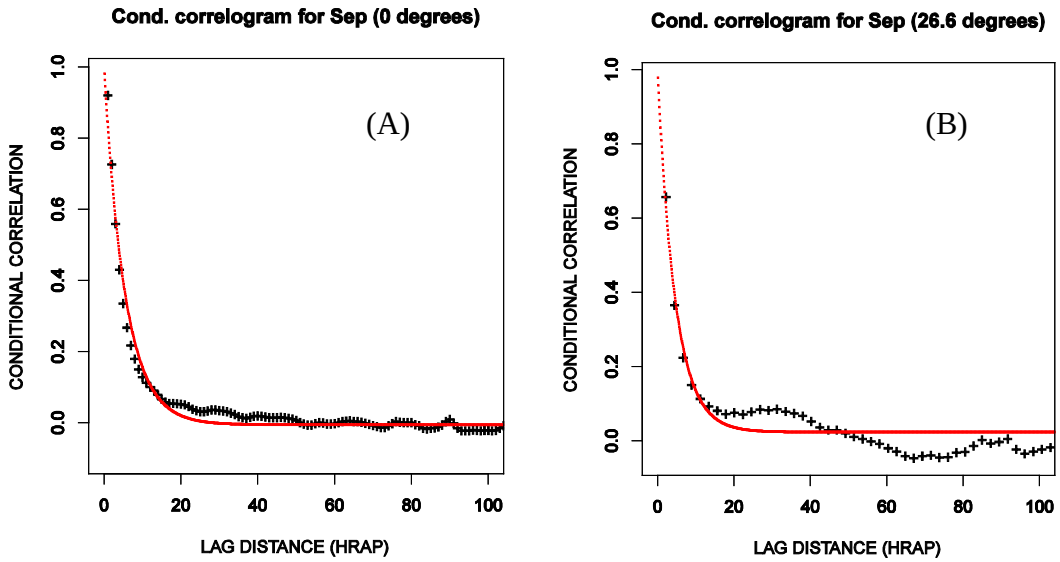

Fig 5-1 Conditional correlograms of the Southeast extreme event along (A) 0 degree and (B) 26.6 degrees for hourly precipitation.

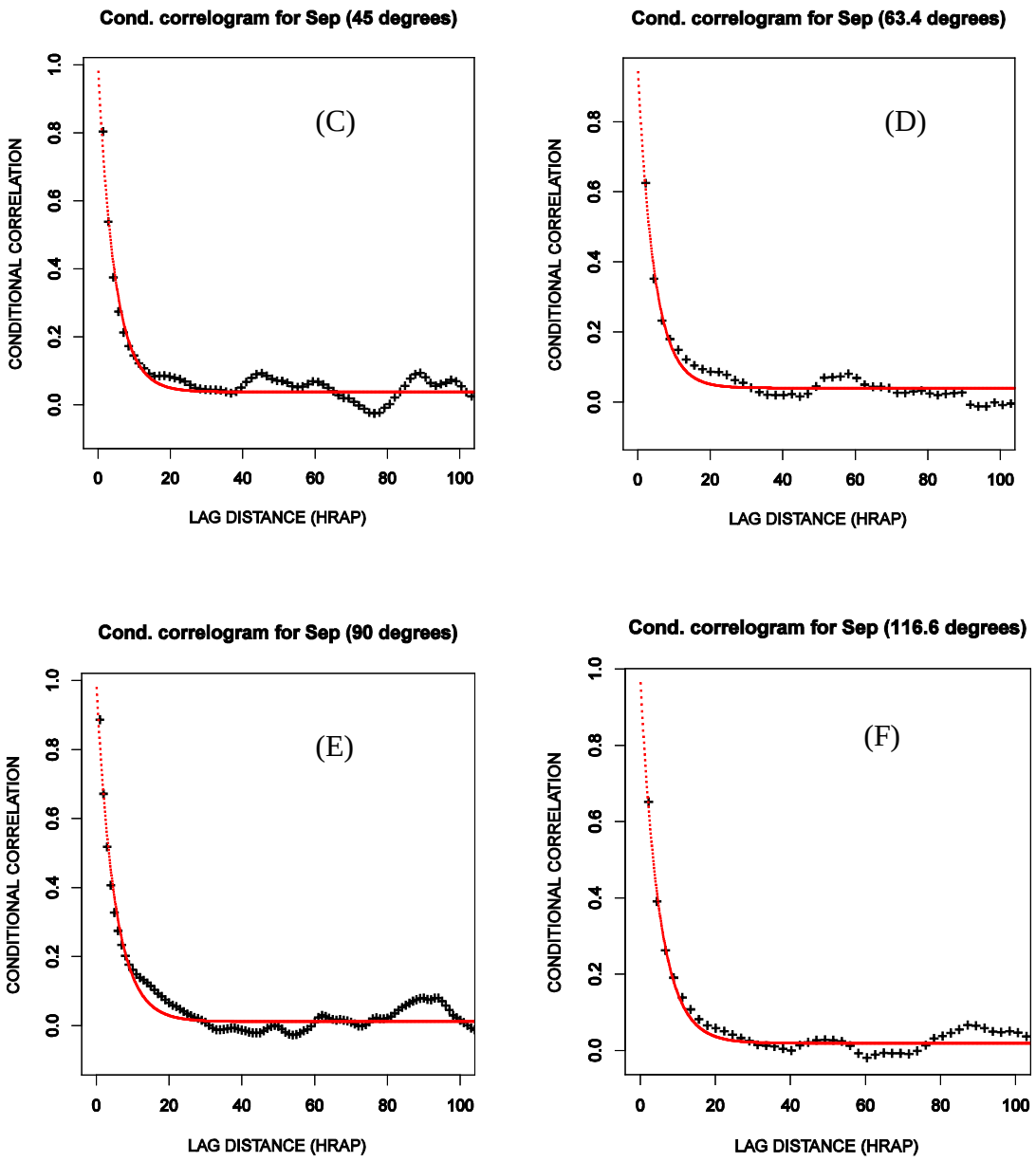

Fig 5-2 Conditional correlograms of the Southeast extreme event along (C) 45 degrees, (D) 63.4 degrees, (E) 90 degrees and (F) 116.6 degrees for hourly precipitation.

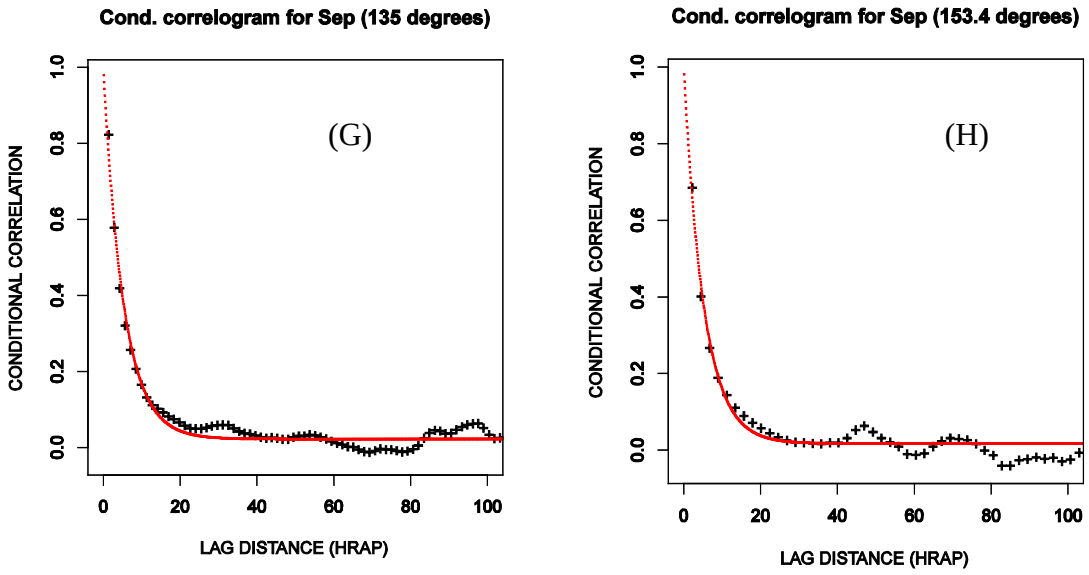

Fig 5-3 Conditional correlograms of the Southeast extreme event along (G) 135 degrees and (H) 153.4 degrees for hourly precipitation.

Some of the directional Indicator correlograms (Fig 5-4, Fig 5-5 and Fig 5-6) show outlier correlation values at very small lag distances. They have been considered as plotting artifacts and ignored.

