Fall 2009

Role of rainfall variability in the statistical structure of peak flows

Pradeep Mandapaka Venkata
University of Iowa

ROLE OF RAINFALL VARIABILITY IN THE STATISTICAL STRUCTURE OF PEAK FLOWS

by

Pradeep Mandapaka Venkata

An Abstract

Of a thesis submitted in partial fulfillment of the
requirements for the Doctor of Philosophy degree
in Civil and Environmental Engineering in
the Graduate College of
The University of Iowa

December 2009

Thesis Supervisor: Professor Witold F. Krajewski

ABSTRACT

This thesis examines the role of rainfall variability and uncertainties on the spatial scaling structure of peak flows using the Whitewater River basin in Kansas, and Iowa River basin in Iowa as illustrations. We illustrate why considering individual hydrographs at the outlet of a basin can lead to misleading interpretations of the effects of rainfall variability. The variability of rainfall is characterized in terms of storm intensity, duration, advection velocity, zero-rain intermittency, variance and spatial correlation structure. We begin with the simple scenario of a basin receiving spatially uniform rainfall of varying intensities and durations, and advection velocities. We then use a realistic space-time rainfall field obtained from a popular rainfall model that can reproduce desired storm variability and spatial structure. We employ a recent formulation of flow velocity for a network of channels and calculate peak flow scaling exponents, which are then compared to the scaling exponent of the channel network width function maxima. The study then investigates the role of hillslope characteristics on the peak flow scaling structure. The basin response at the smaller scales is driven by the rainfall intensities (and spatial variability), while the larger scale response is dominated by the rainfall volume as the river network aggregates the variability at the smaller scales. The results obtained from simulation scenarios can be used to make rigorous interpretations of the peak flow scaling structure obtained from actual space-time model, and actual radar-rainfall events measured by the NEXRAD weather radar network.

An ensemble of probable rainfall fields conditioned on the given radar-rainfall field is then generated using a radar-rainfall error model and probable rainfall generator. The statistical structure of ensemble fields is then compared with that of given radar-rainfall field to quantify the impact of radar-rainfall errors on 1) spatial characterization of the rainfall events and 2) scaling structure of the peak flows. The effect of radar-

rainfall errors is to introduce spurious correlations in the radar-rainfall fields, particularly at the smaller scales. However, preliminary results indicated that the radar-rainfall errors do not significantly affect the peak flow scaling exponents.

Abstract Approved: _____
Thesis Supervisor

_____
Title and Department

_____
Date

ROLE OF RAINFALL VARIABILITY IN THE STATISTICAL STRUCTURE OF PEAK FLOWS

by

Pradeep Mandapaka Venkata

A thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Civil and Environmental Engineering
in the Graduate College of
The University of Iowa

December 2009

Thesis Supervisor: Professor Witold F. Krajewski

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Pradeep Mandapaka Venkata

has been approved by the Examining Committee
for the thesis requirement for the Doctor of Philosophy
degree in Civil and Environmental Engineering at the December 2009
graduation.

Thesis Committee: _____
Witold F. Krajewski, Thesis Supervisor

_____
Allen Bradley

_____
Dale L. Zimmerman

_____
Anton Kruger

_____
Nandita Basu

To my Mom Devaki Kumari

ACKNOWLEDGMENTS

This thesis would not have been possible without the help of great many individuals. I consider myself really fortunate to work under the guidance of Prof. Witold Krajewski. I thank him for continuous support and encouragement throughout the course of this thesis. I truly enjoyed his classes and every discussion with him. He offered full freedom to explore various topics, although not entirely related to this thesis, and steered me in the right direction whenever required. I learned from him not just science but also other aspects which will remain with me forever.

I would like to extend my gratitude to Professors Allen Bradley, Dale Zimmerman, Anton Kruger, and Nandita Basu for serving on my thesis committee, and for their useful suggestions. I acknowledge Dr. Grzegorz Ciach for helpful discussions related to Chapter 5. I thank Dr. Thomas Bell of GSFC, NASA for helpful discussions related to the precipitation model. Many thanks to Dr. Ricardo Mantilla, who was instrumental in shaping this thesis to its present form. The discussions I had with him over the last couple of years have greatly improved my understanding of hydrologic modeling.

I would like to thank my desk neighbor for four years, Dr. Gabriele Villarini for patiently answering my questions, and for being the devil's advocate in all the research discussions. One could not wish for a better colleague/friend. My lab mates Piotr, Renato, Luciana, and Bongchul deserve special mention for their help during various phases of the thesis. Thanks to Radoslaw Goska for his help with the graphics often on a very short notice. I would also like to thank the administrative and computer staff of IIHR-Hydroscience and Engineering for their efficient handling of all non-research matters.

I take this opportunity to thank my roommates over the years for their friendship, for all the good times, and for those never-ending debates and opinions on everything on

the planet. Friends and family back in India deserve appreciation not for their question "When are you graduating?" but for always being there for me in tough and happy times alike. It is impossible to put in words, the role of my Mom and Dad in shaping my career. I am grateful to them for their unconditional love and support and for instilling in me the value for education. Last but not the least, I would like to thank my loving sister for her constant encouragement.

ABSTRACT

This thesis examines the role of rainfall variability and uncertainties on the spatial scaling structure of peak flows using the Whitewater River basin in Kansas, and Iowa River basin in Iowa as illustrations. We illustrate why considering individual hydrographs at the outlet of a basin can lead to misleading interpretations of the effects of rainfall variability. The variability of rainfall is characterized in terms of storm intensity, duration, advection velocity, zero-rain intermittency, variance and spatial correlation structure. We begin with the simple scenario of a basin receiving spatially uniform rainfall of varying intensities and durations, and advection velocities. We then use a realistic space-time rainfall field obtained from a popular rainfall model that can reproduce desired storm variability and spatial structure. We employ a recent formulation of flow velocity for a network of channels and calculate peak flow scaling exponents, which are then compared to the scaling exponent of the channel network width function maxima. The study then investigates the role of hillslope characteristics on the peak flow scaling structure. The basin response at the smaller scales is driven by the rainfall intensities (and spatial variability), while the larger scale response is dominated by the rainfall volume as the river network aggregates the variability at the smaller scales. The results obtained from simulation scenarios can be used to make rigorous interpretations of the peak flow scaling structure obtained from actual space-time model, and actual radar-rainfall events measured by the NEXRAD weather radar network.

An ensemble of probable rainfall fields conditioned on the given radar-rainfall field is then generated using a radar-rainfall error model and probable rainfall generator. The statistical structure of ensemble fields is then compared with that of given radar-rainfall field to quantify the impact of radar-rainfall errors on 1) spatial characterization of the rainfall events and 2) scaling structure of the peak flows. The effect of radar-

rainfall errors is to introduce spurious correlations in the radar-rainfall fields, particularly at the smaller scales. However, preliminary results indicated that the radar-rainfall errors do not significantly affect the peak flow scaling exponents.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Motivation

Floods are one of the deadliest natural hazards. Recent examples are Cedar River and Iowa River floods in June 2008, which caused unprecedented damage. According to the 1997 International Red Cross report on world disasters, across the globe, more than 318,000 people were killed due to floods in the 25-year period from 1972 to 1997 [*IFRCRCS* 1997]. *Ashley and Ashley* [2008] reported a total of 4586 flood related fatalities in contiguous United States during 47-year period from 1959 to 2005. The Fourth Assessment Report (2007) of the Intergovernmental Panel on Climate Change (IPCC) predicts that 'heavy precipitation events, which are very likely to increase in frequency, will augment flood risk'. Floods result from complex interaction of rainfall with various aspects of landscape such as topography, land cover, and antecedent moisture. In addition, the changing climate and man-made changes to the landscape make the flood prediction, a very complex problem.

Based on historical streamflow records, the U.S. Geological Survey (USGS) developed the regional quantile regression approach in the 1960s for predicting annual flood frequencies [e.g., *Dawdy*, 1961; *Benson*, 1962]. The regression relations expressed flood quantiles as power laws involving one or more explanatory variables such as drainage area, basin geomorphology, and climatic characteristics. Drainage area is usually the most important explanatory variable, and many times the only variable, used [e.g., *Riggs*, 1973; *Jennings et al.*, 1994]. The quantile regression method was extensively employed to obtain flood frequencies for several basins including ungauged locations. Many have supported the above statistical relations and showed that despite all the complexity of floods at the hillslope scale, their aggregated behavior is a power-law (scale-invariant) with respect to the drainage areas [e.g., *Smith*, 1992; *Gupta et al.*, 1994;

Goodrich et al., 1997; *Vogel and Sankarasubramanian*, 2000; *Eaton et al.*, 2002; *Ogden and Dawdy*, 2003; *Furey and Gupta*, 2005]. This provides us with an opportunity to develop a unified geophysical framework for studying the hydrologic response of a basin across a wide range of scales. Such a framework would be invaluable for prediction of flood peaks particularly in ungauged basins [e.g., *Sivapalan et al*., 2003]. However, a comprehensive understanding of the role of physical processes in the statistical scaling relations is crucial for building and implementing such a framework [e.g., *Gupta et al.* 2007].

In parallel to developing the statistical regression equations for flood quantiles, hydrologic community also developed models to predict hydrographs at the outlet or at specific locations [e.g., *Crawford and Linsley*, 1962; *Freeze and Harlan*, 1969; *Burnash et al.*, 1973; *Beven and Kirby*, 1979; *Abbott et al.*, 1986a,b; *Beven*, 2001; *Singh and Frevert*, 2002a,b]. Several studies have used these models for understanding the effect of variability in processes such as rainfall, antecedent moisture, and soil infiltration capacity on the hydrologic response of the basin [e.g., *Dawdy and Bergmann*, 1969; *Beven and Hornberger*, 1982; *Milly and Eagleson*, 1988; *Krajewski et al.*, 1991; *Obled et al.*, 1994; *Shah et al.*, 1996; *Finnerty et al.*, 1997; *Singh*, 1997; *Koren et al.*, 1999; *Arnaud et al.*, 2002; *Brath and Montaniri*, 2003; *Brath et al.*, 2004; *Smith et al.*, 2004b; *Tetzlaff and Uhlenbrook*, 2005; *Morin et al*., 2006; *Fenicia et al.*, 2008; *Saulnier and Le Lay*, 2009]. One can draw upon this large and growing pool of (deterministic) hydrologic modeling studies to understand the physics behind the statistical scaling behavior of peak flows.

However, one fundamental problem with most hydrologic models is that the governing equations used to describe the physical processes in hydrologic models are sensitive to the scale [e.g., *Klemes*, 1983; *Beven*, 1985; *Beven*, 1989; *Bhaskar*, 1989; *Grayson et al.*, 1992; *Bloschl and Sivapalan*, 1995; *Beven*, 2001; *Kirchner*, 2006]. For example, the Darcy's law (subsurface flow) and the Richard's equation (unsaturated flow) with "effective" parameters are usually employed by hydrologic models to

represent the corresponding processes. Although physically-based, their applicability was debated even for a laboratory column scale [e.g., *Hornberger et al.*, 1991], and therefore they are certainly not valid at the model grid scale. With huge number of parameters (particularly for distributed models), estimating their values from local measurements continues to be a major problem [e.g., *Gupta et al.*, 2004]. Therefore hydrologic models usually resort to calibration strategies often leading to unrealistic values for the parameters. Further, the calibration is carried out with over emphasis on the outlet hydrographs, which do not provide adequate information regarding the hydrologic response of a basin [e.g., *Hooper et al.*, 1988; *Jakeman and Hornberger*, 1993].

The scale issue in hydrological modeling is partially responsible for the lack of consensus regarding the effect of spatial variability in processes such as rainfall and infiltration on the streamflow prediction. For instance, based on simulations on a 287 km^2 Friends Creek watershed, *Beven and Hornberger* [1982] suggested that correct representation of spatial variability of rainfall is important for accurate prediction of streamflow hydrographs. Using a semi-distributed model on a 71 km^2 watershed in southeast France, *Obled et al.* [1994] concluded that the hydrographs are not very sensitive to the spatial variability in rainfall. Based on simulations on a 10.5 km^2 Wye River basin, *Shah et al.* [1996] reported that spatially averaged rainfall could result in streamflow predictions with reasonable accuracy under wet antecedent moisture conditions. *Arnaud et al.* [2002] used hydrologic model with three different runoff generation mechanisms on fictitious basins of sizes ranging from 20 km^2 to 1500 km^2 and concluded that runoff response varies significantly between spatially uniform and spatially distributed rainfall. Based on the results from the Distributed Model Intercomparison Project (DMIP) [*Smith et al.*, 2004a], *Smith et al.* [2004b] suggested that distributed models are amenable only when the basins cannot effectively filter the rainfall variability. For basins of size ranging from 50 km^2 to 2240 km^2, *Saulnier and Le Lay*

[2009] concluded that the accurate estimation of rainfall intensity distribution and localization is more important than the areal rainfall.

Due to the aforementioned scale and calibration issues, we avoid the use of popular distributed models based on sub-basin or grid-based discretization for our investigations on the peak flow scaling behavior. Instead, we adopt a systematic diagnostic framework to further understand the role of rainfall variability and hillslope characteristics in the scaling structure of peak flows from river basins in Iowa and Kansas. In the next section, we describe the basic concepts in the scaling theory of floods. A brief review of the relevant literature is presented in Section 1.3 followed by statement of objectives in Section 1.4. Section 1.5 outlines the methodology followed in this thesis.

1.2 Basic Concepts

Herein, we give a brief overview of the basic concepts in scaling-based analyses that we encounter at various points in this thesis. The Horton ratio R_X is defined as a ratio of the averages $E[X_{\omega}+1]/E[X_{\omega}]$, where X_{ω} is a generic random field indexed by Horton order ω, a stream ordering system developed by *Horton* [1945] and later modified by *Strahler* [1952,1957]. For instance, the field X can be the upstream areas or width function maxima or peak flows. For more details on the Horton order and the Horton ratios, please see *Rodriguez-Iturbe and Rinaldo* [1997] and *Peckham and Gupta* [1999].

The width function of a river network is a measure of the river network branching structure. There are basically two types of width functions: topologic and geometric. Throughout this study, we employ the topologic width function, which is defined as the number of links which are s links upstream of the outlet of the basin as a function of s [e.g., *Veitzer and Gupta*, 2001]. Under idealized conditions of runoff generation and constant flow velocity, the width function represents the response of the river network to spatially uniform instantaneous rainfall. The statistical structure of the width function

and its relation to the hydrologic response of the basin has been the object of several recent studies [e.g., *Veitzer and Gupta*, 2001; *Lashermes and Foufoula-Georgiou*, 2007; *Moussa*, 2008; *DiLazarro*, 2009]. *Veitzer and Gupta* [2001] showed that the width function maxima of the simulated random self-similar channel networks follow distributional simple scaling. That is, the generalized Horton law in terms of probability distributions [e.g., *Peckham and Gupta*, 1999] holds for the width function maxima, and the Horton ratios of width function maxima R_Θ and upstream areas R_A are related by a power law of the form.

$$R_\Theta = R_A^\beta \tag{1.1}$$

where β is the scaling exponent of the width function maxima. Similarly, the Horton ratios for the peak flow R_Q and upstream areas R_A are related by a power law

$$R_Q = R_A^\phi \tag{1.2}$$

when peak flow distributions exhibit statistical self-similarity, which has been shown to be the case under certain conditions of flow and rainfall [e.g., *Mantilla*, 2006]. The exponent Φ in equation 2 is referred to as the peak flow scaling exponent. A scale break is defined in this work as a transition point in the log-log plot of peak flows vs. drainage areas.

1.3 Brief Review of Literature

In the past two decades, numerous simulation and data-based studies were conducted to determine the physical basis of scale-invariance [e.g., *Gupta and Dawdy*, 1995; *Robinson et al.*, 1995; *Gupta et al.*, 1996; *Bloschl and Sivapalan*, 1997; *Robinson and Sivapalan*, 1997; *Menabde et al.*, 2001; *Menabde and Sivapalan*, 2001; *Ogden and Dawdy*, 2003; *Furey and Gupta*, 2005; *Furey and Gupta*, 2007]. A general consensus emerging from these studies is that the rainfall, channel network topology, hydraulic geometry of the channels and the hillslope runoff generation mechanisms play key roles

in determining the scaling exponents of the power laws in peak flows. While most of the research on peak flow scaling behavior was related to the annual peak flows, there has been a recent shift towards investigating single-event peak flows [e.g., *Gupta et al.*, 1996; *Ogden and Dawdy*, 2003; *Furey and Gupta*, 2005; *Mantilla et al.*, 2006; *Furey and Gupta*, 2007]. Understanding scale-invariance of the peak flows in terms of physical processes is easiest at the scales of single events. Recent studies also suggested that the scaling exponents of annual peak flows are related to those of single-event peak flows [e.g., *Ogden and Dawdy*, 2003; *Gupta et al.*, 2007]. In the following subsections, we give a brief description of key results in the literature related to the statistical structure of single event peak flows. We first discuss basic concepts and then proceed to simulation-based and data-based studies.