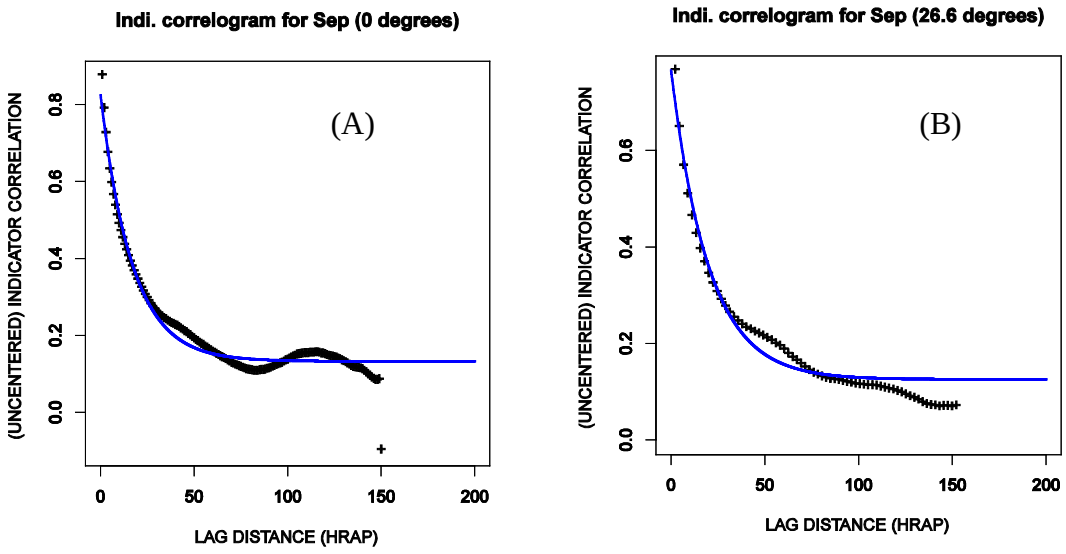

Fig 5-4 Indicator correlograms of the Southeast extreme event along (A) 0 degree and (B) 26.6 degrees for hourly precipitation.

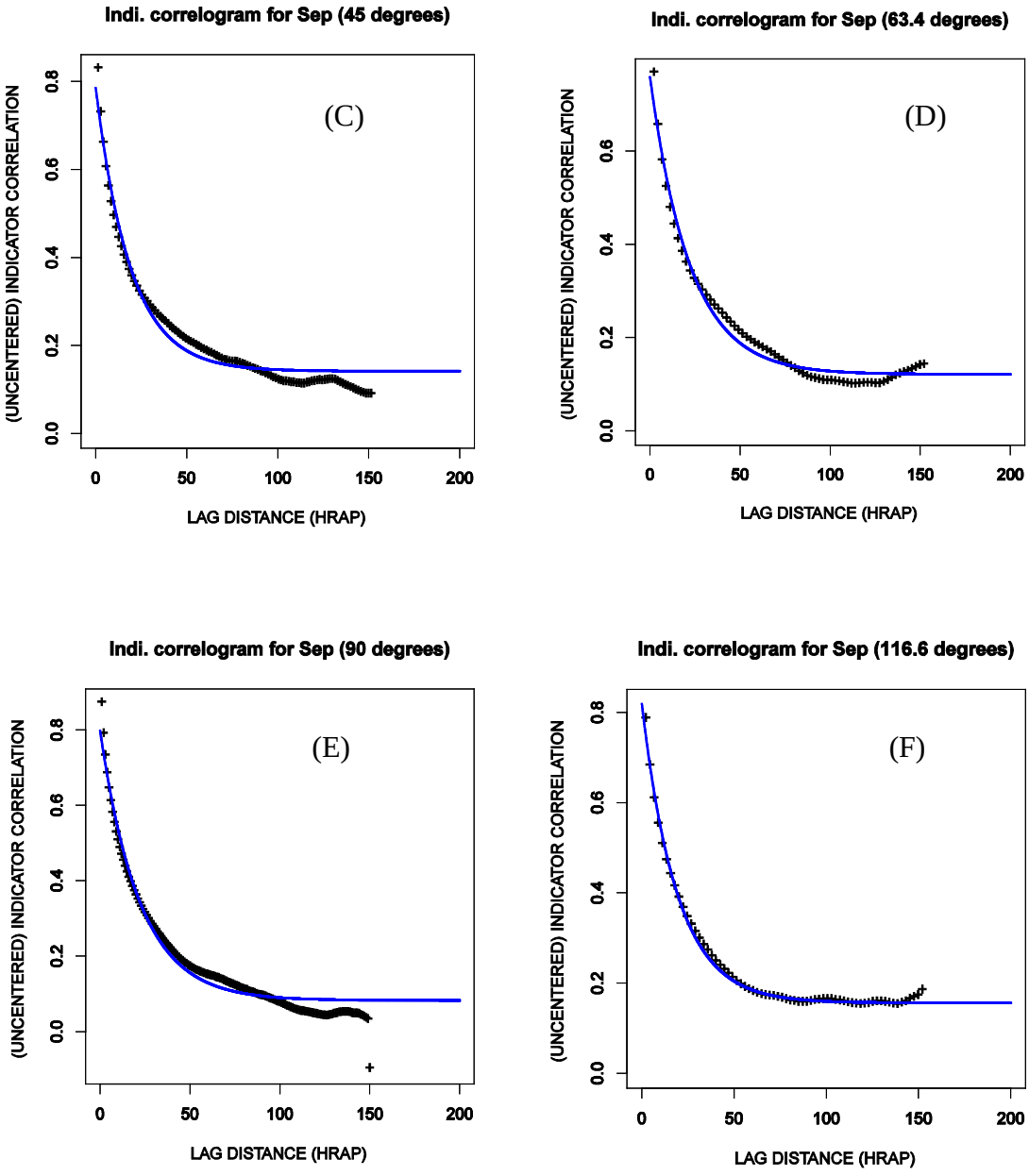

Fig 5-5 Indicator correlograms of the Southeast extreme event along (C) 45 degrees, (D) 63.4 degrees, (E) 90 degrees and (F) 116.6 degrees for hourly precipitation.

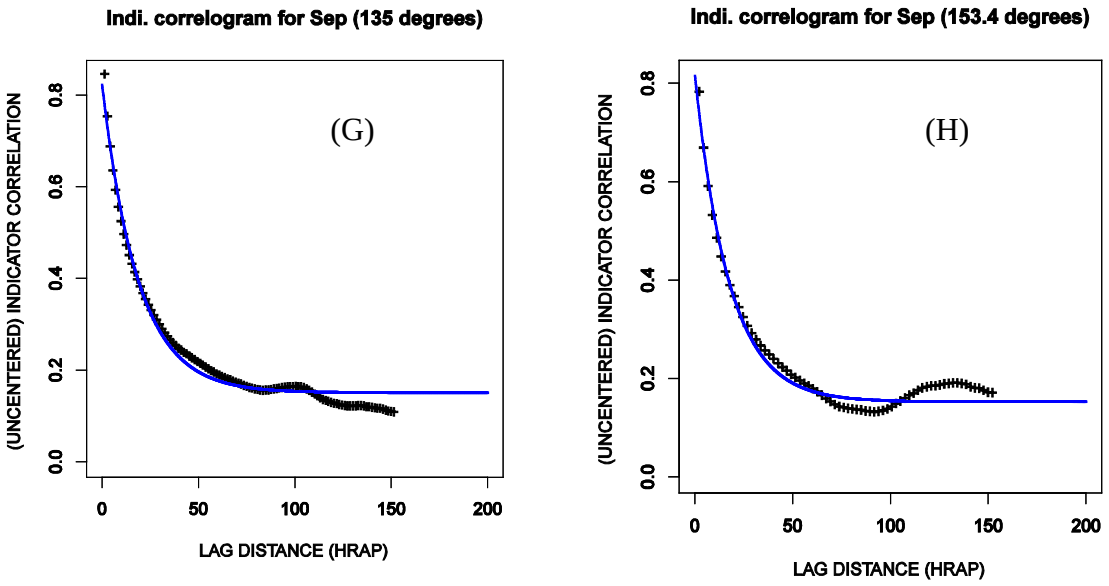

Fig 5-6 Indicator correlograms of the Southeast extreme event along (G) 135 degree and (H) 153.4 degree for hourly precipitation.

5.1.2 Cross Validation

For comparative evaluation of ECBPK, several real world experiments have been carried out using hourly gauge precipitation data for the three precipitation cases in the United States. For evaluation, cross validation was carried out for the real world experiments for both OK and ECBPK. Cross validation allows comparison between the estimated and true values using the available data set (Isaacs and Srivastava,1989). In cross validation, an observed data is withheld and then a set of neighboring data are used to estimate the value at the withheld location. Accordingly, if there are 500 gauge observations, cross validation yields 500 estimates. In this study, the number of neighbors used in the cross validation is 30 for all the real world cases.

5.2 Synthetic Experiment

In order to evaluate the MAP estimates using ECBPK, the hourly Stage IV data were used for the Southeast extreme event. The Stage IV data are considered to be the most accurate and it is reasonable to assume that they represent a plausible realization of the truth at the HRAP scale. The procedure for the synthetic experiment using the Stage IV data is described below:

1) Generate synthetic rain gauge networks of varying density by randomly selecting grid boxes from the 150x150 HRAP domain for the Southeast extreme event. Nine different traces of randomly selected gauge networks are created to reduce the sampling uncertainty. The densities used are 125, 500, 1,000, 2,000 and 4,000 gauges in the domain. In the United States, the hourly rain gauge network density is about 1/700 km^2 (Kim et al. 2009). This is very close to the density of 500 gauges in a 150 x 150 HRAP or 600 x 600 km^2 area used for the synthetic experiment in this study.
2) Perform estimation for all the HRAP grid cells and then calculate the MAP for different square basin areas of 4 x 4, 8 x 8, 16 x 16, 32 x 32 and 64 x 64 km^2 for OK and CBPK estimates. Implicit in the above procedure in which, the Stage IV estimates over HRAP grid boxes are used as point gauge precipitation is the assumption that there is no microscale variability in precipitation, and that variability at point scale is comparable to that at the HRAP scale.