1.3.1 Simulation-Based Studies

Gupta et al. [1996] was the first study to focus on the effect of rainfall and channel network on the scale-invariance of single-event peak flows from a deterministic Peano network. Using a numerical simulation framework, they showed that peak flows exhibit simple scaling for uniform rainfall, with the scaling exponent dependent on the fractal dimension of the channel network width function maxima. For spatially variable rainfall, they reported that the peak flows display multi-scaling, with the exponent being a function of the channel network characteristics and the spatial variability of the rainfall. *Troutman and Over* [2001] derived analytical expressions for channel networks and rainfall mass exponents for the general class of recursive replacement trees and instantaneous multifractal rainfall. *Menabde et al.* [2001] focused on the attenuation due to storage in channel networks and its effect on the scaling exponents of peak flows from deterministic (Mandelbrot-Viscek and Peano networks) and random self-similar networks with linear routing and for spatially uniform rainfall. For the deterministic self-similar networks (SSNs), the scaling exponent of peak flows is smaller than the one predicted for

the width function maxima (i.e., ignoring the attenuation due to storage in channel networks). *Menabde et al.* [2001] also showed that for random SSN with smaller bifurcation ratios, the peak flows scale asymptotically.

To better understand and predict the scaling behavior of peak flows, *Menabde and Sivapalan* [2001] introduced a dynamic and spatially distributed hillslope-link rainfall-runoff model based on representative elementary watershed (REW) consisting of three main components: a space-time model of rainfall, a hillslope model and a channel network model. The rainfall model can generate storms whose spatial structure is characterized by a discrete random cascade. The hillslope model partitions the rainfall into Hortonian runoff, subsurface flow and evaporation, which are assumed to be zero during periods of rainfall. They further assumed that all of the surface runoff reaches the channel instantaneously. The channel network is a deterministic Mandelbrot-Viscek network in which the hydraulic geometry properties at every link are obtained from observed empirical relationships. They investigated the effect of rainfall on the scaling structure of the peak flows, starting from a spatially uniform rainfall scenario and moving to the individual storms based on discrete random cascade. They also extended the study to include continuous rainfall and annual flood peaks. The results from event-based simulations with spatially uniform rainfall and the rainfall based on the random cascade model demonstrated that the interplay between the catchment response time and the storm duration controls the scaling exponent of peak flows.

Mantilla et al. [2006] discussed the difficulties in generalizing the scaling theory to the real networks and tested whether the random spatial variability of the real channel networks and their hydraulic geometry properties, coupled with flow dynamics, produce Hortonian scaling in peak flows. Based on the results from *Veitzer and Gupta* [2001], the value of the scaling exponent of the network width function was computed for the 149 km^2 Walnut Gulch basin in Arizona [e.g., *Goodrich et al.*, 2008]. The runoff rates were estimated from two very small gauged sub-basins within the Walnut Gulch, assuming

that rainfall was spatially uniform. For an instantaneously applied runoff rate, the system of ordinary differential equations describing the runoff dynamics was solved for three different scenarios: (a) constant velocity (b) constant Chezy and (c) spatially varying Chezy constant. They showed that the scaling exponent of peak flows is larger than the exponent of the width function maxima, which contradicted the results from the studies performed on the idealized basins, where the flow scaling exponent is always smaller than the exponent of the width function maxima. The contradiction is explained in terms of the relative roles of flow attenuation and flow aggregation in the river networks that were considered.

1.3.2 Data-based studies

Ogden and Dawdy [2003] investigated the single-event and annual peak flows from the 21.2 km^2 Goodwin Creek watershed in Mississippi [e.g., *Alonso and Bingner*, 2000], where the Hortonian mechanism of runoff generation is dominant. They considered 279 events for which flows were recorded at several interior gauging stations. The results showed that the peak flows follow simple scaling but the exponents vary from event to event and depend on the runoff production efficiency. The mean of scaling exponents is 0.831 with a standard error of 0.10. Some events are then filtered out with a threshold on correlation coefficient (0.93) between logarithm of peak flows and the upstream areas. The mean of scaling exponents from 226 remaining events is equal to 0.826, with a standard error of 0.047 and a mean correlation coefficient of 0.98.

Furey and Gupta [2005] explained this event-to-event variability in the peak flow power laws in Goodwin Creek watershed in terms of variability in the rainfall's excess depth and the duration. To understand the physical origin of the observed peak flow scaling, *Furey and Gupta* [2007] proposed and applied a 5-step framework on the Goodwin Creek watershed. *Gupta et al.* [2007] provided further observational evidence on scaling in single-event peak flows for the Walnut Gulch basin, Arizona. They

reported two different sets of scaling exponents for smaller and larger scales with a scale break at around 1 km^2. They also noticed that for the events that cover almost the entire basin, the single-event scaling exponents are quite close to the scaling exponents of the annual flood quantiles.

All the studies discussed in this section focused on the fundamental question, "How is the peak flow scaling exponent linked to the channel network characteristics such as width function maxima and variability in the rainfall?" In the studies that addressed this question using numerical simulations under idealized conditions, the complexity in the simulations increased from *Gupta et al.* [1996] to *Mantilla et al.* [2006]. The rainfall varied from spatially uniform to the complex cascade-based case, and the networks ranged from deterministic self-similar to random self-similar and actual river networks with linear and nonlinear routing mechanisms [e.g., *Gupta et al.*, 1996; *Veitzer and Gupta*, 2001; *Troutman and Over*, 2001; *Menabde et al.*, 2001; *Menabde and Sivapalan*, 2001; *Mantilla et al.*, 2006]. In the data-based analyses [e.g., *Ogden and Dawdy*, 2003; *Furey and Gupta*, 2005; *Gupta et al.*, 2007; *Furey and Gupta*, 2007], the variability in the scaling exponents were explained in terms of variability in antecedent conditions and storm characteristics. However, the smaller size of the basins (21.2 km^2 Goodwin Creek and the 149 km^2 Walnut Gulch basins) limited the range of scales available to explore the effect of rainfall variability on the peak flow scaling structure. Regardless of the approach followed, these studies enhanced our understanding of the relationship between the statistical structure of flood peaks and the characteristics of rainfall and channel network. However, there is further need to understand and generalize the role that rainfall plays in the statistical structure of peak flows from actual river basins across a range of scales.

1.4 Objectives

The broad motive of this study is to contribute towards the development of a comprehensive geophysical theory of floods. The specific goal is to investigate the sensitivity of the spatial scaling structure of peak flows to the variability of rainfall for different routing and hillslope characteristics. We investigate the effect of rainfall variability on the scatter, the scale break and the regression exponent for the power law fitted to the peak flows vs. upstream areas on the higher side of the scale break. Remotely sensed rainfall products (e.g., radar-rainfall products) are being increasingly employed to study the spatial structure of floods due to their large area coverage and high resolution. Another objective of this study is to investigate the effect of rainfall estimation errors, which are inherent in the remotely sensed rainfall products, on the spatial characterization of rainfall and peak flows. Rigorous understanding of the role of rainfall on the scaling structure of peak flows provides the basis for the scaling based framework to predict the peak flows from real basins. The current study represents a step forward and improves our understanding of the relative roles of rainfall and the channel network on the scale-invariance of peak flows.

1.5 Methodology

We follow systematic simulation-based framework for most of the study. We perform a series of simulation experiments to investigate the sensitivity of the scaling behavior to the rainfall characteristics. We also investigate the sensitivity of the scaling behavior to linear and nonlinear channel routing mechanisms. We selected the simulation framework instead of a data-based analysis since it allows complete freedom to systematically explore various aspects of scale-invariance. Also, there are very few basins in the United States where streamflow data necessary for rigorous scaling analyses are available. Our simulation covers a range of scale from ~0.1-7250 km^2, thus addressing the peak flow scaling for basin response times ranging from minutes to days.

Throughout the thesis, we employ CUENCAS [e.g., *Mantilla et al.* 2005], a distributed hydrologic model with a hillslope-link discretization scheme. In the latter part of the thesis, we employ rainfall datasets from Iowa and Kansas and perform data-based analysis.

In Chapter 2, we briefly describe the hydrological model and illustrate why considering individual hydrographs at the outlet of a basin can lead to misleading interpretations of the effects of rainfall variability. We then investigate the hydrologic response of a basin to the simple scenario of spatially uniform rainfall of varying intensities, durations and storm advection velocities. In Chapter 3, we investigate the role of storm variability characterized by variance, spatial correlation and intermittency. We employ realistic space-time rainfall fields obtained from a popular rainfall model that combines the aforementioned features. All the simulation experiments in Chapters 2 and 3 are performed on ~1100 km^2 Whitewater River basin in Kansas.

In Chapter 4, we change the river network to Iowa River basin, which is an order of magnitude larger than the Whitewater River basin, and repeat the analysis with spatially uniform and realistic rainfall fields. Besides the role of rainfall variability, Chapter 4 also focuses on the effect of channel network extraction, and the hillslope characteristics on the scaling behavior of peak flows. For all the scenarios, a recent formulation of flow velocity for a network of channels is utilized to obtain the peak flows and the scaling exponents. The exponents are in turn compared to the scaling exponent of the width function maxima. The results obtained from simple scenarios are used to make rigorous interpretations of the peak flow scaling structure obtained from rainfall simulated from the space-time rainfall model.

It is well known that the variability in the remotely sensed precipitation fields is a combination of natural variability and the rainfall estimation errors. Therefore, it is equally important to quantify how the rainfall estimation errors affect the spatial characterization of rainfall fields and then propagate into the peak flow scaling

exponents. Chapter 5 presents a theoretical framework for estimating the radar-rainfall error spatial correlation (ESC) using data from relatively dense rain gauge networks. The method is analogous to the error variance separation that corrects the error variance of a radar-rainfall product for gauge representativeness errors. We then apply the method to estimate the ESC of the National Weather Service's standard hourly radar-rainfall products, known as digital precipitation arrays (DPA). However, to quantify the effect of errors, we need a comprehensive error model capable of characterizing not only the first and second-order statistics of errors, but also their conditional distributions.

In Chapter 6, we employ an in-house radar-rainfall error model developed by *Ciach et al.* [2007] and study the impact of radar-rainfall estimation errors on the spatial characterization of ten warm season rainfall events over Wichita, Kansas. We then propagate the ensemble of probable rainfall fields through the hydrologic model and obtain the peak flow scaling structure. Chapter 7 presents preliminary results regarding the effect of radar-rainfall errors on the statistical structure of peak flows. In Chapter 8, we summarize the thesis and present concluding remarks including a discussion on the implications of this study with respect to developing scaling-based geophysical flood forecasting tool. In the Appendix A, we quantified the effects of sampling on the estimation of spatial correlation, the statistic, which is used throughout the study.

CHAPTER 2

EFFECT OF RAINFALL INTENSITY AND DURATION: SIMULATIONS WITH SPATIALLY UNIFORM RAINFALL*

2.1 Introduction

In this chapter, our goal is to clarify the role of basic characteristics of rainfall such as storm intensity, duration and advection velocity on the scaling structure of peak flows. The chapter is organized as follows. Section 2.3 provides a brief description of the study area. In Section 2.3, we present the rainfall-runoff model based on decomposition of the landscape into hillslopes and channel links. Section 2.4 lists the assumptions in the hydrologic model. In Section *2.5, we illustrate with a simple simulation experiment, how examining the basin response in terms of outlet hydrograph can be misleading. Section 2.6 describes the simulation results related to the sensitivity of the peak flow scaling behavior to the basic characteristics of rainfall. In the Section 2.7, we present an analysis of scatter seen in the scaling structure of peak flows, followed by our conclusions in Section 2.8.

2.2 Study Area - Whitewater River basin

The Whitewater River basin (Figure 2.1), with an area of 1100 km^2, stretches between latitudes 37° 46'E and 38° 09'E and longitudes 96° 51'W and 97° 18'W. The river network extraction was based on the maximum gradient method, also known as the D8 algorithm [e.g., *O'Callaghan and Mark*, 1984]. *Mantilla and Gupta* [2005] compared the network extracted from CUENCAS with those extracted from popular GIS software such as ArcInfo, GRASS and RiverTools and found no major differences when high resolution DEMs were used. They showed that a 30m resolution DEM is sufficient to

*Adapted from Mandapaka, P.V., W.F. Krajewski, R. Mantilla, and V. K. Gupta (2009), Dissecting the effect of rainfall variability on the statistical structure of peak flows, *Advances in Water Resources*, 32, 1508-1525.

extract the drainage network that is close to the terrain's actual network. We use the one arc-second resolution (~30m) digital elevation model (DEM) from USGS to extract the channel network. This results in some 20,000 hillslopes and, thus, channel links for this basin. In Figure 2.1, we show the extracted channel network with links of Horton orders 4 to 7.

Section 1.2 indicated that the width function maxima play an important role in understanding the scaling structure of the peak flows. Figure 2.2 shows the topologic width function for the outlet of the Whitewater River basin. We estimated the width functions for each link in the river network and obtained the corresponding maximum values. Figures 2.3(a) and 2.3(c) show the Horton plots for drainage areas and width function maxima of links of various orders for the Whitewater River basin, Kansas. If the channel network is self-similar, the averages of drainage areas and width function maxima display linearity with respect to the corresponding Horton orders in the log-linear domain [e.g., *Strahler*, 1957; *Peckham and Gupta*, 1999; *Furey and Troutman*, 2008). The log-linearity in Figures 2.3(a) and 2.3(c) confirm the statistical self-similarity of the upstream areas and the width function maxima. In the regression analysis, we use the areas and width function maxima corresponding to the Horton orders 2 to 6. The order 7 stream is not used in the Horton regression due to sampling reasons: we have only one point corresponding to the order 7. Although, averages corresponding to order 1 streams do not suffer from sampling issues, they are usually not considered in the regression [e.g., *Peckham and Gupta*, 1999; *Mantilla and Gupta*, 2005] as they represent the finest detail in a stream network, and therefore the corresponding basins do not contain a "network". The Horton ratios for the areas and width function maxima are then obtained by exponentiation of the slopes from the regression analysis. The scaling exponent of width function maxima obtained through Horton ratios in (1) is 0.49.

If the upstream areas and width function maxima display log-linearity, as shown in Figures 2.3(a) and 2.3(c), then $E[X_\omega] = E[X_1] \bullet (R_X)^{\omega-1}$, where X is either the upstream

area or the width function maxima and R_X is the corresponding Horton ratio. The rescaled upstream areas and width function maxima are obtained by dividing each value of X_ω by $E[X_1]\bullet(R_X)^{\omega-1}$. The probability distribution of the quantity $X_\omega/[E[X_1]\bullet(R_X)^{\omega-1}]$ is called the rescaled probability distribution. In Figures 2.3(b) and 2.3(d), we show the statistical self-similarity of areas and width function maxima in terms of their rescaled probability distributions for orders 1 to 5. Although order 1 basins were not considered in the regression analysis, it can be seen that their rescaled probability distribution collapses onto those of orders 2 to 5.

2.3 Hydrologic Model

Because of the fundamental effect of the river network structure on peak flows, it is necessary to have a distributed hydrologic model that can calculate hydrographs for all river network links in order to carry out a systematic investigation. In this study, we used the CUENCAS model, developed by *Mantilla and Gupta* [2005], which is based on hillslope-link decomposition of the landscape and mass conservation equations [e.g., *Gupta and Waymire*, 1998]. The model can be run with linear routing with constant flow velocity throughout the channel network or nonlinear routing with velocity that depends on the discharge in each link and the corresponding upstream area. For the nonlinear case, the velocities are estimated using [*Mantilla*, 2007]

$$V_c(t) = v_R \cdot \left(\frac{q(t)}{Q_R}\right)^{\lambda_1} \cdot \left(\frac{A}{A_R}\right)^{\lambda_2} \tag{2.1}$$

where $V_c(t)$ is the velocity in the channel and A is the upstream area of the corresponding channel. The coefficients λ_1 and λ_2 are the velocity scaling exponents for discharge and upstream area, respectively, and v_R, Q_R and A_R, are reference velocity, discharge and area, whose values are taken in this study to be 1.0 m/s, 200 m^3/s and 1100 km^2. These values are obtained from measurements during the rainfall-runoff events in the Whitewater

River basin. The above equation gives the instantaneous velocity as a function of discharge *q(t)* in the channel link, which in turn gives rise to a non-linear ordinary differential equation that represents fluxes coming out of the channel link. Please see equations (6) and (9) – (11) in *Mantilla et al.* [2006] for more details.

Although the nonlinear routing mechanism is closer to reality, we also included the linear routing analysis in this study as it is a good starting point to investigate the effect of rainfall variability on the peak flow scaling structure. Throughout this study, we use a value of 0.5 m/s for the V_c for the linear routing scenario and λ_1 and λ_2 of 0.3 and -0.1 for the nonlinear routing scenario, obtained based on field data from the region. In Figure 2.4, we show the velocity obtained using (3) for the link that corresponds to the largest upstream area of each Horton order for the Whitewater River basin. Throughout this chapter, we employed a rainfall grid of size 40×40 km^2 with a spatial resolution of 1 km.

2.4 Assumptions

In all of our simulation scenarios, we assume (1) negligible evaporation; (2) purely surface runoff (i.e. no infiltration and no subsurface runoff); and (3) instantaneous flow of runoff into the channel. Evaporation rate is often an order of magnitude lower than storm rainfall rate, and Hortonian runoff generation is one of the main flood producing mechanisms. From the brief review of literature presented in Section 2, one can infer that the complexity in the simulation-based studies that were carried out to understand the scaling behavior of peak flows have steadily increased since the early nineties. For instance, one of the first studies was based on the deterministic Peano network and uniform rainfall [e.g., *Gupta et al.* 1996]. Some recent studies have used random self-similar networks to mimic the river network behavior [e.g., *Veitzer and Gupta*, 2001; *Mantilla*, 2007]. We continue on this trajectory by introducing complexity one step at a time. Therefore, in this study, the complexity is in terms of rainfall

variability and the river network structure, which is why we limited our analysis to the Hortonian runoff generation mechanism. We understand that in reality, other runoff producing mechanisms are also possible in the selected study area. The hydrologic model we used can account for the saturation excess mechanism, for instance. However, including it in the study would only add additional variability, and it is difficult to separate the role of rainfall variability and the variability introduced by the saturation excess mechanism. The third assumption regarding the instantaneous flow to the channel plays a key role in shaping the hydrologic response of smaller basins (< ~ 10 km^2) as the hillslope travel times are comparable to the time spent in the channel network [e.g., *D'Odorico and Rigon*, 2003]. But the net effect is smoothed out for larger basins. Therefore, the assumptions are reasonable in the context of exploring the roles of rainfall and channel network on the scaling exponents of peak flows, i.e., floods, for individual rainfall-runoff events.

2.5 Hydrographs vs. Scaling-based Framework

This section illustrates via simple simulation experiments the advantages of the scaling-based analysis of hydrologic response. The hydrologic model CUENCAS is forced with two simple rainfall scenarios of changing intensity (60 mm/h and 10 minutes) and duration (5 mm/h and 120 minutes), while keeping the total rainfall volume constant (1.1×107 m^3). The simulated rainfall is spatially uniform over the basin for the given duration. We also assumed that the runoff is Hortonian and reaches the channel instantaneously. The discharges are normalized with respect to the peak flow corresponding to the rainfall scenario of 60 mm/h for 10 minutes. The time of occurrence is then normalized with the time at which the normalized discharge corresponding to the scenario of 60 mm/h and 10 minutes reaches 0.01. In Figure 2.5, we show the normalized hydrographs at six different locations in the Whitewater River basin (Figure 2.1). Although we show the normalized hydrographs at only six locations, we simulated

hydrographs for all the interior sub-basins as well as for the outlet of the Whitewater River basin by solving the mass and momentum equations throughout the river network. Figure 2.5 demonstrates that at smaller scales, the values of flow peaks differ greatly from each other and occur at different instances. However, the flow hydrographs are indistinguishable as we move to the larger scales.

We then relax the spatial uniformity assumption and assume that the rainfall is randomly distributed in space over the hillslopes of that same basin. We obtained ten realizations of the rainfall following a uniform distribution over the range of 20 to 100 mm/h with the average intensity equal to 60 mm/h and the duration kept at 10 min. That is, for each rainfall field of size 40 × 40 km^2, we generated 1600 random numbers following a uniform distribution with a range of [20,100] and a mean of 60 mm/h. It should be noted that these fields do not possess any spatial correlation. In Figure 2.6, we compare the normalized hydrographs obtained with these ten rainfall fields with the one obtained for the spatially uniform case of Figure 2.5. It is clear from Figure 2.6 that for spatially random rainfall, the variability in the hydrographs at smaller scales is higher compared to those of larger scales. Therefore, to develop a comprehensive understanding of river basin response, it is imperative that we study the hydrographs throughout the basin across multiple scales.

In this context, the results from spatially uniform rainfall (Figure 2.5) can be alternatively represented in the form of Figures 2.7(a) and 2.7(b). Similarly, the results from spatially variable rainfall (Figure 2.6) for two of the simulated realizations are shown in Figures 2.7(c) and 2.7(d). This framework allows us to study the basin response across multiple scales. Figure 2.7 illustrates that our simulated peak flows display scaling structure with respect to the drainage area, and the scaling regime depends on the intensity, duration and variability of the rainfall. Figure 2.7 also demonstrates that the effect of rainfall variability on the basin response is scale-dependent. While peak flows are sensitive to the intensity, duration and spatial distribution of rainfall at small

scales (~10 km^2), the variability in rainfall is dampened at larger scales (~1000 km^2) by the river network via aggregation of flows.

2.6 Simulation Scenarios and Results

2.6.1 Sensitivity to the Intensity and Duration of Spatially Uniform Rainfall

We start with the simple scenario of a basin receiving spatially uniform rainfall of a certain intensity for a certain duration. Figure 2.8 shows the peak flows versus drainage areas for different rainfall intensities and durations with a linear channel routing mechanism. Three important features of the peak flow scaling structure apparent in the plots are the scatter, the scale break and the scaling exponent. For a fixed rainfall intensity, the scatter decreases as the duration of the event increases. For each link, there is an upper limit for the peak flow that is not exceeded. This upper limit corresponds to the equilibrium discharge reached when the rainfall duration is larger than the concentration time. With the increase in the duration of rainfall, more hillslopes reach saturation, thereby decreasing the scatter. The peak flows for the links that reached steady-state correspond to the well known rational method ($Q = c\ I\ A$). We obtain the scale break by comparing the peak flows obtained from our simulations to those from the above rational method equation. A window of fixed size in the logarithmic domain is moved along the upstream area axis of each panel in Figure 2.8. Within such a window, we compute the following ratio

$$\chi = \frac{\sum_{i=1}^{n_p} I_{p,i}}{n_p} \tag{2.2}$$

where, $I_{p,i} = 1$ if $0.9Q_{rat} \le Q_{p,i} \le Q_{rat}$, $Q_{p,i}$ is the peak flow for the link i, Q_{rat} is the corresponding peak flow obtained from the rational method and n_p is the total number of links in the network. If the ratio χ is less than 0.75, the scale break is considered to be

at the average of the upstream areas within that window. We realize that this definition of scale break is subjective. However, it serves the purpose of a qualitative comparison only. The scale break is indicated by a red line in Figure 2.8. As the duration of the rainfall increases, more links reach saturation and the scale break moves towards the larger areas. Because of the large scatter, we do not estimate the scale break for the shortest duration of 5 minutes.

Figure 2.8 shows that for a fixed duration, the scale break and scatter in the peak flow scaling structure remain unchanged with intensity. Similar to the Section 2.3, we fit these peak flows in the Hortonian framework using equation 1.2. In Figure 2.9, we show the Horton plots of peak flows for all the intensities and durations. The Horton ratio of peak flows is estimated considering only the orders that lie on the higher side of the scale break. Since the scale break is not obvious in the Horton plots, we select the orders for regression based on Figure 2.8. Horton ratio of peak flows obtained by exponentiation of the regression slope is shown in each panel of Figure 2.9. We obtained ordinary least squares fit to the peak flows, but in the Hortonian regression framework, similar to the width function analysis presented in the previous section. The Horton ratio of peak flows and the upstream area are plugged into equation 1.2 to obtain the scaling exponent of the peak flows. That is, instead of obtaining the scaling exponent by regression of peak flows with upstream areas, we use equation 1.2 to obtain the scaling exponent. The coefficient of the power law is obtained so that the regression line passes through the average of the peak flows corresponding to the top three orders (Figure 2.8). The regression equations in Figure 2.8 allow us to conclude that, for a fixed duration, the peak flows are linearly related to the rainfall intensities when the routing mechanism in the channels is linear. This result is similar to the one observed by *Furey and Gupta* [2005] over the Goodwin Creek Watershed. For a fixed intensity, the scaling exponents range from 0.50 to 0.56 as the duration changes from 5 to 360 minutes. We have also noticed that the peak flow at the outlet of the basin changes linearly with the duration. The

scaling exponents for all the cases are larger than the width function scaling exponent of 0.49, which confirms the result of *Mantilla et al.* [2006] for the Walnut Gulch watershed.

Figure 2.10 shows the effect of the intensity and duration of spatially uniform rainfall when the channel routing mechanism is nonlinear. The parameters selected in this study for the nonlinear routing mechanism result in different velocities in different links, and a straightforward panel to panel comparison between Figures 2.8 and 2.10 is therefore not meaningful. We do not obtain the scale break for the shortest duration simulations for the same reason mentioned in the linear routing case. Regression is not performed for the longest duration of 360 minutes, as most of the links have reached saturation and the scale break is not well defined. For the 120 minute duration, the fitted regression equations reveal that the scaling exponent decreases as the rainfall intensity decreases. It is also clear from Figure 2.10 that the relationship between flow peaks and rainfall intensities is nonlinear. The peak flow scaling exponents for all the cases of nonlinear routing are larger than the exponent of the width function maxima. They range from 0.55 for the shortest duration of 5 minutes to 0.66 when the rainfall intensity is 50 mm/h and the duration is 120 minutes. For the longest duration of 360 minutes, the scaling exponent is close to 1.0.

2.6.2 Sensitivity to the Advection Velocity of Spatially Uniform Rainfall

A spatially uniform rainfall of intensity 30 mm/h and approximately half the size of the basin (40×20 km^2) is moved from west to east at five different velocities (4, 8, 16, 32 and 64 km/h). The peak flows are fitted in the Hortonian framework and the regression equations are obtained. Figure 2.11 plots the scaling exponents versus advection velocities for linear and nonlinear routing mechanisms. As expected from the results shown in Figure 2.10, there is no scale break for smallest advection velocity of 4 km/h for the nonlinear routing mechanism, and thus we did not perform any regression

analysis. For both channel routing mechanisms, the scaling exponent decreases with the increase in advection velocity. Our motive for plotting linear and nonlinear routing mechanisms in the same figure is not to compare them point-to-point but to compare how the exponents decrease with advection velocity. Figure 2.11 also shows the power law fit for both routing mechanisms. The fitted equations reveal that the trend is the same for both routing mechanisms. The decreasing trend can be explained in terms of the effect of duration discussed in the previous subsection. With the increase in advection velocity, the duration for which the block of rainfall stays over the basin decreases, and therefore the scaling exponent also decreases.

2.6.3 Sensitivity to the Spatio-Temporal Variability

We have thus far assumed that the rainfall is spatially uniform throughout the basin. In this subsection, we investigate the effect of spatio-temporal variability on the scaling exponents of peak flows. We relax the spatial uniformity of rainfall over the basin by breaking it into two components: a block of uniform rainfall with an intensity of 25 mm/h for a duration of 30 minutes on the western half of the basin and a rainfall of 50 mm/h for a duration of 120 minutes over the eastern half of the basin. The channel network routing is assumed to be linear with a velocity of 0.5 m/s throughout the network. Figure 2.12 shows the scaling structure of peak flows for this scenario. The banded structure apparent in Figure 2.12 is a direct manifestation of the different rainfall intensities received by western and eastern regions of the basin. The scale break for these two bands also occurs at different locations because of the different durations for which the rainfall lasts over the western and eastern parts of the basin. We noticed a similar trend for the nonlinear routing in channels.

2.7 Analysis of Scatter

The scatter in the peak flow scaling structure for the lower order basins can be explained in terms of peak flows reaching equilibrium (basins reaching saturation), while

the scatter for the higher order basins can be explained in terms of aggregation and attenuation of flows. We illustrate this by analyzing the peak flows for two of the rainfall scenarios in Section 2.6.1 and Figure 2.8. Specifically, we compare the probability distributions of the rescaled peak flows corresponding to spatially uniform rainfall of 5 mm/h and durations of 5 and 120 minutes with the probability distributions of rescaled areas and width function maxima. It can be seen from the Figure 2.13(a) that for a spatially uniform rainfall intensity of 5 mm/h and duration of 120 minutes, the order 1 probability distributions of rescaled peak flows and areas are indistinguishable (negligible scatter in Figure 2.8) but for shorter duration of 5 minutes, the probability distributions are very different (large scatter in Figure 2.8). Since the width function has the signature of the aggregation of flows in the channel network, we compare the order 5 probability distributions of rescaled peak flows and width function maxima in Figure 2.13(b). We intended to compare the probability distribution of peak flows with that of width function maxima for higher order basins. However, for orders 6 and 7, there are not enough points to obtain the probability distributions. Therefore, we limit the comparison to order 5 basins. Unlike the order 1 distribution, the order 5 distribution is not very sensitive to the duration of the rainfall event. For both the rainfall scenarios, the distribution of rescaled peak flows match reasonably well with that of width function maxima (Figure 2.13b).

2.8 Summary and Conclusions

In this chapter, we discussed the advantages of the studying the hydrologic response of a basin in a scaling-based framework. We carried out a systematic investigation to understand the role of rainfall on the spatial structure of peak flows. Due to the lack of adequate field data, i.e. numerous stream gauges as well as highly accurate rainfall maps, we used simulations. Our simulation experiments consisted of simple scenarios aimed at dissecting the effects of rainfall variability on peak flows scaling

structure. We demonstrated that rainfall variability has a different impact on the magnitude of peak flows for basins of different scales.

While our results are subject to the usual limitations of a simulation study, our experiments contain many realistic aspects. First, our river basin has a size (scale) substantially larger than many small experimental basins that are the basis for many hydrologic studies. Second, our assumptions regarding the runoff generation mechanism represent a realistic yet limited subset of actual hydrologic conditions. They are particularly realistic for flood-prone conditions, which are clearly the motivation of our study. We selected the Whitewater River basin in Kansas for this study and a distributed hillslope-link based hydrological model to obtain the peak flows for each link within the basin. The channel network extracted is characterized in terms of width function maxima. The width function maxima of the Whitewater River basin displayed scaling behavior with respect to the Horton orders. The scaling exponent of width function maxima was estimated to be 0.49.

We focused on three aspects of the peak flow scaling structure for all the scenarios: scatter, scale break and the scaling exponents. The results showed that the peak flow scaling exponents for all the scenarios considered in this study are greater than the width function scaling exponent. This result is in agreement with the hypothesis of *Mantilla et al.* [2006] that in the river networks, the peak flow scaling exponent is governed by the competition between attenuation and aggregation of the flows. For a fixed intensity, the scaling exponent increases with an increase in the rainfall duration, and for a fixed duration, the scaling exponent does not change with intensity for linear channel routing and decreases with intensity for nonlinear channel routing. For the two hour duration, the fitted regression equations reveal that the scaling exponent decreases as the rainfall intensity decreases. Based on simulations with spatially uniform rainfall of varying depths and a fixed duration of 10 minutes on a deterministic Mandelbrot-Viscek network, *Menabde and Sivapalan* [2001] reported that the scaling exponent increases as

the rainfall depth decreases. For the Whitewater River basin, and therefore a real river network, we did not notice such a trend for 30 and 5 minute duration simulations and noticed a reverse trend when the duration is 120 minutes. The results obtained from the simple scenarios contribute to a rigorous understanding of the peak flow scaling structure obtained from actual space-time variable rainfall, which we address in the following chapter.

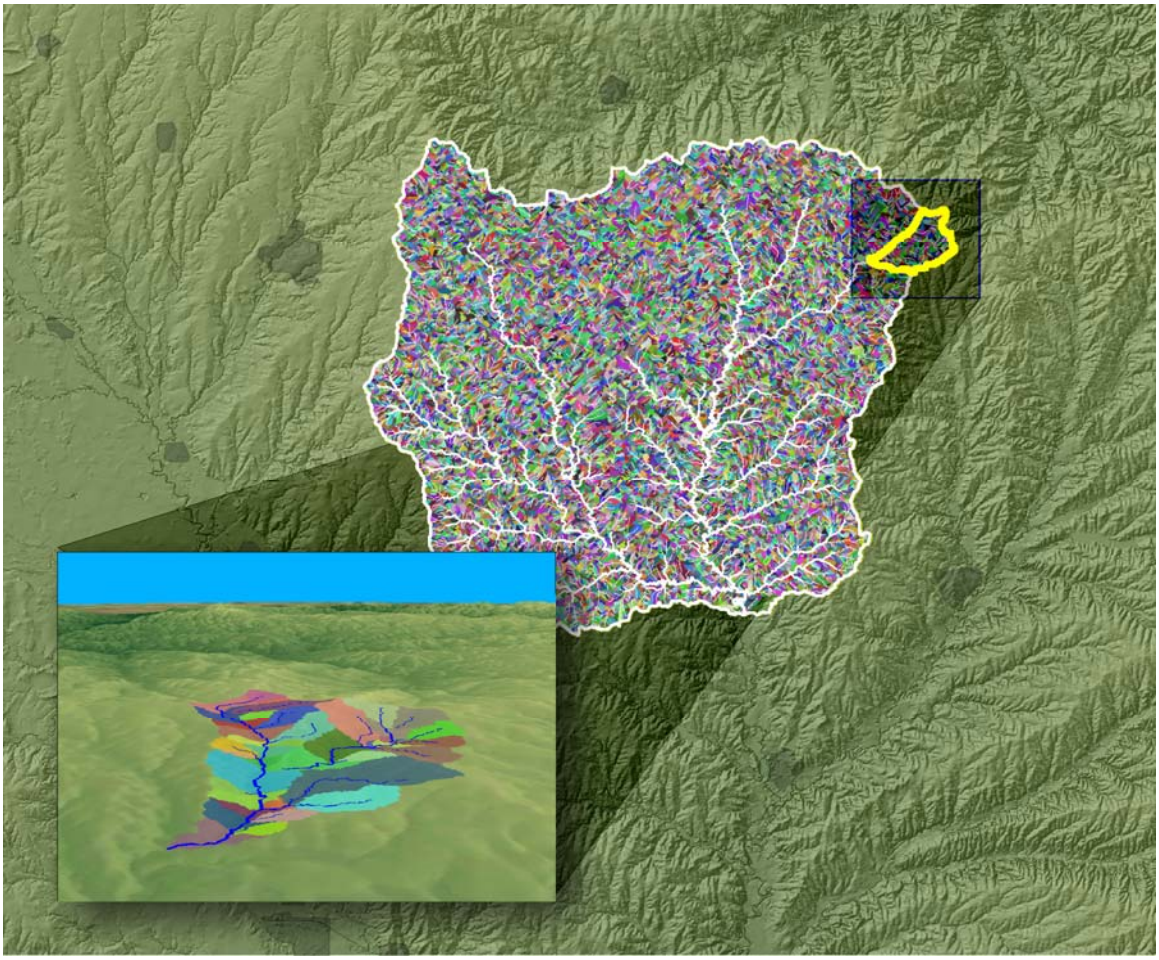

Figure 2.1 A shaded relief map of the Whitewater River basin showing the hillslope and channel link structure of the CUANCAS model. The channel network with links of order 4 to 7 is shown.