Chapter 6

Results and Discussion

6.1 Point Estimation

The Single Optimal Estimator (SOE), currently used by NWS is a variant of OK. In this section, comparative results between ECBPK and OK are presented for both point and MAP estimation.

6.1.1 RMSE

The root mean squared error (RMSE) is one of the most widely used error statistics for comparing estimates with observations:

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

where Z and Z' denote the observation and the estimate, respectively and n denotes the total number of observations.

In this study, RMSE's were calculated for OK and ECBPK estimates for different thresholds of true precipitation. For all true values equal to or greater than the cutoff (i.e. threshold), the corresponding OK and ECBPK estimates are used to calculate the RMSE. Fig 6-1 shows the percent reduction in RMSE by ECBPK over OK at different cutoff values of truth. Note that the ECBPK estimates are more accurate than the OK estimates for cutoff larger than 5 to10 mm. The figure indicates that, for large cutoff values, reduction in RMSE by ECBPK over OK is almost 9 to 14 percent. For very small thresholds, however, negative reduction is seen.

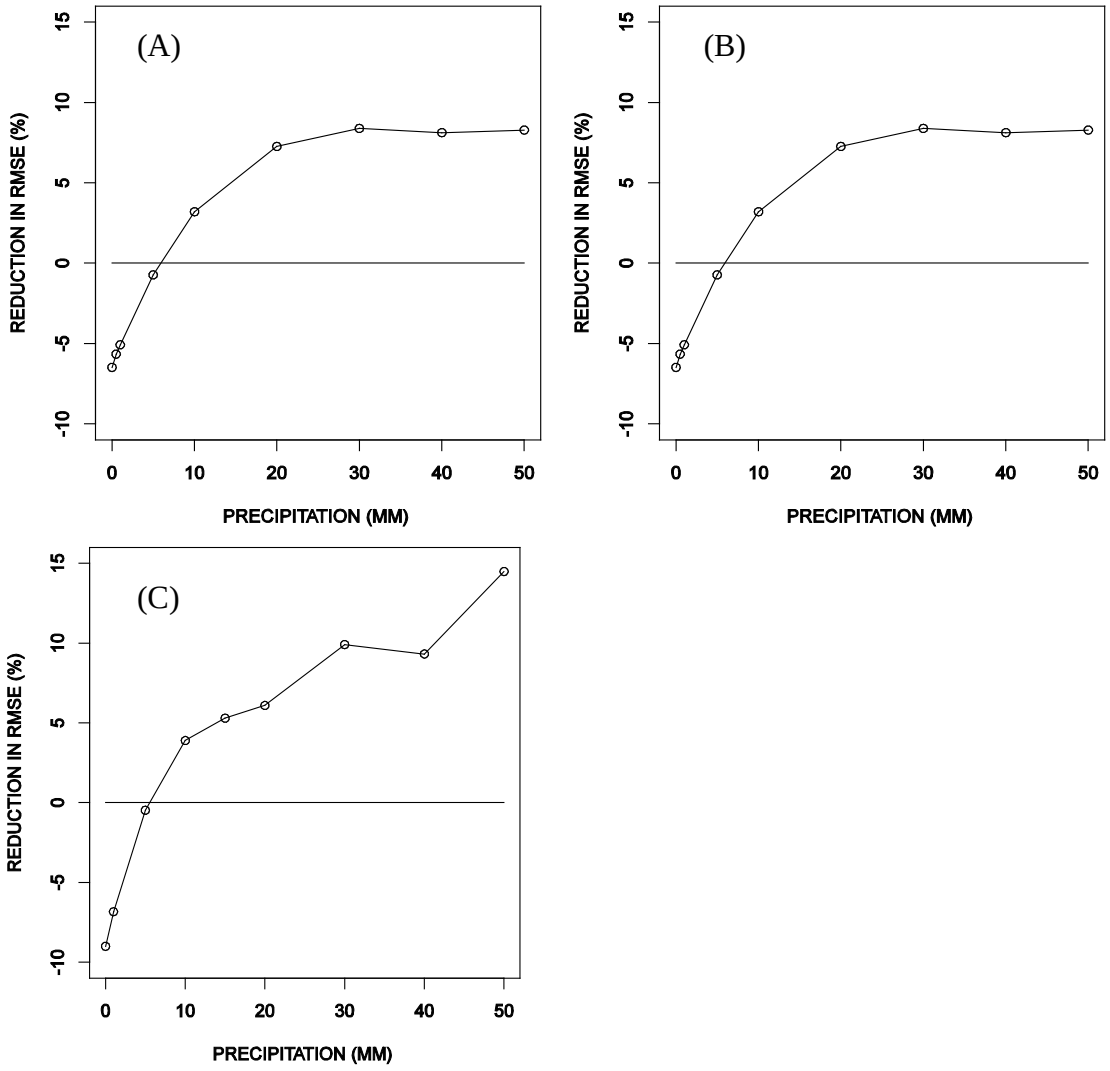

Fig 6-1 Percent reduction in RMSE by ECBPK over OK for the (A) LCRA service area in Texas (B) ABRFC service area in Oklahoma and (C) Southeast extreme event.

Fig 6-2 shows the scatter plots of the OK estimates vs the observations; note the underestimation for precipitation larger than 40 mm to 60 mm. The ECBPK estimates, on the other hand, significantly reduce underestimation for large precipitation amounts for all three cases.

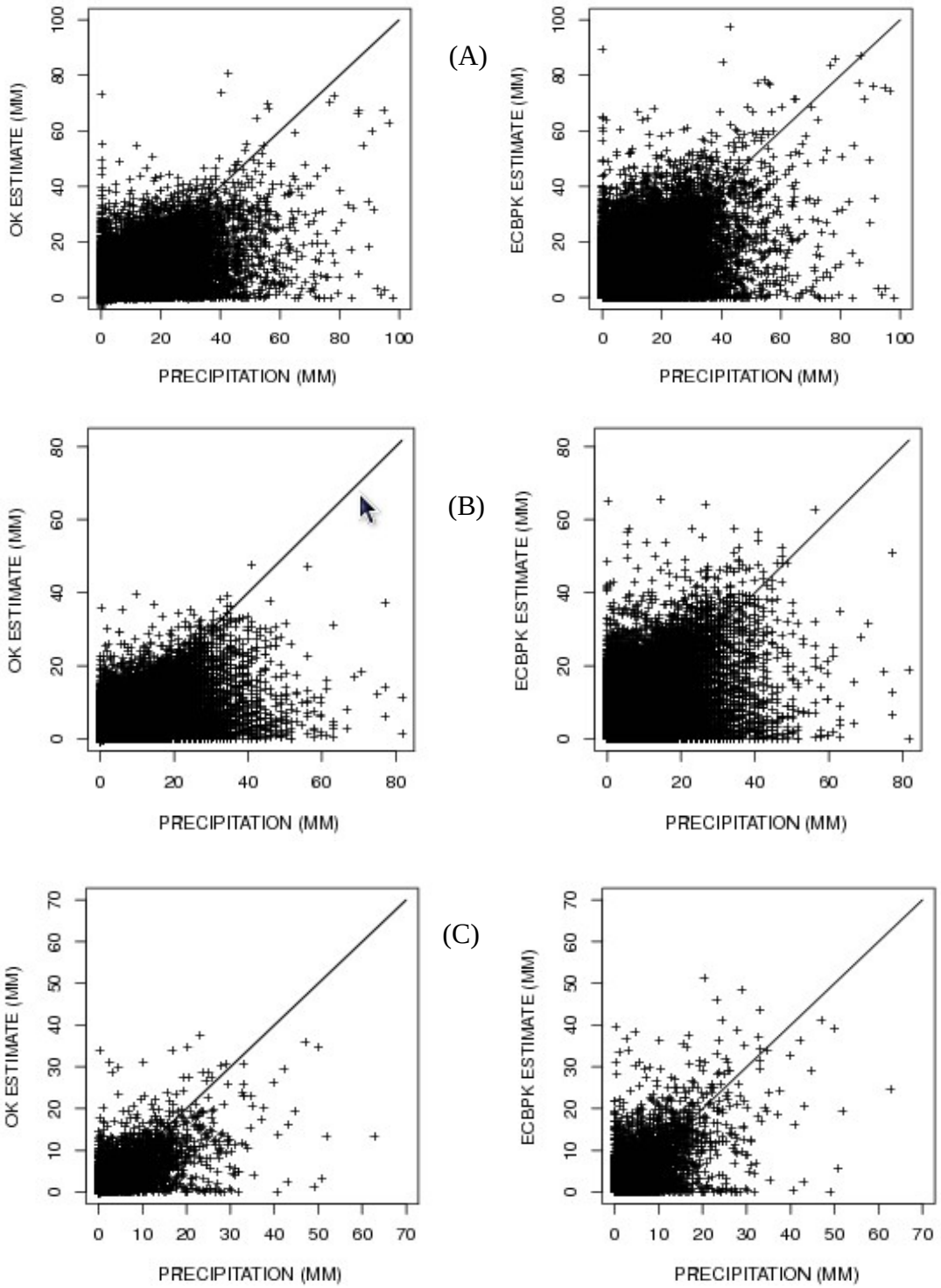

Fig 6-2 Scatter plots of the OK and ECBPK estimates vs. the observed for (A) LCRA service area in Texas (B) ABRFC service area in Oklahoma and (C) Southeast extreme event.