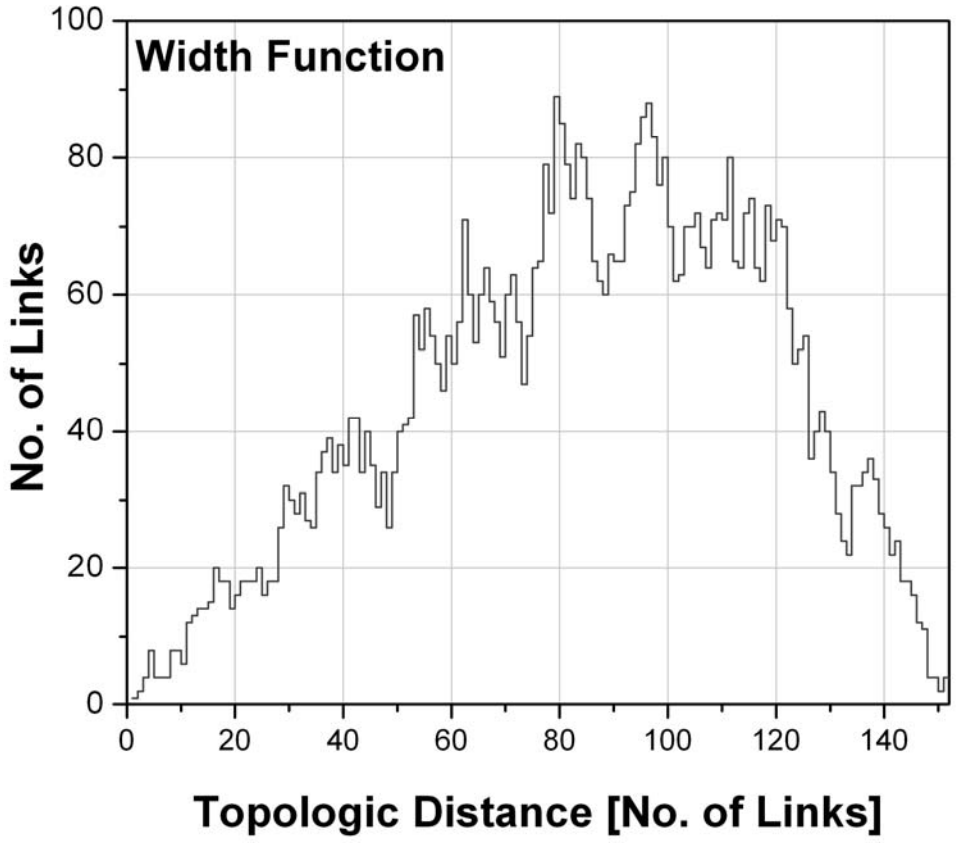

Figure 2.2 Topologic width function at the outlet of the Whitewater River basin, Kansas

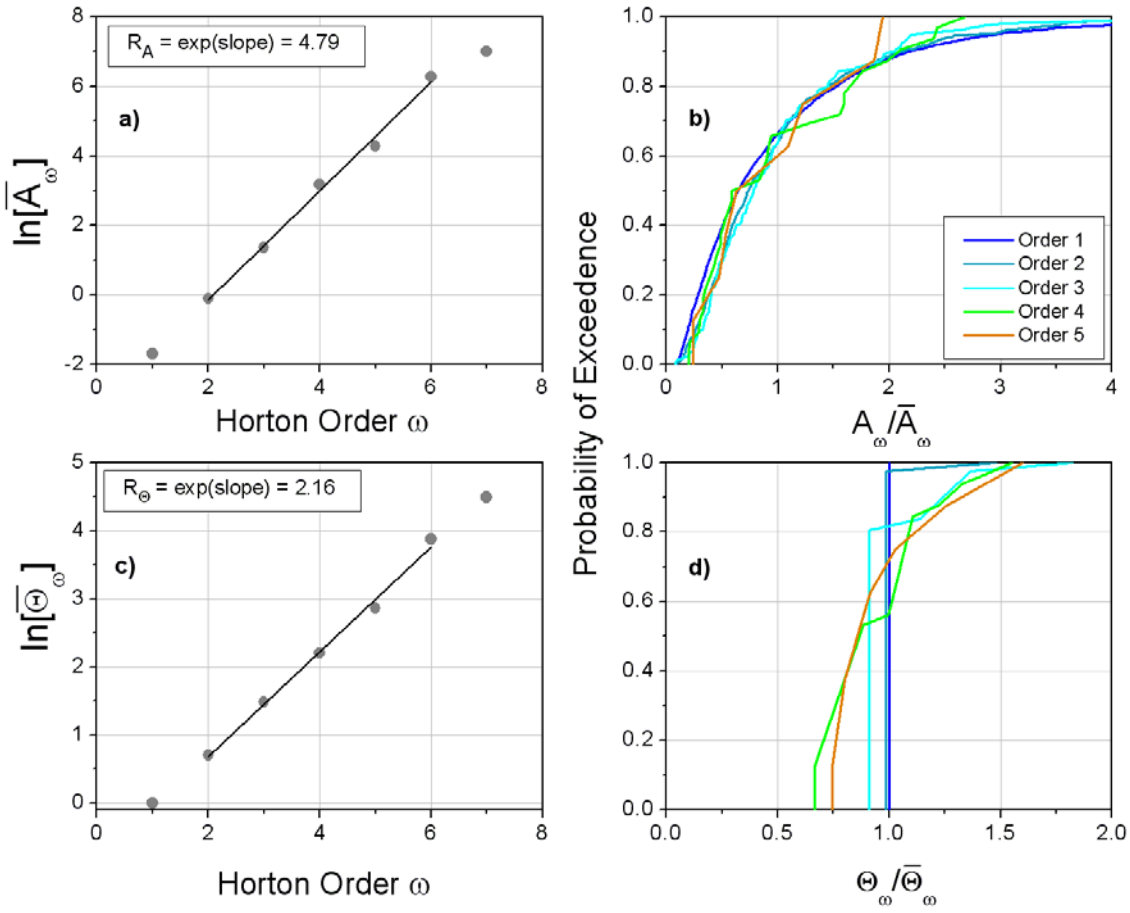

Figure 2.3 Statistical self-similarity of upstream areas and width function maxima in terms of Horton plots (left panels) and rescaled distributions (right panels). The ordinary least square regression is used to obtain the corresponding Horton ratios. The first order and seventh order links are not considered in fitting.

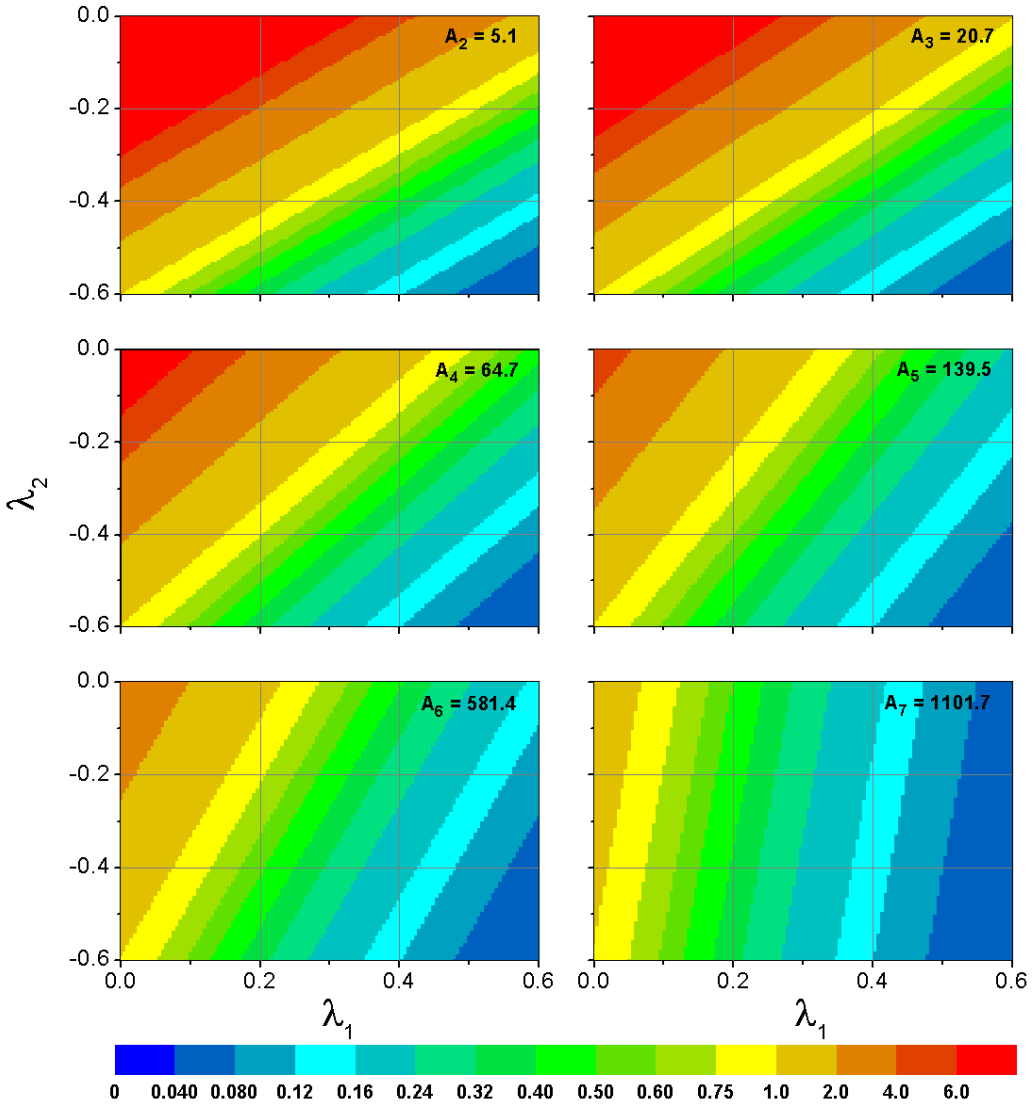

Figure 2.4 Sensitivity of the channel velocity (m/s) to the λ_1 and λ_2 in equation 2.1. The velocities are shown only for the channels that correspond to the largest are (displayed on each panel in km^2) for the Horton orders 2 to 7.

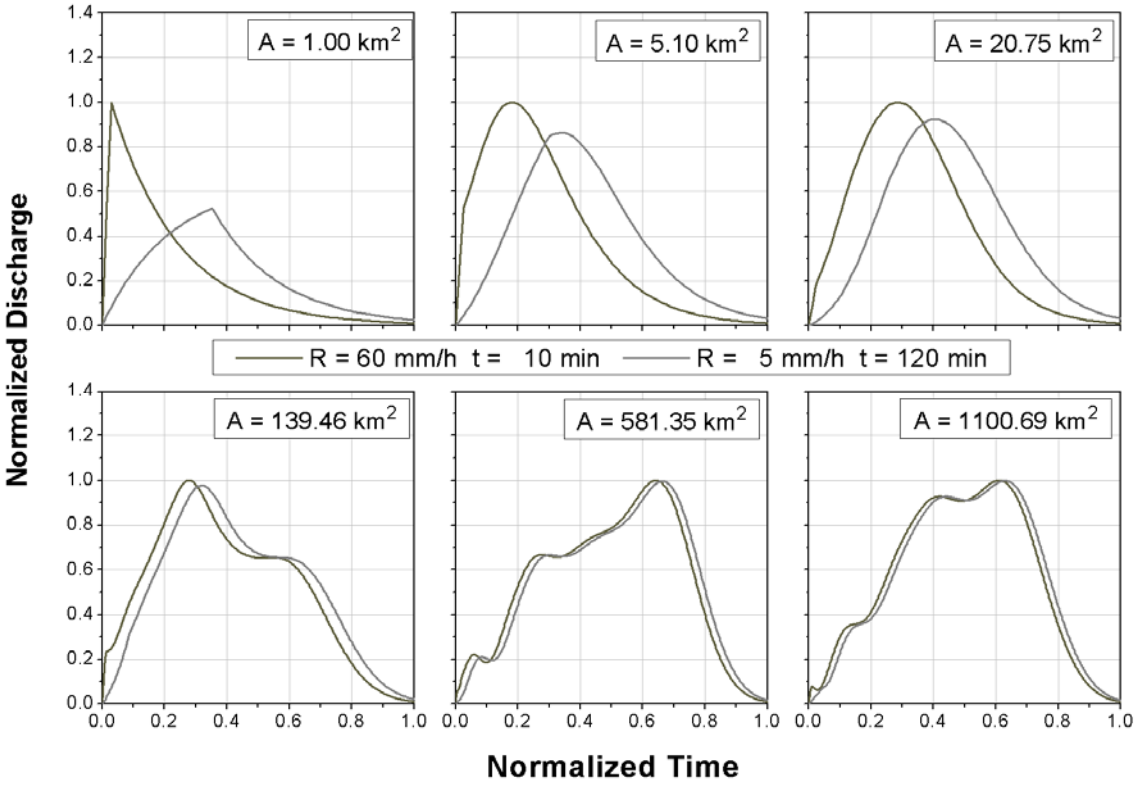

Figure 2.5 Hydrographs at six locations in the Whitewater River basin obtained from a distributed hydrologic model for a spatially uniform rainfall.

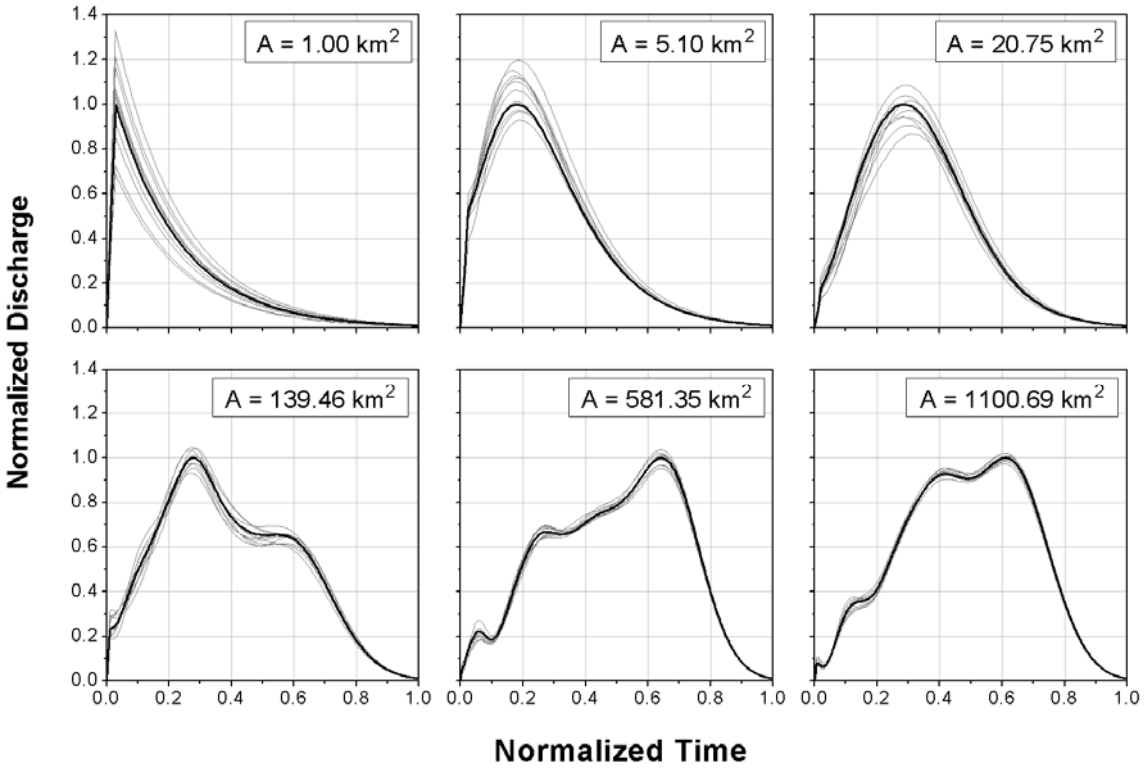

Figure 2.6 Hydrographs at six locations in the Whitewater River basin obtained from a distributed hydrologic model. The gray lines are the hydrographs for each of the 10 rainfall realizations assumed to be random in space with the intensities following the uniform distribution U[20,100] mm/h for a duration of 10 minutes. The solid line represents hydrographs for the spatially uniform rainfall of 60 mm/h for 10 minutes.