6.1.2 Conditional Mean

Conditional mean of the observed, OK-estimated and ECPK-estimated precipitation are calculated for different thresholds of hourly observed precipitation (i.e. truth). Fig 6-3 shows that the conditional mean of the ECBPK estimates is significantly less conditionally-biased than that of the OK estimates.

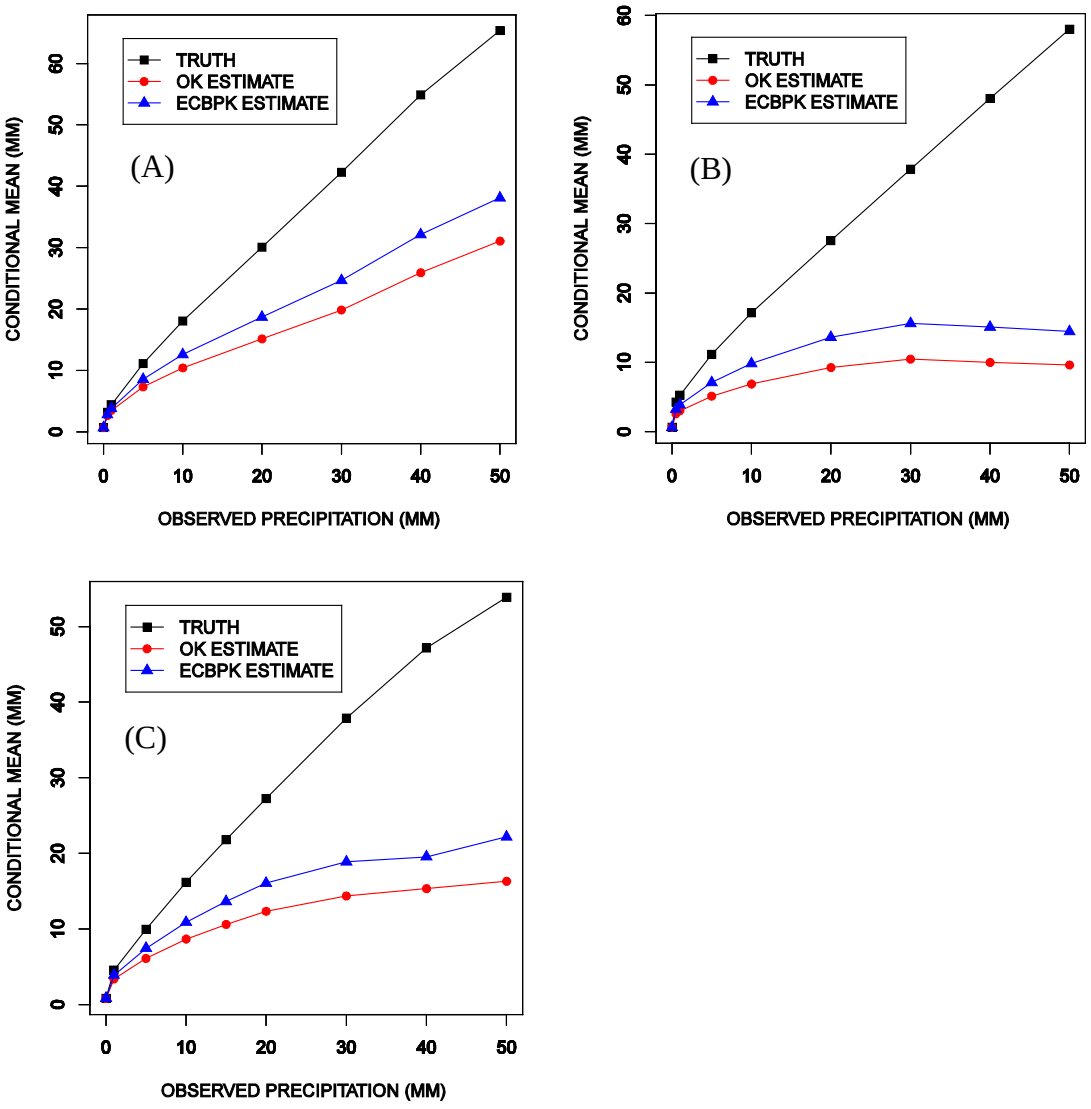

Fig 6-3 Conditional mean of the observed and estimated precipitation for (A) LCRA service area in Texas (B) ABRFC service area in Oklahoma and (C) Southeast extreme event.

Figure 6-3 indicates that, for the LCRA service area, the margin of improvement increases as the threshold increases. For the other cases, o the other hand, the margin of improvement remains steady for larger thresholds. To investigate the variation in the margin of improvement, the histograms of the average distance to the gauge observations used in the estimation are examined. Fig 6-4 shows that the LCRA network is the densest which results in the largest margin of percent reduction in Type-II CB by ECBPK. On the other hand, the ABRFC gauge network is the sparsest in the mean sense and shows the smallest percent reduction in Type-II CB by ECBPK (see Fig 6-3).

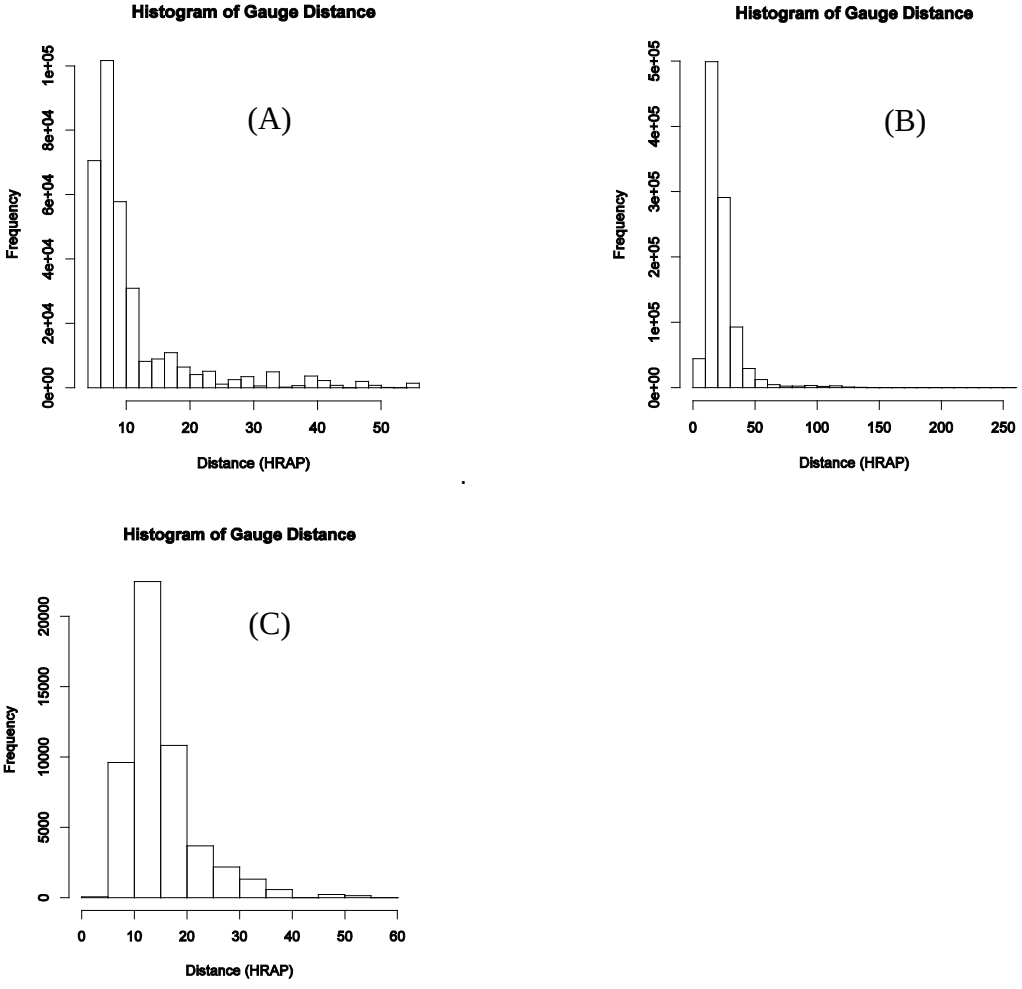

Fig 6-4 Histograms of average distances to the neighboring gauges for (A) LCRA service area in Texas (B) ABRFC service area in Oklahoma and (C) Southeast extreme event.

6.1.3 Conditional Bias and Correlation

Fig 6-5 is similar to Fig 6-4 and shows the Type-II CB in the ECBPK and OK estimates for different thresholds of precipitation amount. Fig 6-6 shows that the ECBPK estimates also have slightly larger correlation with the observed than the OK estimates for large thresholds.