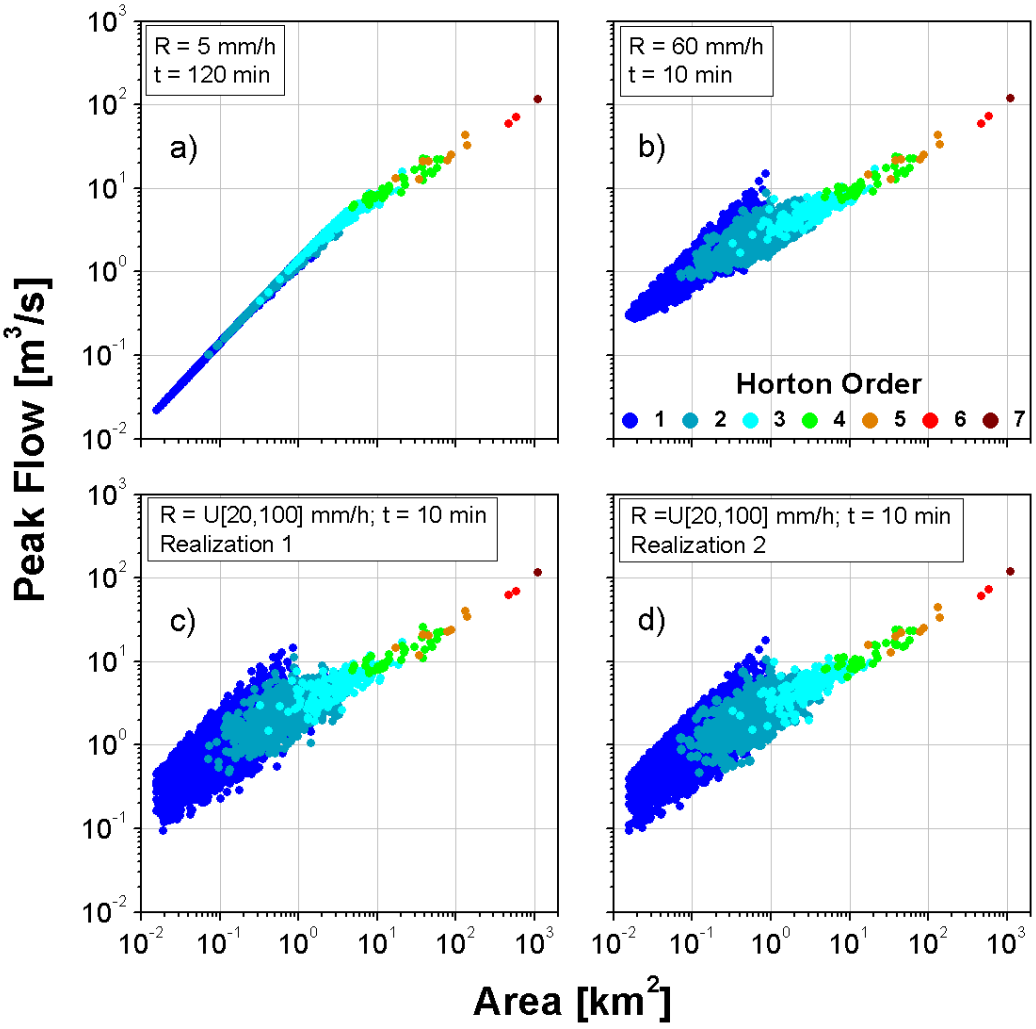

Figure 2.7 Scaling of peak flows with respect to the upstream areas of all the subbasins in the Whitewater River basin. For (a) and (b), the rainfall is spatially uniform with intensity and duration indicated on the panels. For (c) and (d), the rainfall is random in space with the intensities following the uniform distribution U[20,100] mm/h and the duration is equal to 10 minutes.

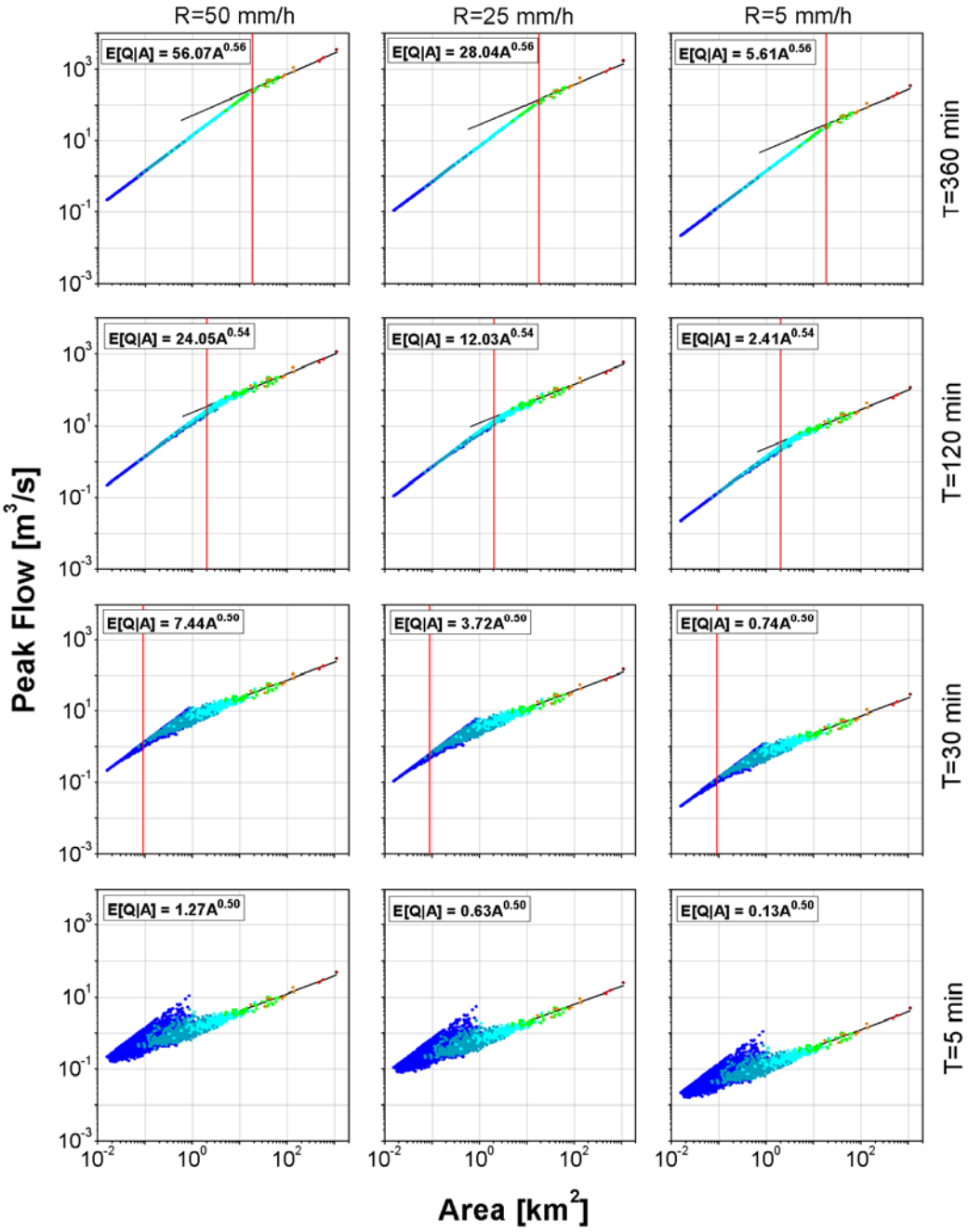

Figure 2.8 Sensitivity of peak flow scaling structure to intensity and duration of spatially uniform rainfall and linear channel routing with a velocity of 0.5 m/s. The solid black line represents the ordinary least squares fit (equation on each panel) obtained in the Hortonian framework. The solid red line indicates the scale break. The color scheme is same as Figure 2.7

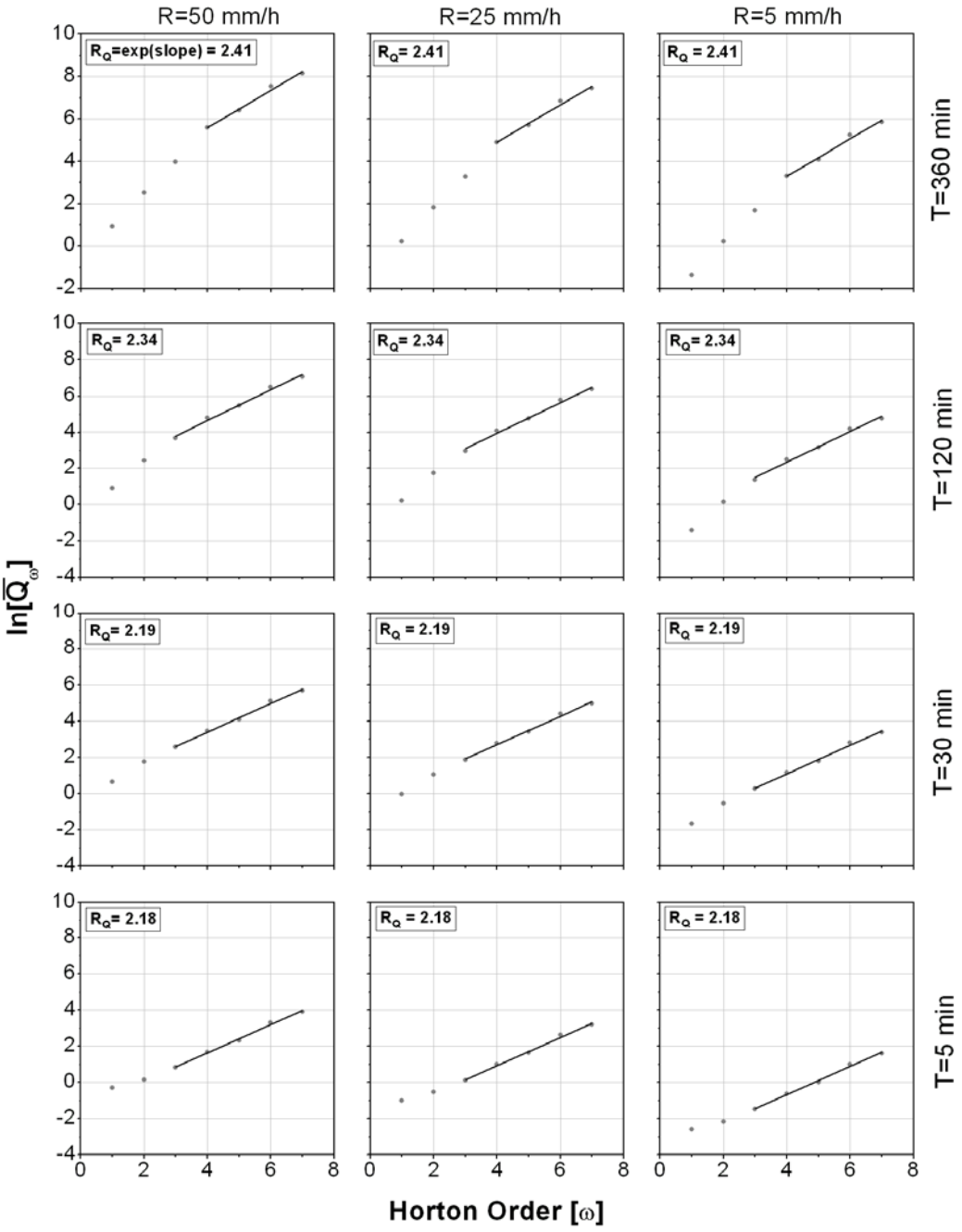

Figure 2.9 Horton plots of peak flows for different combinations of intensity and duration of spatially uniform rainfall applied throughout the basin. The solid line indicates the ordinary least squares regression fit. The corresponding Horton ratios are also indicated on each panel.

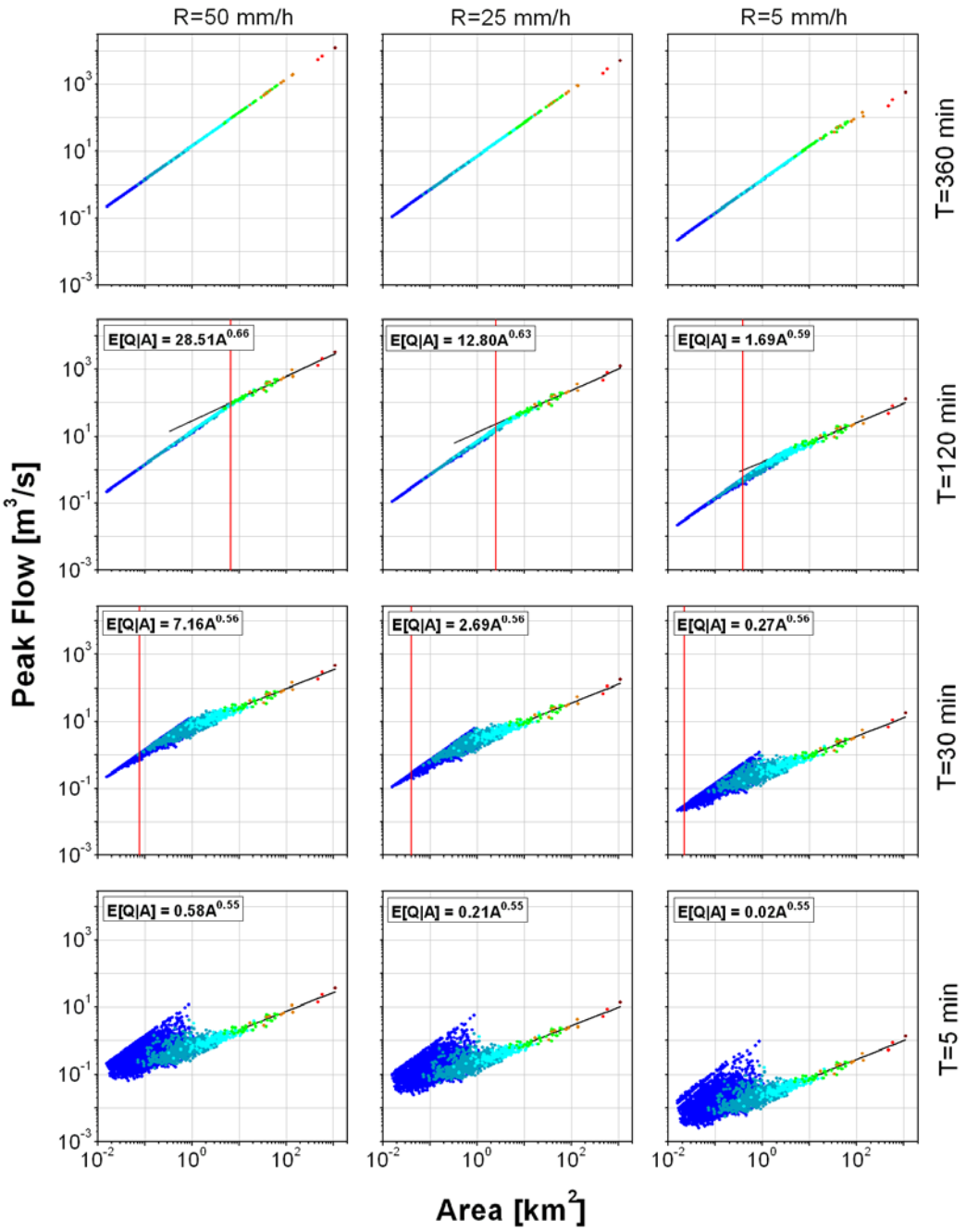

Figure 2.10 Sensitivity of peak flow scaling structure to intensity and duration of spatially uniform rainfall and nonlinear channel routing. The solid black line represents the ordinary least squares fit (equation on each panel) performed in the Hortonian framework. The solid red line indicates the scale break. The color scheme is same as Figure 2.7.