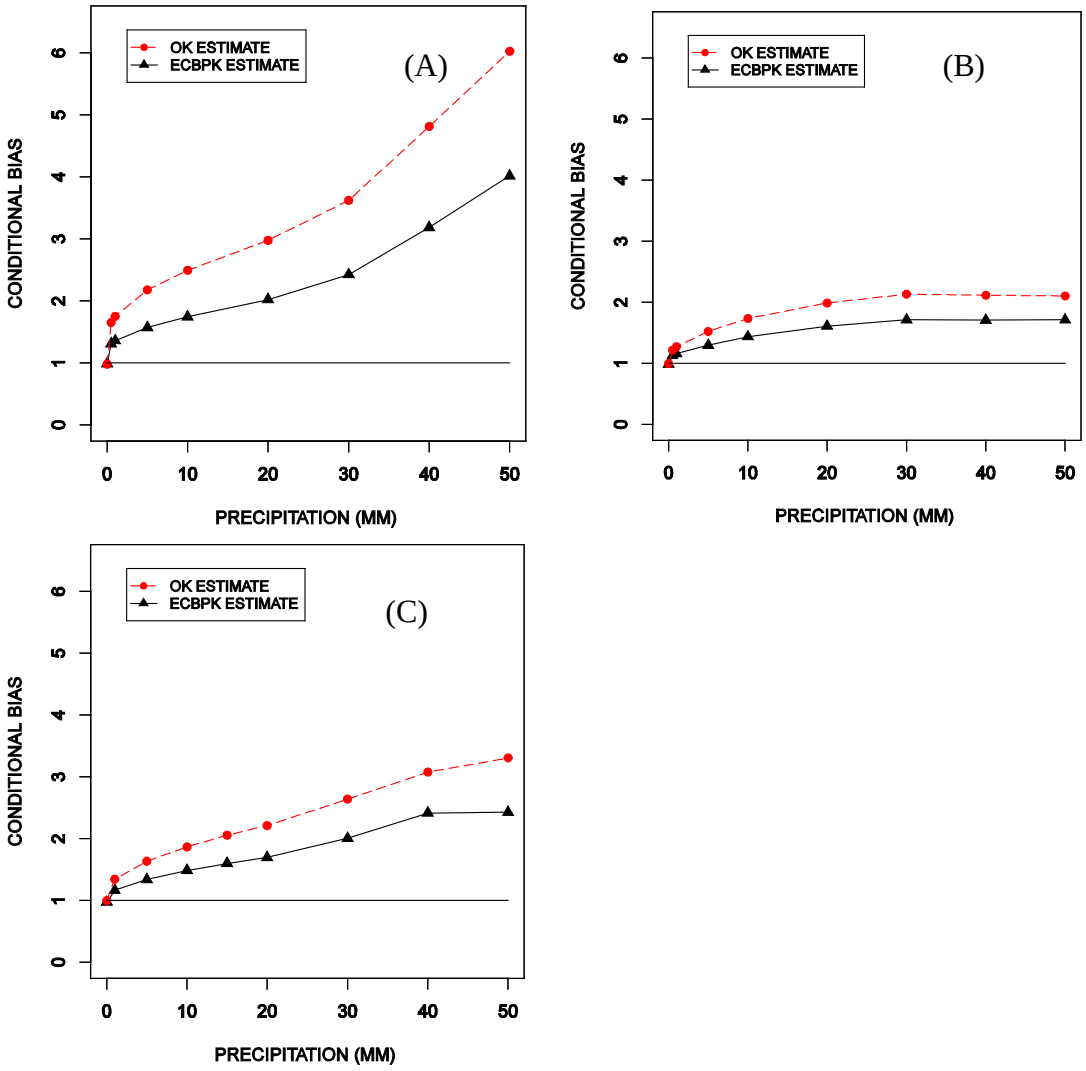

Fig 6-5 Type-II CB for (A) LCRA service area in Texas (B) ABRFC service area in Oklahoma and (C) Southeast extreme event.

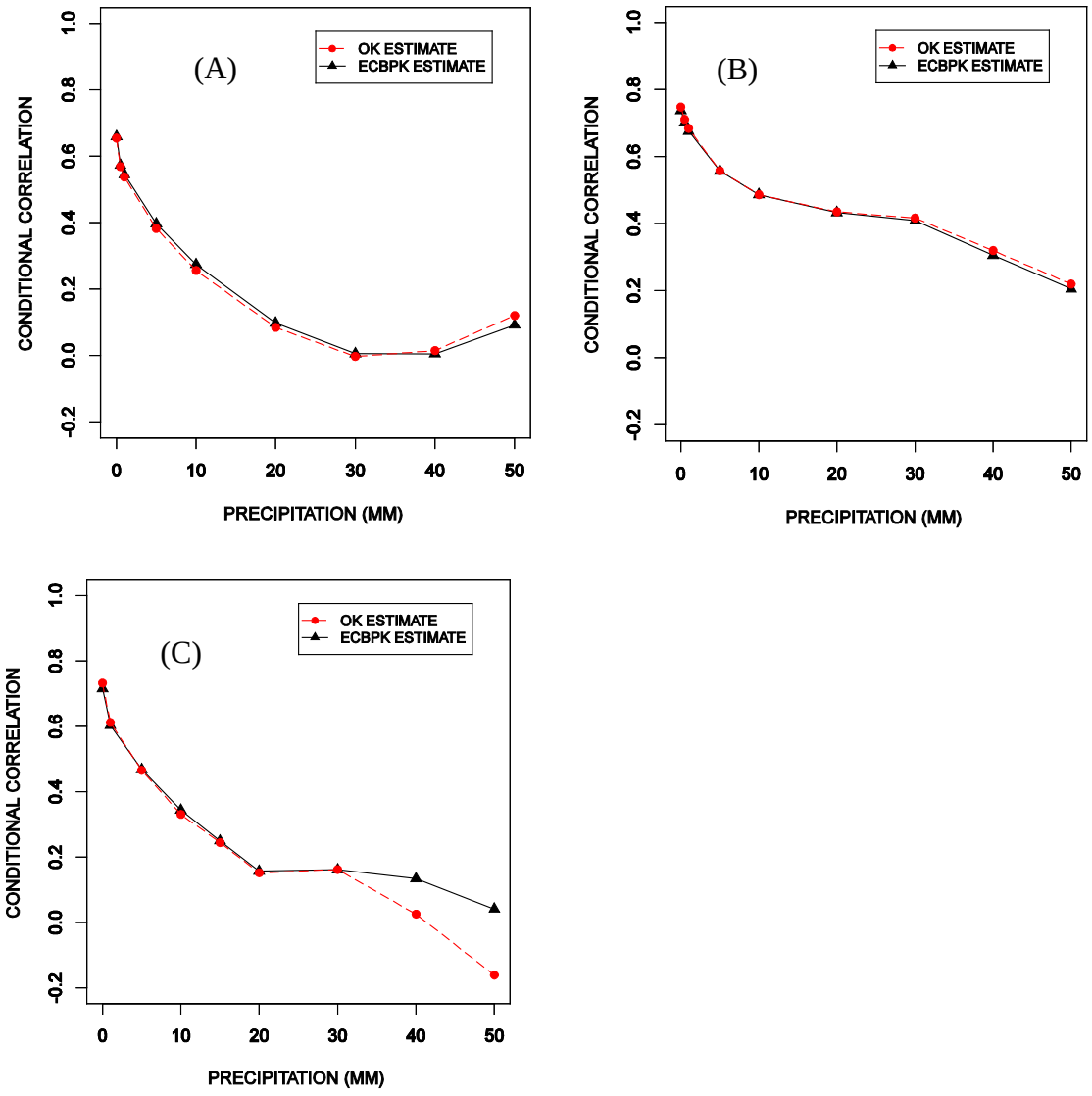

Fig 6-6 Correlation coefficient for (A) LCRA service area in Texas (B) Southeast extreme event and (C) ABRFC service area in Oklahoma.

6.1.4 Hourly Precipitation Maps

Figure 6-7 shows examples of the hourly ECBPK- and OK-estimated precipitation fields. Also shown for comparison is the hourly Stage IV precipitation field. These precipitation

fields in the figure represent the total precipitation accumulation for the Southeast extreme event that occurred in September, 2009. As the Stage IV precipitation field is the closest representation of the truth, it provides a sense of how good OK and ECBPK estimates are in comparison with truth.