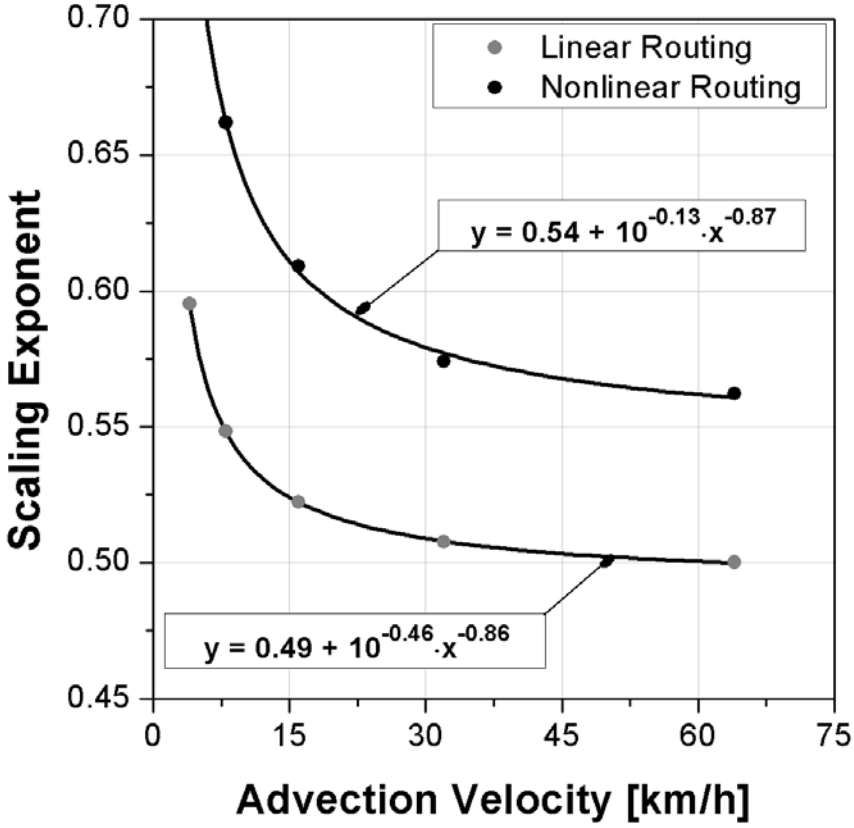

Figure 2.11 Sensitivity of scaling exponent of peak flows to the advection velocity of spatially uniform rainfall of intensity 30 mm/h and linear and nonlinear channel routing mechanisms.

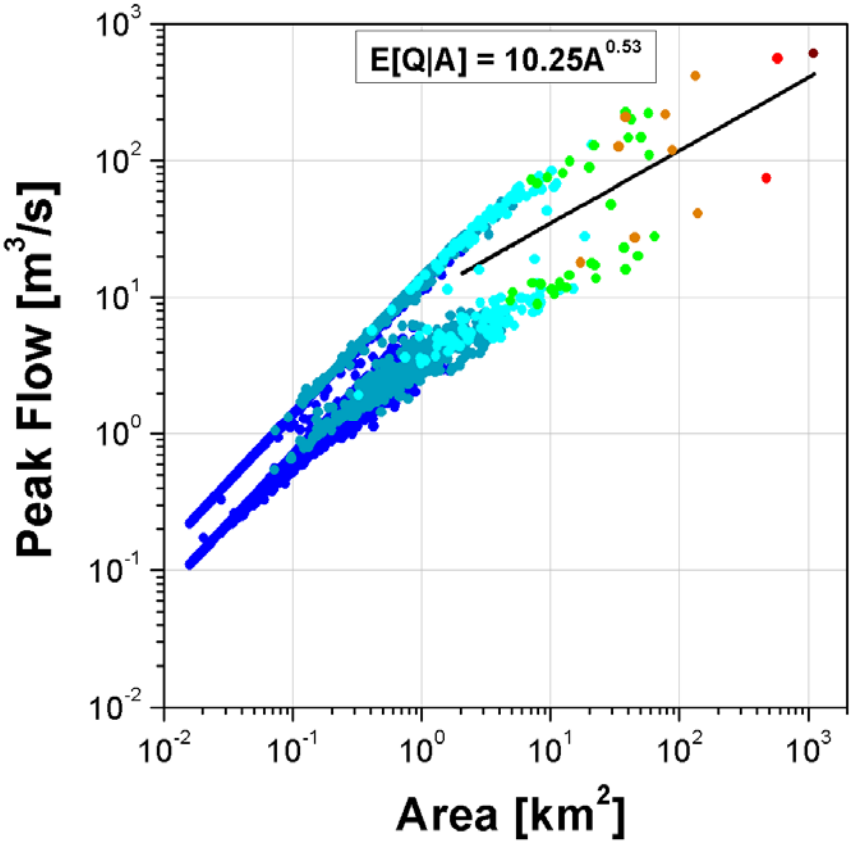

Figure 2.12 Effect of space-time variability of rainfall on the peak flow scaling structure. The rainfall is taken to be spatially uniform with the intensity of 25 mm/h for 30 minutes over western half of the basin. For the eastern half of the basin, the rainfall is 50 mm/h for 120 minutes. The color scheme is same as Figure 2.7

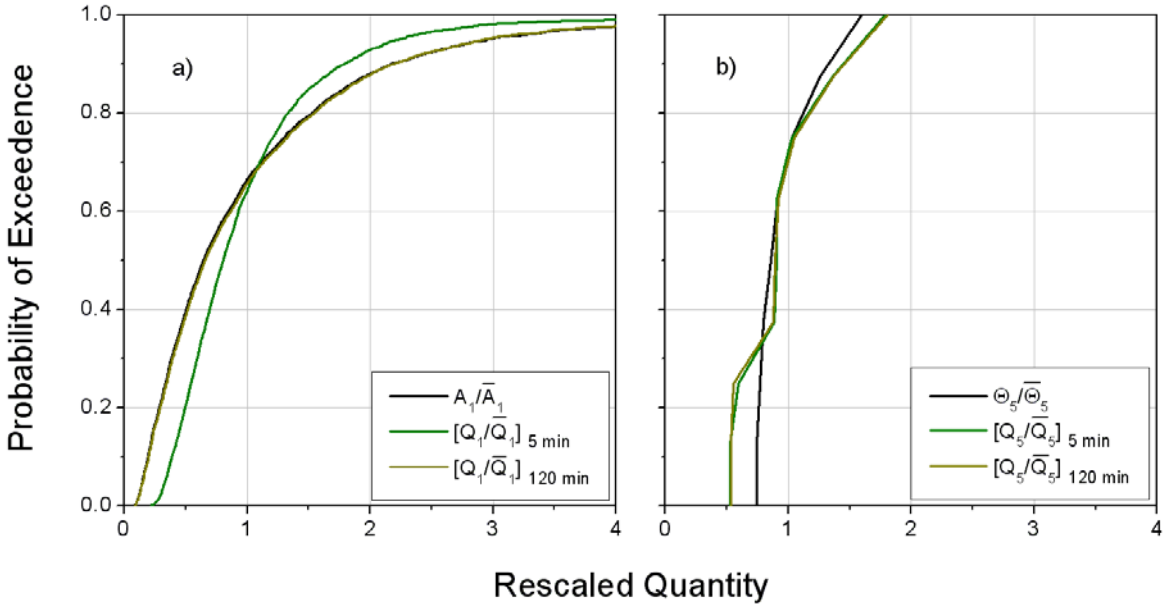

Figure 2.13 Probability distributions of rescaled areas, width function maxima and peak flows for order 1 and order 5 basins.

CHAPTER 3

EFFECT OF RAINFALL VARIABILITY: SIMULATIONS WITH SPACE-TIME VARIABLE RAINFALL*

3.1 Introduction

In the previous chapter, we investigated the sensitivity of the peak flow scaling structure to the basic characteristics of rainfall such as intensity, duration, and advection velocity employing idealistic spatially uniform rainfall scenarios. However, it is well known that rainfall is extremely variable both in space and time. Several studies analyzed the effect of spatial variability of rainfall on the output hydrographs from either lumped or distributed hydrological models [e.g., *Beven and Hornberger*, 1982; *Milly and Eagleson*, 1988; *Naden*, 1992; *Pessoa et al.*, 1993; *Obled et al.*, 1994; *Shah et al.*, 1996; *Finnerty et al.*, 1997; *Singh*, 1997; *Koren et al.*, 1999; *Smith et al.*, 2004b; *Morin et al.*, 2006; *Saulnier and Le Lay*, 2009]. Nevertheless,* the results did not give a complete picture of the effect of rainfall variability sensitivity as the analyses were based on the hydrographs at the outlet or at specific locations. Instead of outlet hydrographs, we focus on the spatial scaling structure of the peak flows, and therefore shed more light on the sensitivity of the hydrologic response to the variability in the rainfall.

To investigate the statistical structure of peak flows for space-time variable rainfall, it is necessary to have information on spatial-temporal distribution of rainfall events. Such information can be conveniently provided by the ground-based weather radar network. We obtained radar estimates of three rainfall events that occurred in 2007 over Whitewater River basin, Kansas. We forced the hydrologic model CUENCAS with radar-rainfall estimates and obtained the hydrographs for all the interior sub-basins and

*Adapted from Mandapaka, P.V., W.F. Krajewski, R. Mantilla, and V. K. Gupta (2009), Dissecting the effect of rainfall variability on the statistical structure of peak flows, *Advances in Water Resources*, 32, 1508-1525.

outlet of the Whitewater River basin. The simulation framework in terms of the study area, hydrologic model parameters and the assumptions are essentially the same as in the previous chapter. We also assumed that flow in the channels is constant throughout the river network. Figure 3.1 shows the peak flow structure for the 6th of May 2007 event that lasted for approximately 25 hours. Figure 3.1 reveals that peak flows display scaling with a scaling exponent of 0.70. The scaling exponent of 0.70 is larger than the scaling exponent of the width function maxima. For the other two 2007 events that we analyzed, the scaling exponents were 0.68 and 0.77.

From the studies discussed in Section 2.6 of previous chapter, we know that when a spatially uniform rainfall is applied instantaneously, the peak flow scaling exponent is very close to that of the width function maxima. A real rainfall event is far from being spatially uniform and lasts for a certain duration. Therefore, the scaling exponent in Figure 3.1 is very different from that of the width function maxima. Figure 3.1 also shows that the scatter at small scales is different from that of simple scenarios considered in the previous chapter. Another conspicuous feature in Figure 3.1 is that the scale break is not well defined, possibly because of the inherent space-time variability of the rainfall event such as zero-rain intermittency and its spatio-temporal correlation structure. It is well known that radar-rainfall estimates are subject to considerable uncertainties [e.g. *Bras and Rodriguez-Iturbe*, 1993; *Krajewski and Smith*, 2002; *Ciach et al.*, 2007]. The errors propagate through the hydrologic models and contribute to the variability of the predicted peak flows across scales. However, we do not attempt to separate the effects of rainfall estimation uncertainties from the effects of natural variability of rainfall on the peak flow scaling structure. The natural variability of rainfall itself has a great impact on the statistical structure of peak flows, and understanding its role is the main goal of this study.

The simulation experiments discussed in previous chapter can only partially explain the effect of various characteristics of rainfall that resulted in Figure 3.1. To

investigate the effect of spatial variability of rainfall on the peak flow scaling structure, we extend the simulation framework of the previous chapter with spatially variable rainfall fields. Though there are many ways in which spatial variability can be characterized, we explore it in terms of variance, correlation structure and zero-rainfall intermittency of Gaussian random fields and intermittent lognormal random fields. Our experiments are designed so that we depart from the uniform rainfall scenarios in a gradual, simple manner to keep from losing the benefits of the previously gained understanding of the effects of the uniform intensity and duration. We start with Gaussian uncorrelated fields, and then proceed to Gaussian and lognormal correlated fields with varying degrees of intermittency. We employ a well-known space-time rainfall model developed by *Bell* [1987] to obtain lognormal rainfall fields with desired spatial correlation and intermittency.

This chapter is organized as follows. Following the introduction, Section 3.2 describes the peak flow scaling structure for Gaussian uncorrelated fields with changing variance. The effect of rainfall correlation structure on the peak flows is discussed in Section 3.3. The peak flow scaling structure obtained using the rainfall from the space-time model is discussed in Section 3.4.

3.2 Effect of Spatial Variability

3.2.1 Gaussian Uncorrelated Fields

In this scenario, we simulate Gaussian random fields with a mean of 25.0 mm/h and the standard deviation ranging from 0.1 mm/h to 6 mm/h. By gradually varying the variance, we gently depart from the well-understood case of uniform intensity. The duration of the rainfall is fixed at 120 minutes. The peak flow scaling structure for four different cases of standard deviation and linear routing mechanism is shown in Figure 3.2. In the Hortonian regression, we used the orders 3 to 7. Table 3.1 lists the average rainfall, intercept, scaling exponent and the peak flow at the outlet of the basin for all the

cases and for both routing mechanisms. Figure 3.2 and Table 1 reveal that the increasing variance has no significant effect on the fitted regression equations. The main effect of the variance is to increase the scatter in the peak flow scaling structure. The scatter is averaged out by the basin at the larger scales, as seen from the peak flow values at the outlet (Table 3.1). The slight variation in the intercepts and outlet peak flows is expected given that we are using realizations of a random process. This is further evident from the estimated values of the mean, which are different from the theoretical value of 25 mm/h.

3.2.2 Gaussian Correlated Fields

To investigate the effect of spatial correlation of the rainfall field on the scaling exponents, we obtain Gaussian fields with a mean of 25 mm/h and a standard deviation of 2 mm/h and that is characterized by an exponential correlation structure

$$\rho(d) = \exp\left[-\left(\frac{d}{\theta_1}\right)^{\theta_2}\right] \qquad (\theta_1 > 0, 0 \le \theta_2 \le 2), \tag{3.1}$$

where d is the distance between any two points in the field, θ_1 is the correlation distance defined as the distance at which the correlation drops to 1/e and θ_2 is the shape parameter that controls the shape of the correlation function at the origin. We fixed the shape factor at one and generated the random fields with the correlation distances varying from 5 km to 50 km. Each field is then applied for 120 minutes over the basin, and the peak flows are estimated for linear and nonlinear routing mechanisms. Figure 3.3 shows the scaling structure of peak flows for two extreme cases of correlation distances and a linear routing mechanism. The effect of increasing correlation is to decrease the scatter in the scaling structure (Figure 3.3). Table 3.2 shows that the larger scale basin response is almost independent of the correlation structure, although there is some variability in the intercepts and outlet peak flows, which is mainly due to the fact that we are using realizations of a random process.

3.2.3 Zero-Rain Intermittency: Uncorrelated Fields

To investigate the effect of zero-rainfall intermittency, we simulated random rainfall with varying degrees of zero-rainfall intermittency and a duration of 120 minutes. The value of rainfall over each pixel was drawn from uniform distribution U[10,30], and intermittency is introduced randomly but maintains an overall mean fixed at 20 mm/h. We considered four values of intermittencies: 0.0, 0.05, 0.25 and 0.50 (corresponding rainy area fractions are 1.0, 0.95, 0.75 and 0.50). These rainfall scenarios were supplied as input to the CUENCAS model, and the peak flow scaling structure was obtained for linear routing mechanisms. The sensitivity of the peak flow scaling to the intermittent random fields is shown in Figure 3.4. With the increase in intermittency (or decrease in rainy area), the scatter for the smaller scale basin peak flows increased. However, the effect of intermittency is reduced for the larger scale basins, as evidenced by the linear regression equations shown in each panel of Figure 3.4 and also from the outlet peak flow values shown in Table 3.3. The simulations are repeated for the nonlinear routing mechanism, and we found a similar pattern to the pattern found in linear routing, although the intercepts and scaling exponents differed (Table 3.3).

3.2.4 Zero-Rain Intermittency: Correlated fields

To study the effect of intermittency in a more realistic manner, we selected the spatial component of the rainfall model developed by *Bell* [1987]. The model belongs to the class of meta-Gaussian models [e.g., *Mejia and Rodriguez-Iturbe*, 1974; *Bell*, 1987; *Guillot and Lebel*, 1999] and generates a two-dimensional isotropic, correlated random field using spectral analysis. A non-linear transformation and an external threshold are then applied to obtain a rainfall field with desired intermittency, average intensity and correlation structure. In *Bell* [1987], the use of exponential transformation resulted in lognormally distributed rainfall. The parameters for the model are the log-transformed (Gaussian) mean and variance of the rainy area, the zero-rainfall intermittency factor and

the spatial correlation structure. The parameters we selected are 0 and 0.5 for the log-transformed mean and variance of the rainy area and exponential correlation structure with a correlation distance of 20 km. We observed that the realizations from the model, besides having the desired spatial correlation structure, also displayed spatial scaling behavior. The duration of the event is fixed at 120 minutes. To keep the volume constant with changing intermittency, we simulated a single realization with a given correlation structure on a large (256×256) domain and selected the portion that yielded the desired intermittency and volume. The spatial structure of the field thus obtained will remain the same as the larger one. Figure 3.5 shows the scaling structure of peak flows for four different intermittency factors starting from 0 to 0.50 for the linear channel routing mechanism. Unlike in Figure 3.4, significant scatter was observed even for higher order basins (particularly for the bottom panels of Figure 3.5) when the pixels are correlated. The large scatter is due to the high probability of concentrated intermittent pixels present in correlated intermittent fields. Whereas, for uncorrelated intermittent fields, the river network efficiently aggregates the randomness in the fields. However, the overall behavior - of increasing scatter with increasing intermittency - is similar for both scenarios.

3.3 Simulations from Space-Time Rainfall Model

The scenarios investigated so far have offered insight into the effects of different characteristics of rainfall on the spatial scaling structure of peak flows. We will now investigate the basin's response to more realistic space-time rainfall events. Therefore, we have simulated a space-time rainfall event from a model developed by *Bell* [1987] and applied it over the basin. The spatial component of the model is described in the previous section. The temporal evolution of the rainfall is modeled as an autoregressive process with parameters based on the correlation time of area-averaged rainfall. The parameters for the model are obtained by analyzing several storms over the Midwest. In this study,

we simulate two different storm events with characteristics listed in Table 3.4. Storm 2 is more variable than storm 1, as seen from the values of the coefficient of variation and correlation distance. Another important difference is that storm 1 lasts longer and has larger values of mean and rainy fraction compared to storm 2. The values of the shape factor suggest that storm 2 is more correlated at very small scales than storm 1.