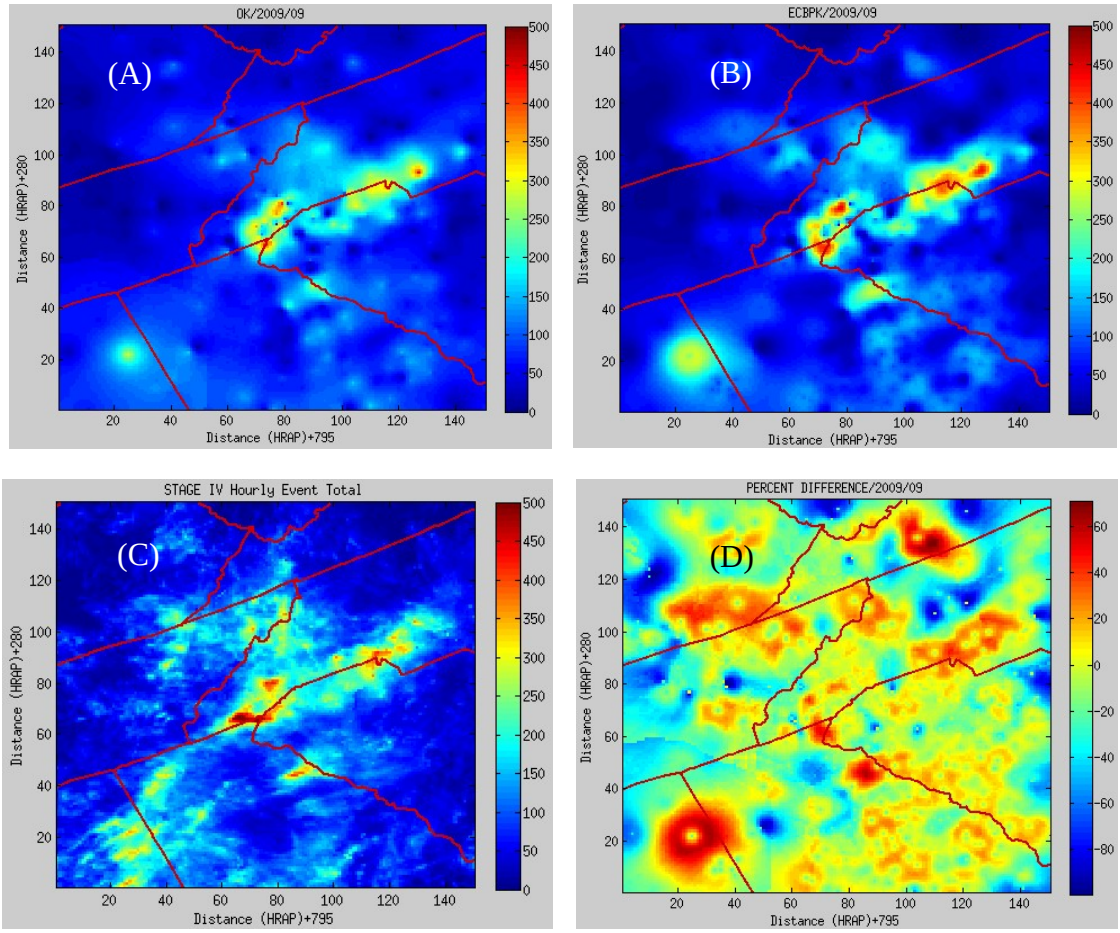

Fig 6-7 Event-total precipitation fields for the Southeast extreme event for hourly analysis of (A) OK (B) ECBPK and (C) Stage IV data and (D) percent difference between OK and ECBPK precipitation fields.

Precipitation fields in Figure 6-7 show that the ECBPK estimates are better than the OK estimates in three different ways. Firstly, ECBPK picks up the large precipitation values better than OK. ECBPK also does better in estimating very small precipitation. In addition, ECBPK captures the natural pattern of precipitation better than OK.

6.2 Estimation of MAP

Estimation of MAP was evaluated by performing synthetic experiments using the hourly Stage IV data for the Southeast extreme precipitation event for both OK and ECBPK. Cross validation was carried out for thirty different combinations of the rain gauge network density and the basin scale. The smallest basin scale was 16 km^2 (4 km x 4 km) and the largest was 4,096 km^2 (64 km x 64 km). The simulation was performed for five different synthetic gauge network densities ranging from 125 to 4,000 in a 600 x 600 km^2 area. RMSE's were calculated for all thirty combinations at varying thresholds of MAP. Figure 6-8 shows the percent reduction in RMSE as a function of the basin scale for different rain gauge network densities. The figure shows that reduction in RMSE by ECBPK over OK tends to increase with increasing gauge network density and basin scale. The maximum reduction in percent RMSE was found at a basin scale of 256 km^2.

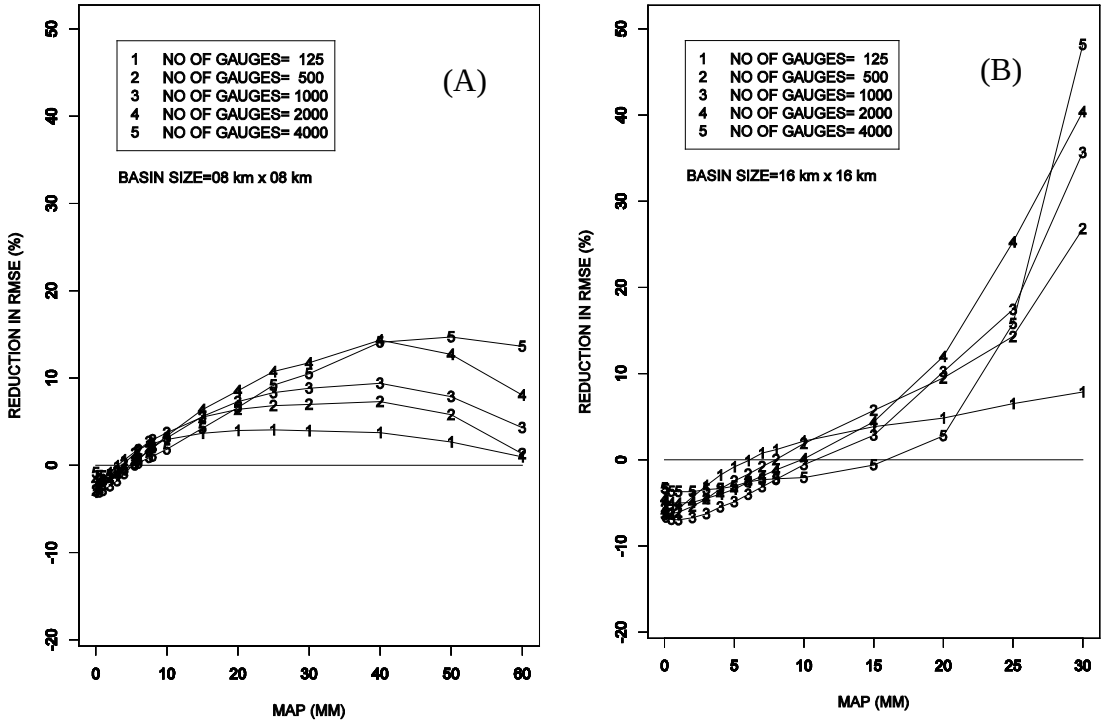

Fig 6-8 Percent reduction in RMSE by ECBPK over OK for estimation of MAP for different gauge densities for basin sizes of (A) 64 km^2 and (B) 256 km^2.

As the basin size increases, the margin of improvement starts to fall (see Fig 6-9). This is due to the decreasing fractional coverage of precipitation within the basin. The fractional coverage of precipitation gets increasingly smaller as the basin size increases. Note that, when the fractional coverage is large (close to 1), ECBPK does exceedingly well. Despite this limitation for large basins, it is clearly seen that ECBPK improves estimation of very large MAP over OK significantly (See the upper end of the scatter plots in Fig 6-10).

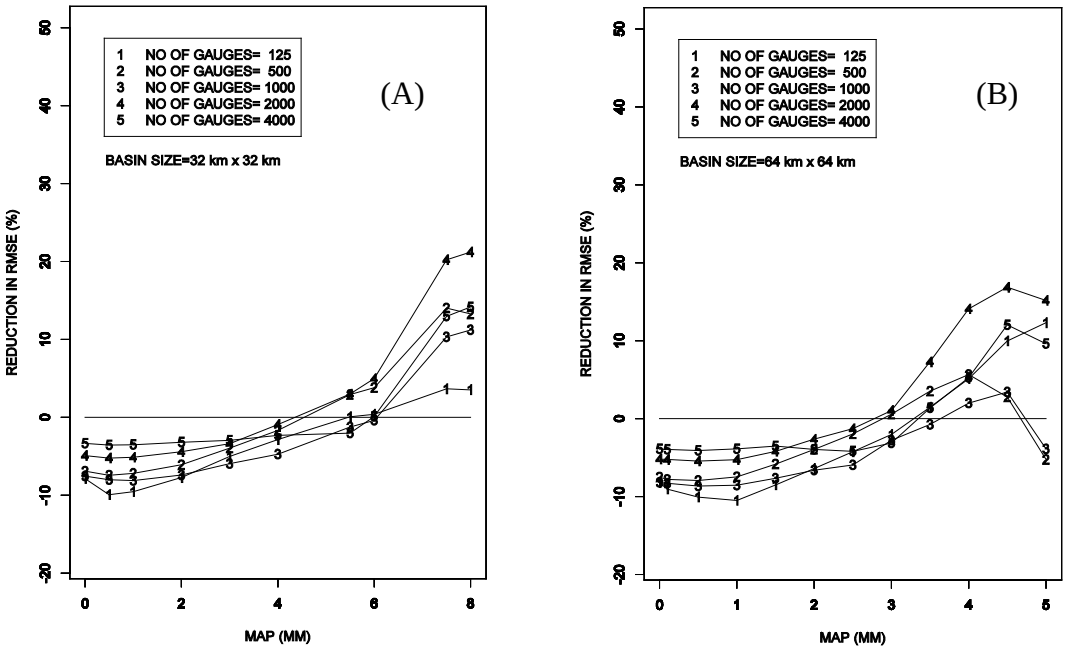

Fig 6-9 Percent reduction in RMSE by ECBPK over OK for estimation of MAP for different gauge network densities for basin sizes of (A) 1,024 km^2 and (B) 4,096 km^2.

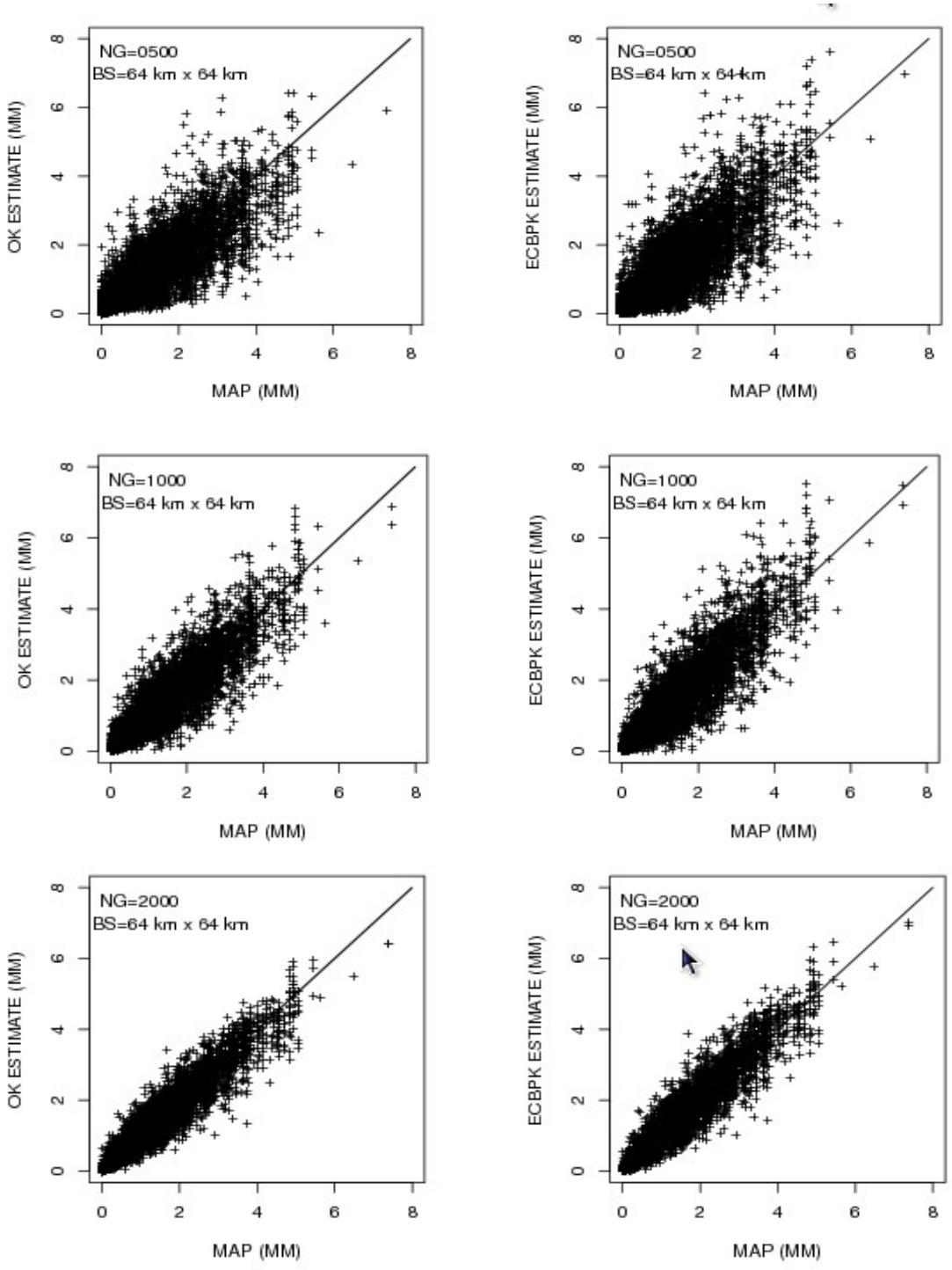

Fig 6-10 Scatter plots of OK and ECBPK estimates vs. the true MAP's for a 4,096 km^2 basin with 500, 1,000 and 2,000 gauges.

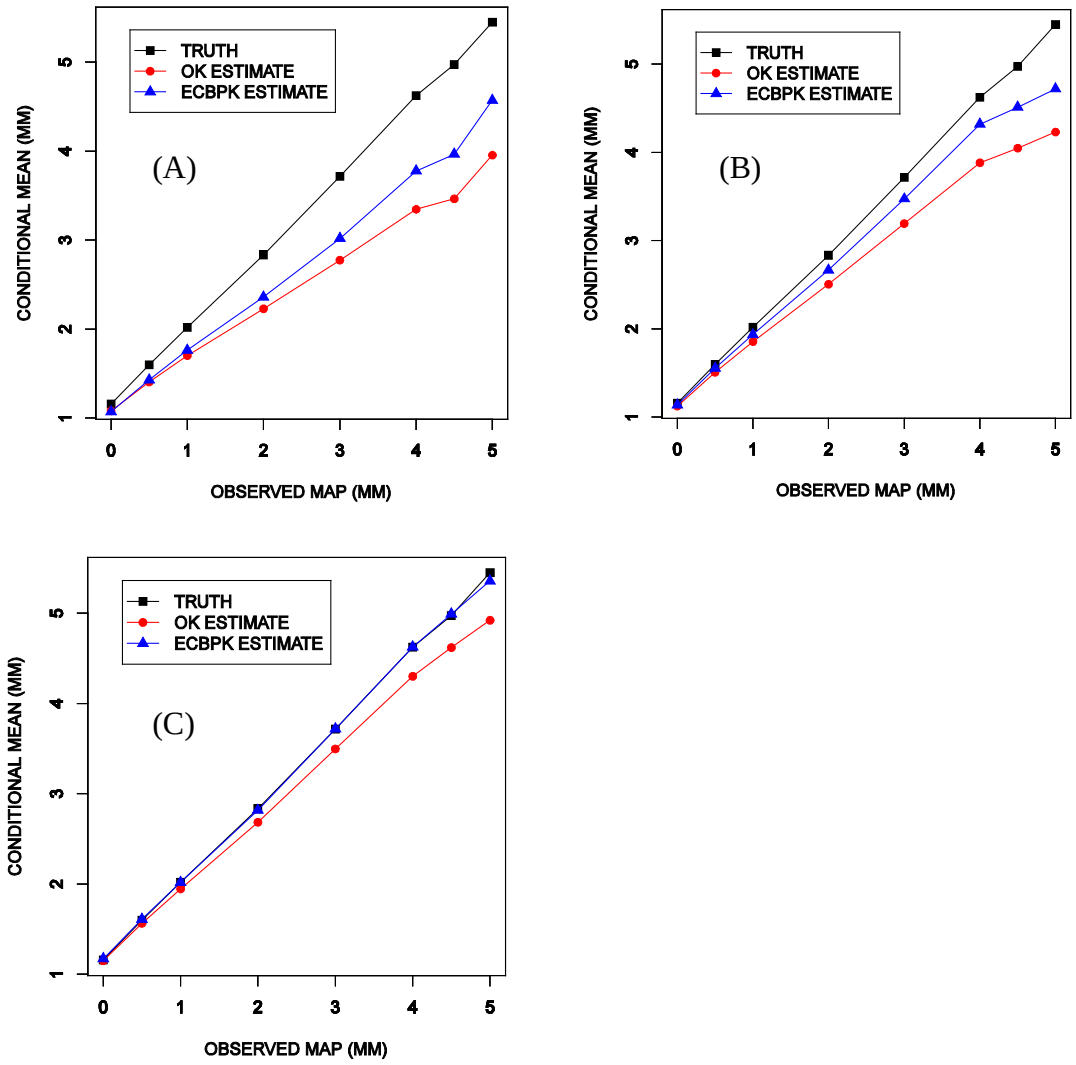

Fig 6-11 Conditional mean of the observed and estimated MAP for a 4,096 km^2 basin with gauge network density of (A) 500 gauges, (B) 1000 gauges and (C) 2000 gauges.

Fig 6-11 shows that, for MAP analysis, conditional means of ECBPK estimates are much closer to the truth in compared to the conditional means of OK estimates. This means, even though ECBPK estimates show negative percent reduction in RMSE over OK estimates for large basin scales, ECBPK is superior in estimating conditional mean for large precipitation.

Chapter 7

Conclusion and Future Recommendation

A new precipitation analysis technique, extended conditional bias-penalized kriging (ECBPK) has been developed which explicitly reduces Type-II CB. It is shown that, in comparison with OK of which the current multi-sensor precipitation estimator (MPE) algorithm used by the National Weather Service (NWS) is a variant, ECBPK improves estimation of heavy to extreme precipitation at points and over a range of catchment scale.

In this study, real world and synthetic experiments are carried out for the evaluation of point and mean areal precipitation using the newly proposed technique ECBPK. Real world cases include several heavy precipitation events in Lower Colorado River Authority's (LCRA) service area, Arkansas-Red Basin River forecast Center's (ABRFC) service area and Southeastern US. Synthetic experiments are carried out using the Stage IV data. The main findings of the real world and synthetic experiments are summarized below:

1) For estimation of point precipitation, ECBPK reduces RMSE over OK almost 9 to 14 percent for precipitation larger than 20 to 30 mm though negative reduction is seen for very small precipitation (less than 5 mm).
2) The OK estimates show increasing underestimation for precipitation larger than 25-30 mm. The ECBPK estimates show much smaller underestimation for large precipitation amounts.
3) When compared to that from OK, the conditional mean of point precipitation estimates from ECBPK at varying thresholds of observed precipitation is much closer to the truth. It is also found that a denser gauge network is able to produce more accurate estimates in terms of conditional mean.
4) The ECBPK precipitation fields are able to capture the natural pattern of precipitation better than the OK precipitation fields.