Figure 3.6 shows the basin response to the two storms for the linear routing mechanism. The peak flow scaling structure for this complex scenario can now be explained using the results from idealized scenarios. Although the scale break is sharp and evident for the simulation scenarios of Section 2.6 under idealized conditions, it is not clearly seen in the peak flow scaling structure resulting from the simulated realistic space-time rainfall event. Since the scaling exponent is not close to 1.0, basin saturation (for instance, top panels of Figure 2.7) is not the reason for the absence of scale break. The lack of sharp scale break for realistic rainfall scenarios can be best explained by revisiting the idealized scenarios in Section 2.6. For instance, a combination of just two different intensities and durations has diffused the scale break in Figure 2.11. The space-time rainfall fields are characterized by different intensities, durations, correlations and intermittencies and move with a certain advection velocity. In a way, these fields are a combination of all of the scenarios considered in Section 2.6. This explains the absence of scale break for storm 1.

The scatter for storm 1 is smaller than that of storm 2. Two main factors responsible for the reduced scatter are the duration of the storm and zero-rain intermittency. Results from Section 2.6.1 indicate that one consequence of longer duration events is the decreased scatter that is most pronounced at smaller scales. Also, from Sections 3.2.3 and 3.2.4 that pertain to the effect of zero-rain intermittency, we note that the scatter in the peak flow structure increases rapidly as the area of rainfall decreases. Though the rainy fraction of storm 1 is 46 %, the advection of the storm eventually increases the effective wetted area of the basin, thereby decreasing the scatter

seen at smaller scales. For the second storm, the large scatter is due to the increased intermittency combined with the shorter duration and large coefficient of variation.

The regression equation seen for storm 1 is obtained in a Hortonian framework using the orders 2 to 7. For storm 2, peak flows corresponding to the Horton orders of 4 to 7 are used in this storm's regression. The scaling exponent for both storms is larger than the scaling exponent of the width function maxima. For the nonlinear routing scenario (not shown), we noticed a similar pattern with larger scatter and higher values for scaling exponents than in the linear routing case.

3.5 Conclusions

In Chapter 2, we investigated the sensitivity role of spatial structure of peak flows to the basic characteristics of rainfall such as intensity, duration and storm advection velocity using simple scenarios of uniform rainfall. This chapter extended the framework by systematically introducing variability in the rainfall fields. The variability was characterized in terms of variance, space-time correlation functions, and zero-rainfall intermittency. We started with Gaussian uncorrelated rainfall fields and proceeded to complex rainfall fields obtained from a space-time rainfall model. Our rainfall variability cases and range of values, though simple, captured the key aspects of natural rain systems. The main effect of variability, as characterized by variance, spatial correlation and the spatial intermittency, is to increase the scatter in the scaling structure of the peak flows. At larger scales, the effect of variability decreases, as seen from peak discharges at the outlet of the basin. For homogeneous rainfall fields and under idealized conditions of flow routing on hillslopes and in channels, we observed that the smaller scale basin response was dominated by the rainfall intensity (and spatial distribution), while the hydrologic response of larger scale basins was driven by rainfall volume, river network topology and flow dynamics. Based on the simulations on a deterministic Mandelbrot-Viscek network, *Menabde and Sivapalan* [2001] reported that the variability in the

rainfall decreases the scaling exponent of peak flows on both sides of the scale break. We did not observe such behavior in our simulations.

The results obtained from the simple scenarios in Chapter 2 and Section 3.2 enhanced our understanding of the peak flow scaling structure obtained from simulated space-time variable rainfall. Storm duration and advection are the key factors that control the effective zero-rain intermittency, which in turn affects the scatter in the peak flows. The peak flow scaling structure for the realistic space-time rainfall scenarios did not present a clear and sharp scale break. The scale break was masked due to the inherent space-time variability in the realistic rainfall fields.

Table 3.1 Sensitivity of intercepts, scaling exponents and outlet peak flows to the variance of the Gaussian rainfall field with a mean intensity of 25 mm/h.

$N(\mu,\sigma^2)$	Mean [mm/h]	Linear Routing			Nonlinear Routing		
		Intercept	Slope	Outlet Peak Flow [m^3/s]	Intercept	Slope	Outlet Peak Flow [m^3/s]
N(25,0.1)	24.99	12.02	0.54	587.54	12.79	0.63	1246.87
N(25,1.0)	24.98	12.00	0.54	587.13	12.73	0.63	1245.98
N(25,4.0)	25.02	12.00	0.54	587.01	12.74	0.63	1246.65
N(25,9.0)	24.93	12.08	0.54	588.67	12.85	0.63	1228.01
N(25,16.0)	25.09	11.95	0.54	580.44	12.70	0.64	1249.81
N(25,25.0)	24.88	11.93	0.54	587.15	12.70	0.63	1243.16
N(25,36.0)	24.88	11.97	0.54	589.58	12.64	0.64	1252.31

Table 3.2* Sensitivity of intercepts and scaling exponents to the spatial correlation structure of the rainfall field.

Correlation Distance [km]	Mean [mm/h]	Linear Routing			Nonlinear Routing		
		Intercept	Slope	Outlet Peak Flow [m^3/s]	Intercept	Slope	Outlet Peak Flow [m^3/s]
5.0	25.06	12.19	0.54	581.92	12.99	0.63	1232.85
10.0	24.41	11.92	0.54	566.09	12.67	0.63	1192.25
20.0	24.83	11.86	0.54	582.06	12.62	0.64	1222.65
30.0	26.49	12.54	0.54	619.69	13.44	0.64	1340.37
40.0	25.48	12.33	0.54	597.79	13.17	0.63	1279.25
50.0	23.46	11.29	0.54	545.59	11.88	0.63	1120.85

*The rainfall field is assumed to be Gaussian with a mean intensity of 25 mm/h and a variance of 2.0 mm/h and is characterized by an exponential correlation function with the correlation distances indicated in the Table.

Table 3.3* Sensitivity of intercepts and scaling exponents to the intermittency in the uncorrelated random fields.

Intermittency [%]	Mean [mm/h]	Linear Routing			Nonlinear Routing		
		Intercept	Slope	Outlet Peak Flow [m^3/s]	Intercept	Slope	Outlet Peak Flow [m^3/s]
0	19.88	11.45	0.50	464.66	12.26	0.58	898.48
5	20.04	11.73	0.50	473.95	12.23	0.58	907.88
25	20.05	11.23	0.50	469.37	11.58	0.58	907.24
50	19.84	11.91	0.50	485.78	12.41	0.58	942.22

*The value of rainfall over each pixel was drawn from a Uniform distribution U[10,30], and the duration of the event is 120 minutes. The mean of the field is kept constant for different intermittencies.

Table 3.4* Characteristics of rainfall events simulated from the space-time rainfall model.

Storm	Mean [mm/h]	Standard Deviation [mm/h]	Coefficient of Variation [mm/h]	Correlation Distance [km]	Shape Factor	Rainy Area [%]	Duration [hr]
1	3.97	10.37	2.61	15.40	0.73	45.84	20
2	1.41	5.97	4.23	5.52	0.92	17.51	4

*A two parameter exponential correlation function is used to characterize the spatial dependence.

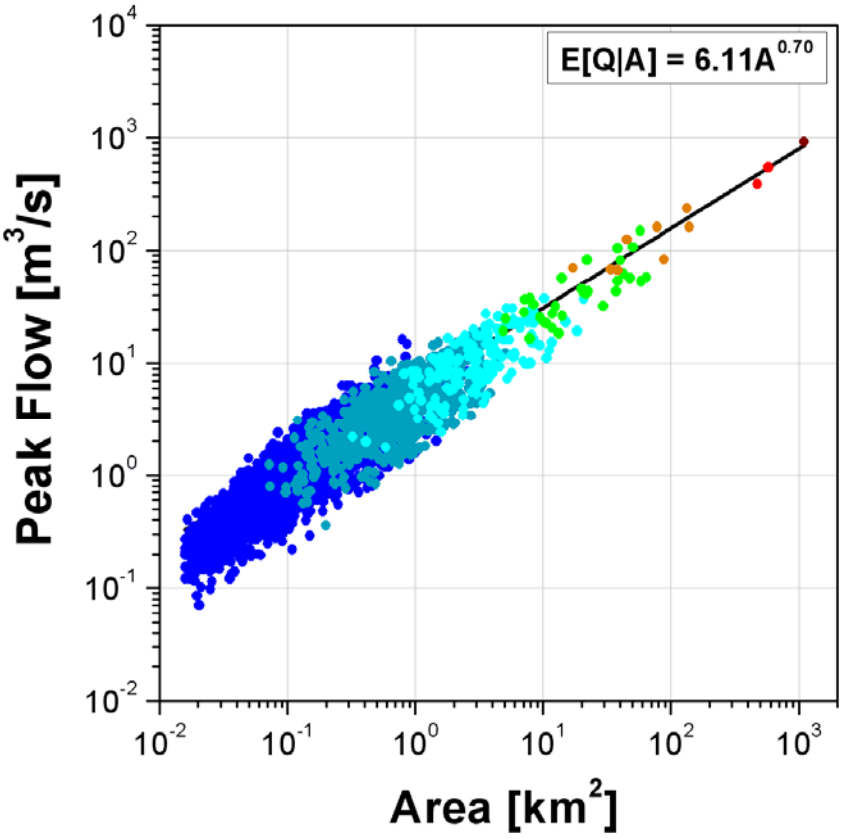

Figure 3.1 Scaling of peak flows with respect to the upstream areas of all the sub-basins in the Whitewater River basin, Kansas. The rainfall data is obtained from the KICT NEXRAD weather radar at Wichita, Kansas. The color scheme indicates the Horton orders as in Figure 2.7.

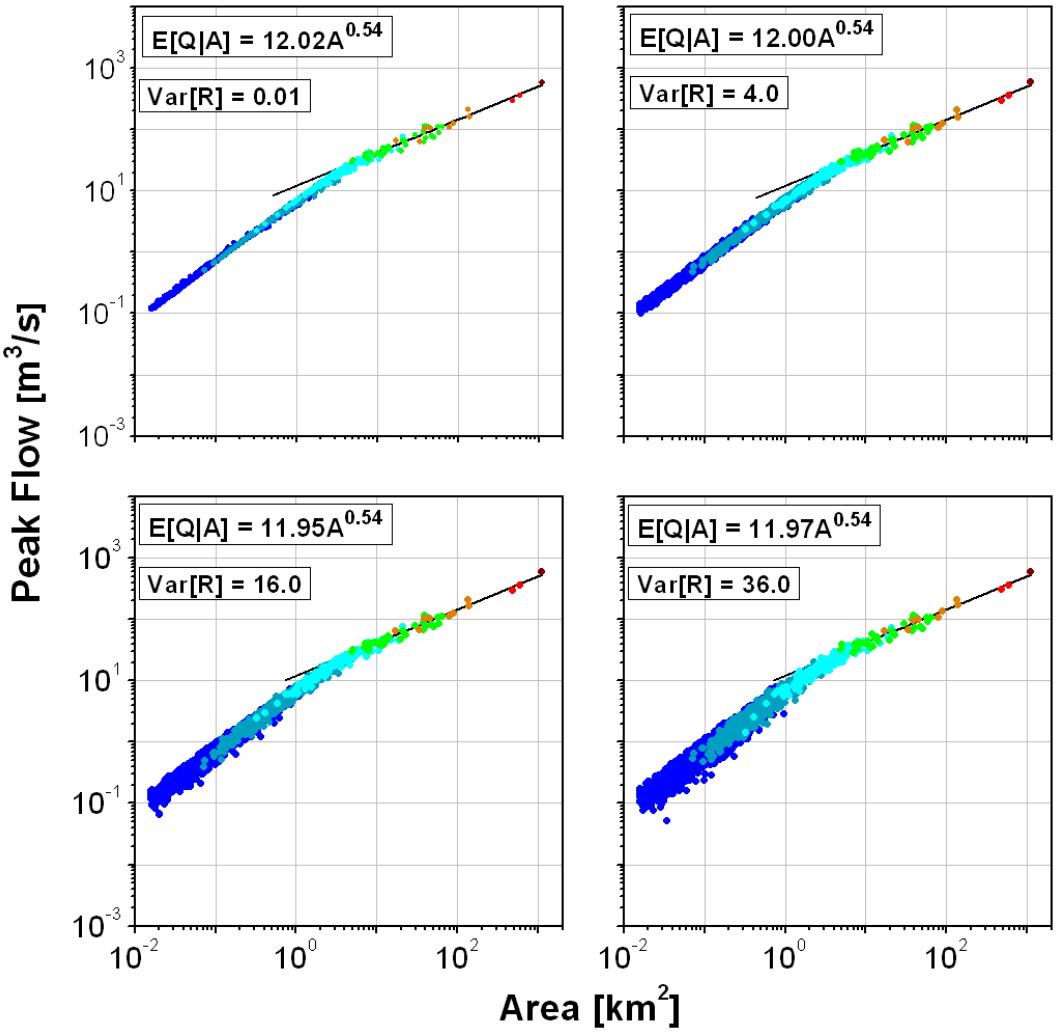

Figure 3.2 Sensitivity of peak flow scaling structure to the variance of the rainfall field. The rainfall field is assumed to be Gaussian with a mean of 25 mm/h and variance indicated on each panel. The color scheme indicates the Horton orders as in Figure 2.7.

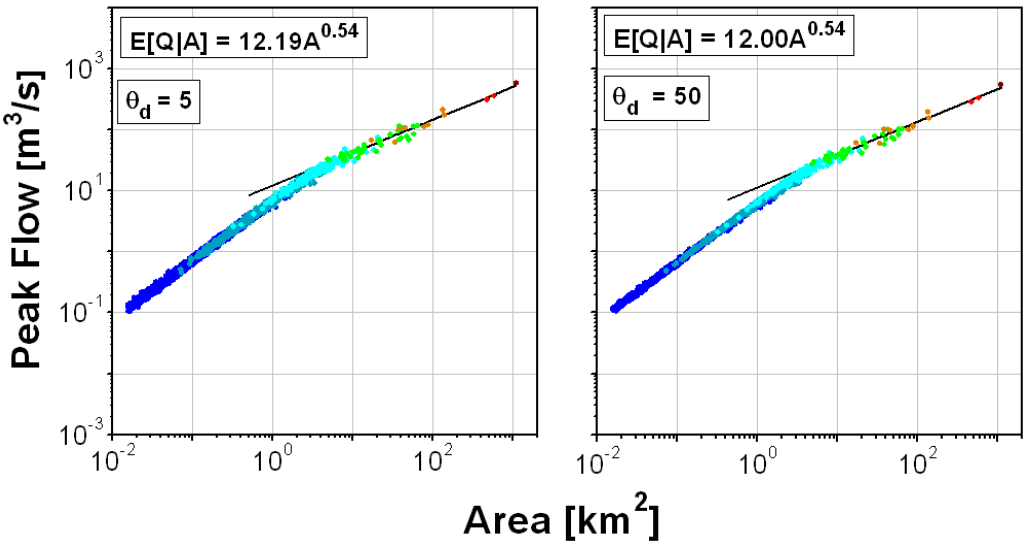

Figure 3.3 Sensitivity of peak flow scaling structure to the spatial correlation of the rainfall field. The rainfall field is assumed to be Gaussian with a mean of 25 mm/h, standard deviation of 2.0 mm/h. The spatial structure of the rainfall field is characterized by an exponential correlation structure with the correlation distance indicated on each panel. The color scheme indicates the Horton orders as in Figure 2.7.