5) The ECBPK estimates show slightly larger correlation with observed precipitation in comparison with the OK estimates for large amounts of precipitation.
6) Type-II CB in ECBPK estimates is 20 to 35 percent smaller than that in OK estimates.
7) The ECBPK estimates show significant improvement over OK for basin areas ranging from 16 km^2 to 256 km^2. The largest improvement was found for the basin scale of 256 km^2 where the reduction in RMSE by ECBPK over OK reaches almost 30 percent. For basin size larger than 256 km^2, the marginal reduction is limited.

For future research, it is recommended that multivariate extension of ECBPK and CB-penalized linear filter be explored. In the context of multisensor QPE, ECBPK is expected to provide improvement when the auxiliary variable(s) are not very skillful. Such technique is hence particularly applicable when merging rain gauge data with satellite QPE, NWP analysis and/or cool-season radar QPE. In the context of linear filtering, one may expect the CB-penalized approach to improve filter performance when the observations and/or the dynamical model used is susceptible to CB.

References

Azimi-Zonooz, A., W. F. Krajewski, D. S. Bowles and D.-J. Seo, 1989. Spatial Precipitation Estimation by Linear and Non-Linear Co-Kriging of Radar-Precipitation and Raingage Data, Stoch. Hydrol. Hydraul., 3.

Bellerby, T., M. Todd, D. Kniveton, C. Kidd, 2000. Precipitation estimation from a combination of TRMM precipitation radar and GOES multispectral satellite imagery through the use of an artificial neural network. J. Appl. Meteor. 39, 2115–2128.`

Brown, J.D. and Seo, D-J., 2010. A nonparametric post-processor for bias correcting ensemble forecasts of hydrometeorological and hydrologic variables. Journal of Hydrometeorology. 11(3), 642-665.

Brown, J. D. and D.-J. Seo, 2012. Evaluation of a nonparametric post-processor for bias correction and uncertainty estimation of hydrologic predictions, Hydrol. Process. Published online in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/hyp.9263.

Chen, S.-T., P.-S. Yu, and B.-W. Liu, 2011. Comparison of Neural Network Architectures and Inputs for Radar Precipitation Adjustment for Typhoon Events, J Hydrol, in press.

Chiang, Y.-M., F.-J. Chang, B. J.-D. Jou, and P.-F. Lin, 2007. Dynamic ANN for precipitation estimation and forecasting from radar observations, J. Hydrol., 334, 250–261.

Chirlin, G., and E. Wood, 1982. On the Relationship Between Kriging and State Estimation, Water Resour. Res., 18(2), 432-438.

Ciach, G. J., M. L. Morrissey, and W. F. Krajewski, 2000. Conditional bias in radar precipitation estimation. J. Appl. Meteor., 39, 1941-1946.

Ciach, G. J., W.F. Krajewski and, G. Villarini, 2007: Product-Error-Driven Uncertainty Model for Probabilistic Quantitative Precipitation Estimation with NEXRAD Data. J. Hydrometeor, 8, 1325–1347. doi: http://dx.doi.org/10.1175/2007JHM814.1

Creutin, J. D. and C. Obled, 1982. Objective analysis and mapping techniques for precipitation field: An objective comparison. Water Resour. Res. 18(2), 413-431.

Grimes, D.I.F., E. Coppola, M. Verdecchia, G. Visconti, 2003. A neural network approach to real-time precipitation estimation for Africa using satellite data. J. Hydrometeor., 4, 1119–1133.

Habib, E., Qin, L., D.-J. Seo, G. Ciach, and B. R. Nelson, 2012. Independent Assessment of Incremental Complexity in the NWS Multi-Sensor Precipitation Estimator Algorithms,Journal of Hydrologic Engineering, doi: http://dx.doi.org/10.1061/(ASCE)HE.1943-5584.0000638.

Hsu, K.L., X. Gao, S. Sorooshian, H. V. Gupta, 2007. Precipitation estimation from remotely sensed information using artificial neural networks, J. Appl. Meteor., 36, 1176–1190.

Journel, A. G., and Ch. J. Huijbregts, 1978. Mining geostatistics, Academic Press, 600pp.

Kim, D., B. Nelson, and D.-J. Seo, 2009. Characteristics of Reprocessed Hydrometeorological Automated Data System (HADS) Hourly Precipitation Data. Weather and Forecasting, 24, 1287-1296.

Kitzmiller, D., and Coauthors, 2011: Evolving Multisensor Precipitation Estimation Methods: Their Impacts on Flow Prediction Using a Distributed Hydrologic Model. J. Hydrometeor, 12, 1414–1431. doi: http://dx.doi.org/10.1175/JHM-D-10-05038.1

Kondragunta, C., D. Kitzmiller, D.-J. Seo, and K. Shrestha, 2005. Objective Integration of Satellite, Rain Gauge, and Radar Precipitation Estimates in the Multisensor Precipitation Estimator Algorithm, 2.8, 19th Conf on Hydrology, AMS Annual Meeting.

Krajewski, W.F., 1987. Cokriging of radar-precipitation and rain gage data, J. Geophys. Res., 92 (D8), 9571-9580.

Kuligowski, R. J., 2002: A self-calibrating real-time GOES precipitation algorithm for short-term precipitation estimates. J. Hydrometeor., 3, 112-130.

Nelson, B., D.-J. Seo, and D. Kim, 2010. Multisensor Precipitation Reanalysis. Journal of Hydrometeorology. 11(3), 666-682.

NOAA service assessment report, 2010, Southeast United States Floods, September 18-23, 2009

Rosenfeld, D., D. B. Wolff, and E. Amitai, 1994. The window probability matching method for precipitation measurements with radar, JAM, 33, 682-693.

Schweppe, F. C., 1973. Uncertain Dynamic Systems. Prentice-Hall, Englewood Cliffs, NJ, 563pp.

Seo, D.-J. and J. A. Smith, 1996. Characterization of the Climatological Variability of Mean Areal Precipitation Through Fractional Coverage, Water Res. Resour., 32(7).

Seo, D.-J., 1996. Nonlinear Estimation of Spatial Distribution of Precipitation - An Indicator Cokriging, Stoch. Hydrol. Hydraul., 10(2), 127-150.

Seo, D.-J., 1998a. Real-time estimation of precipitation fields using rain gage data under fractional coverage, J. Hydrol., 208, 25-36.

Seo, D.-J., 1998b. Real-time estimation of precipitation fields using radar precipitation and rain gage data, J. Hydrol., 208, 37-52.

Seo, D.-J. And J. P. Breidenbach, 2002. Real-time correction of spatially nonuniform bias in radar precipitation data using rain gauge measurements, J. Hydrometeorol., 3, 93-111.

Seo, D.-J., S. Perica, E. Welles, and J. Schaake, 2000. Simulation precipitation fields from Probabilistic Quantitative Precipitation Forecast, J. Hydrol., 239, 203-229.

Seo, D.-J., A. Seed, and G. Delrieu, 2010. Radar-based precipitation estimation, chapter to appear in AGU Book Volume on Precipitation: Microphysics, Measurement, Estimation, and Statistical Analyses, F. Testik and M. Gebremichael, Editors.

Seo D-J. 2012. Conditional bias-penalized kriging. Stochastic Environmental Research and Risk Assessment. DOI 10.1007/s00477-012-0567-z (online first), http://www.springerlink.com/content/f5q475570q08252t/.

Tabios, G. Q. III and J. D. Salas, 1985. A comparative analysis of techniques for spatial interpolation of precipitation. Water Resour. Bull. 21(3), 365-380.

Thiessen, A. H., 1911. Precipitation Averages for Large Areas. Mon. Wea. Rev. 39(7), 1082-1089.

Zhang, Y., D.-J. Seo, D. Kitzmiller, H. Lee, R. Kuligowski, D. Kim, C. Kondragunta , 2012. Comparative Strengths of SCaMPR Satellite QPEs with and without TRMM ingest vs. Gridded Gauge-only Analyses, to appear in JHM.

Zhang, J., K. Howard, C. Langston, S. Vasiloff, B. Kaney, A. Arthur, S. V. Cooten, K. Kelleher, D. Kitzmiller, F. Ding, D.-J. Seo, E. Wells and C. Dempsey, 2011. National Mosaic and Multi-sensor QPE (NMQ) system, AMS Bulletin, Oct, 1321-1338.

Zhang, J., Y. Qi, D. Kingsmill, K. Howard, 2012. Radar-Based Quantitative Precipitation Estimation for the Cool Season in Complex Terrain: Case Studies from the NOAA Hydrometeorology Testbed, J.Hydrometeorology,13,1836-1854.

Biographical Information

Ridwan Siddique was born in Chittagong, Bangladesh on June 23, 1986. He graduated with a Bachelor of Science in Water Resources Engineering from Bangladesh University of Engineering and Technology in August 2009. The author joined the University of Texas at Arlington in Fall 2011 for graduate studies. He had the opportunity to work as a graduate research assistant under Dr. Dong Jun Seo of Civil Engineering Department. After completing the MS, the author's plan is to pursue the doctoral degree. The author's research interests include hydrology, hydrometeorology, hydroclimatology and remote sensing.