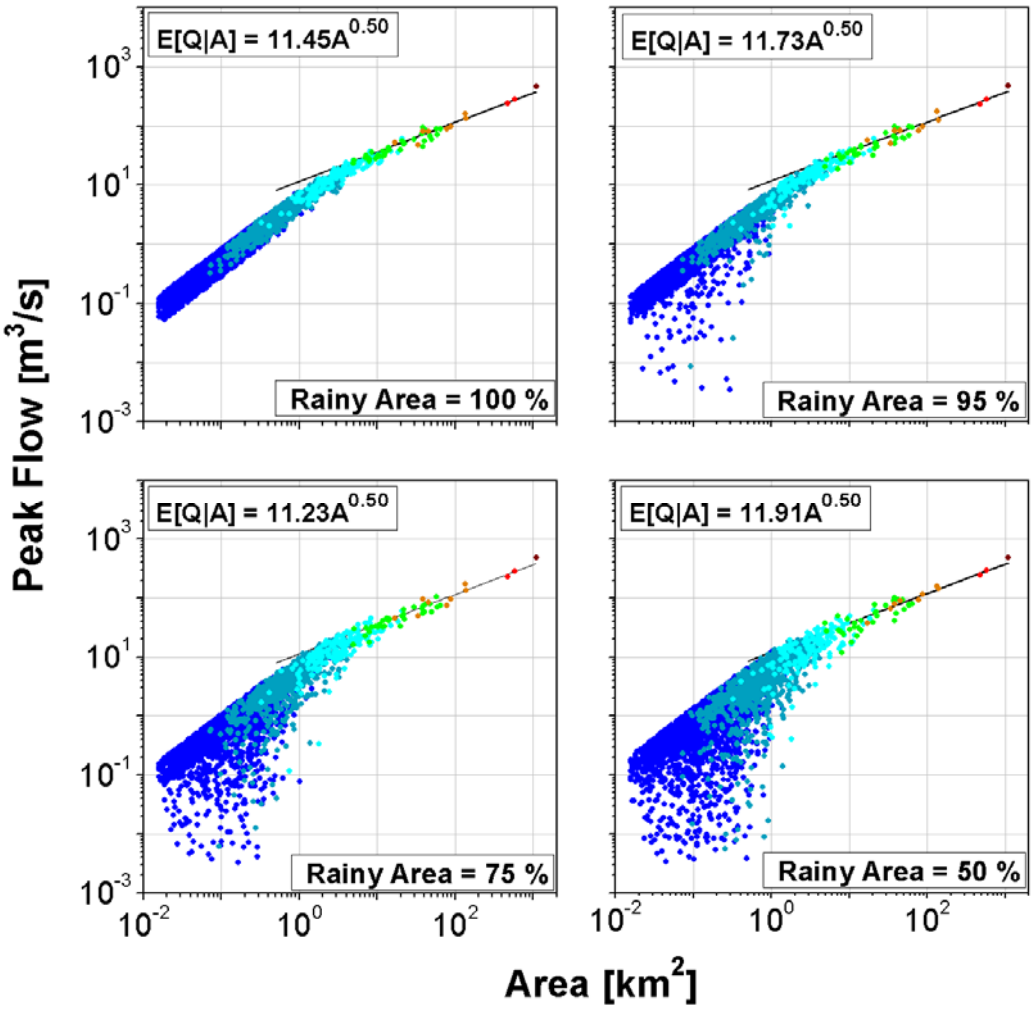

Figure 3.4 Sensitivity of the peak flow scaling structure to the spatial zero-rainfall intermittency in rainfall fields. The rainfall fields are distributed randomly in space with the value at each pixel drawn from uniform distribution U[10,30] with a mean of 20 mm/h and a duration of 120 minutes.

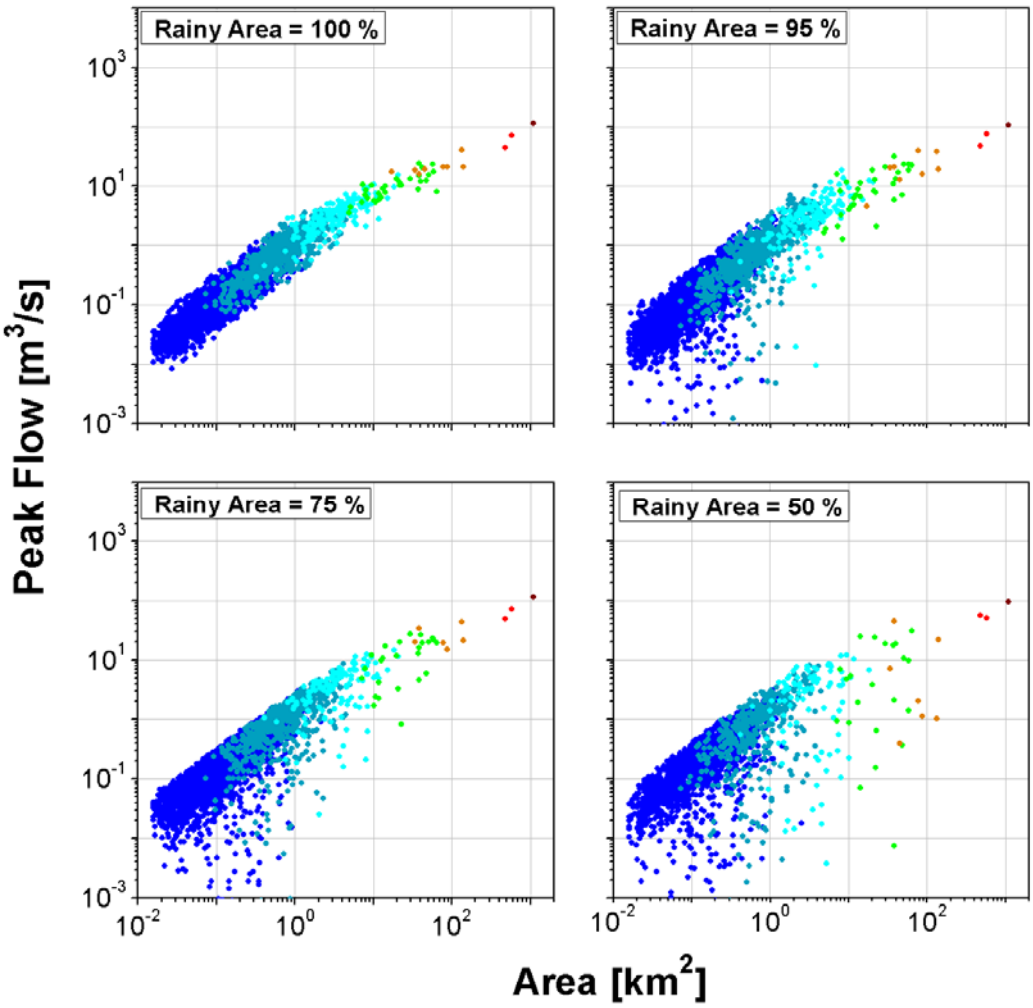

Figure 3.5 Sensitivity of the peak flow scaling structure to the spatial zero-rainfall intermittency of rainfall field. The rainy portion of the field is assumed to follow lognormal distribution. The color scheme indicates the Horton orders as in Figure 2.7.

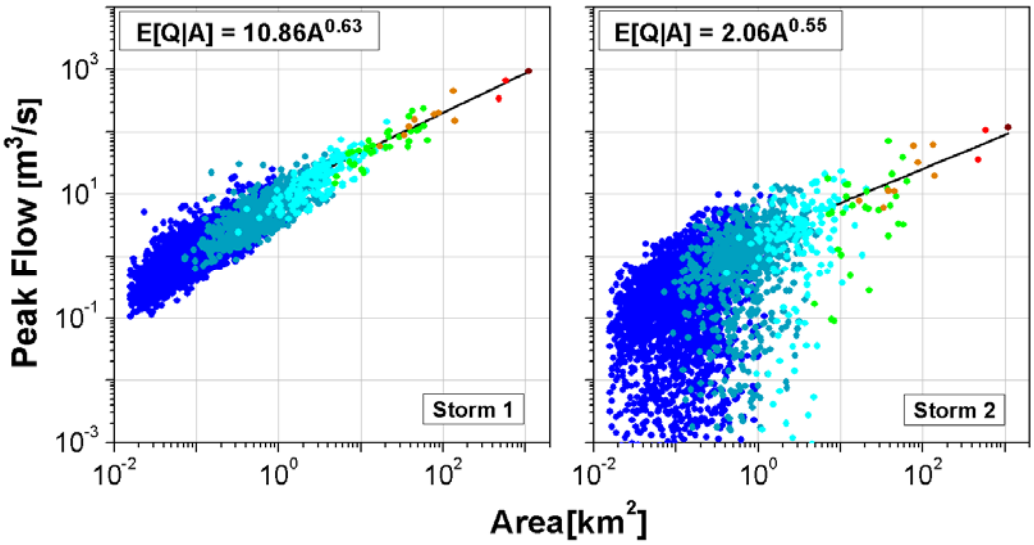

Figure 3.6 Response of the watershed to the simulated space-time rainfall events. The characteristics of the storms are listed in Table 3.3. The color scheme indicates the Horton orders as in Figure 2.7.

CHAPTER 4

EFFECT OF RAINFALL AND HILLSLOPE VARIABILITY: INTRODUCING HILLSLOPE DELAY

4.1 Introduction

The effect of space-time variability of rainfall on the peak flow scaling structure has been investigated in Chapters 2 and 3 using ~1200 km^2 Whitewater River basin as an illustration. The results have improved our understanding on the role of rainfall variability and the river network structure on the scale-invariance of peak flows. However, it is very important to check the validity of the results for different basins with different geomorphologic characteristics. One of the key results from the previous chapters was that the variability in rainfall seen at smaller scales was averaged out by the river network as one moved to larger scales. However, we did not have sufficient large scale basins to strongly support the above statement. In this chapter, we extend the analysis to ~ 7250 km^2 Iowa River basin up to Marengo, Iowa. By moving to a larger basin, we will have a better perspective of the attenuation and aggregation of flows by the river network.

In the previous chapters, we did not consider the role of hillslope travel time on the peak flow scaling structure. We assumed that the runoff generation mechanism was Hortonian and that the flow reached the channel instantaneously. The hillslope hydrology is characterized by variety of processes such as interactions between soil and atmosphere, vegetation and soil moisture, rainfall and soil moisture, surface and subsurface flows, leading to a complex system of runoff generation at very small scales [e.g., *Kirkby*, 1978; *Kirkby*, 1988; *Anderson and Burt*, 1990]. Several studies have investigated the role of hillslopes in shaping the catchment hydrologic response [e.g., *Robinson et al.,* 1995; *Ritsema et al.*, 1996; *Ogden and Watts*, 2000; *D'Odorico and Rigon,* 2003; *Ridolfi et al.*, 2003; *Weiler and McDonnell*, 2004]. Particularly in small

catchments, where channels are relatively short, the travel time in the hillslope is comparable to the travel time in the channels [e.g., *D'Odorico and Rigon*, 2003]. Although such studies have greatly improved our understanding of the hillslope hydrology, we still lack a coherent framework to systematically investigate the effect of hillslope processes across a wide range of scales. Therefore in this study, instead of focusing on each hillslope process, we account for the net effect of hillslopes in terms of hillslope travel time by incorporating hillslope velocity. It should be noted that the assumptions of zero-infiltration and zero-evaporation are still in place.

So far in this thesis, the river network has been extracted from 30 m resolution DEM. Current computational resources limit the use of high resolution DEMS and hillslope-link based hydrologic modeling for large basins such as Mississippi River basin. Another goal of this chapter is to explore the role of the river network extraction on the statistical structure of peak flows.

4.2 Description of the Study Area

The Iowa River basin (Figure 4.1) stretches between latitudes 37° 46'E and 38° 09'E and longitudes 96° 51'W and 97° 18'W. In this chapter, we considered Iowa River basin up to Marengo (Figure 4.1) with an area of approximately 7250 km^2 and highest Horton order of eight, when a 30 m DEM is used to extract the river network. This results in about 100,000 hillslopes and, thus, channel links for this basin. In Figure 4.1, we also show the extracted channel network with links of Horton orders 5 to 9. Figure 4.2 shows the topologic width function for the Iowa River basin at Marengo. We estimated the width functions for each link in the river network and obtained the corresponding maximum values.

Figures 4.3(a) and 4.3(c) show the Horton plots for drainage areas and width function maxima of links of various orders for the Whitewater River basin, Kansas. The log-linearity in Figures 4.3(a) and 4.3(c) confirm the statistical self-similarity of the

upstream areas and the width function maxima [e.g., *Strahler*, 1957; *Peckham and Gupta*, 1999; *Furey and Troutman*, 2008]. In the regression analysis, we used the areas and width function maxima corresponding to the Horton orders 2 to 7. The order 8 stream is not used in the Horton regression due to sampling reasons: we have only one point corresponding to the order 8. Although, averages corresponding to order 1 streams do not suffer from sampling issues, they are usually not considered in the regression [e.g., *Peckham and Gupta*, 1999; *Mantilla and Gupta*, 2005] as they represent the finest detail in a stream network, and therefore the corresponding basins do not contain a "network". The Horton ratios for the areas and width function maxima are then obtained by exponentiation of the slopes from the regression analysis. The scaling exponent of width function maxima obtained through Horton ratios in (1) is 0.46.

If the upstream areas and width function maxima display log-linearity, as shown in Figures 4.3(a) and 4.3(c), then $E[X_\omega] = E[X_1]\bullet(R_X)^{\omega-1}$, where X is either the upstream area or the width function maxima and R_X is the corresponding Horton ratio. The rescaled upstream areas and width function maxima are obtained by dividing each value of X_ω by $E[X_1]\bullet(R_X)^{\omega-1}$. The probability distribution of the quantity $X_\omega/[E[X_1]\bullet(R_X)^{\omega-1}]$ is called the rescaled probability distribution. In Figures 4.3(b) and 4.3(d), we show the statistical self-similarity of areas and width function maxima in terms of their rescaled probability distributions for orders 1 to 7. Although order 1 basins were not considered in the regression analysis, it can be seen that their rescaled probability distribution collapses onto those of orders 2 to 7.

4.3 Sensitivity to the Intensity and Duration of Spatially Uniform Rainfall

4.3.1 Without Hillslope Delay

Spatially uniform rainfall of different intensities was imposed on the study area for different durations. The assumptions in the hydrologic model were same as in

chapters 2 and 3; i.e. 1) negligible evapotraspiration, 2) purely surface runoff, and 3) instantaneous flow to the channels (no hillslope effect). The channel routing was assumed to be nonlinear with velocity that depends on the discharge in each link and the corresponding upstream area (equation 2.1). For all the rainfall scenarios in this chapter, we obtained peak flow regression equations (Figure 4.4) in Hortonian framework described in Sections 2.3 and 2.6. That is, instead of obtaining the scaling exponent by regression of peak flows with upstream areas, we use equation 1.2 to obtain the scaling exponent. And the coefficient of the power law is obtained so that the regression line passes through the average of the peak flows corresponding to the top three orders (Figure 4.4). We noticed that the scaling exponent was always greater than the scaling exponent of the width function maxima (Figure 4.4), which is in confirmation with Whitewater River basin results discussed in Chapter 2 and with *Mantilla et al* [2006] for the Walnut Gulch basin.

For the durations of 5 min and 60 min, the scaling exponent was equal to 0.58 for all three rainfall intensities, and when the duration was changed to 240 min, the scaling exponent was found to decrease with the decrease in rainfall intensity (Figure 4.4). Similar behavior was reported in Chapter 2 for the Whitewater River basin. The scale break (solid red line in Figure 4.4) was estimated using the methodology presented in Chapter 2 (equation 2.2). We did not estimate the scale break for the shorter duration of 5 min due to large scatter at smaller scales. With the increase in rainfall duration, more number of small scale links reached steady state, and the scale break moved towards larger scale basins (Figure 4.4). We noticed similar behavior with an increase in the rainfall intensity; the scale break moved towards larger scale basins (Figure 4.4). Overall, the response of the Iowa River basin to the spatially uniform rainfall was similar to that of Whitewater River basin. In the next subsection, we relax the assumption of instantaneous flow to the channel to study the role of hillslope characteristics.

4.3.2 With Hillslope Delay

The scenario with the instantaneous flow to the channel can be considered as equivalent to the very high value for the hillslope velocity. We gradually decreased the velocity to 1.0 m/s and then to 0.1 m/s to study the net effect of hillslopes on the hydrologic response of a basin. It should be noted that the value of 1.0 m/s was still high for a hillslope flow velocity and was comparable to the instantaneous flow situation. However, it already started to play a role in determining the peak flows at smaller scales, particularly for the shorter duration of 5 min (Figure 4.5). The scatter in the peak flows was reduced as the hillslope velocity was decreased from very high value (instantaneous flow) to 1.0 m/s (Figure 4.5). With further reduction in the hillslope velocity to 0.1 m/s, the scatter in the peak flows decreased, but with formation of a curve shaped envelope towards the smaller scale basins. The behavior can be observed particularly for the shorter duration of 5 min (Figure 4.6). However, the overall sensitivity of the peak flow scaling structure to the rainfall intensity and duration remained same as the instantaneous flow scenario. That is, more number of links reached steady state with the increase in rainfall duration and intensity, resulting in the reduction in the scatter and movement of the scale break towards higher order basins (Figures 4.5 and 4.6). Table 4.1 lists the scale breaks for each combination of rainfall intensity, duration and hillslope velocity. In general, the scale break moved towards smaller scales with the decrease in hillslope velocity; i.e. lesser number of links reached steady state with the introduction of the hillslope delay (Table 4.1).

The scaling of peak flows was much more sensitive for the hillslope velocities between 1.0 m/s and 0.1 m/s than those between instantaneous flow velocities and 1.0 m/s (Figures 4.4, 4.5, 4.6, and Table 4.1). The changes in the small-scale scatter and the location of the scale break did not significantly affect the fitted regression equations. While the scaling exponents remained exactly same for the velocity corresponding to instantaneous flow and velocity of 1.0 m/s, they changed only slightly (except for the

case of 50 mm/h and 240 min duration) as the hillslope velocity reduced to 0.1 m/s (Figures 4.4, 4.5, and 4.6).

4.4 Spatially Correlated Rainfall Fields

In the previous section, we studied the sensitivity of peak flows to the spatially uniform rainfall of varying intensities and duration under different scenarios of hillslope routing. Here, we consider the effect of rainfall spatial variability on the peak flows by systematically increasing the smoothness in the rainfall fields. The rainfall intensity distribution was assumed to be Gaussian with a mean of 25 mm/h and standard deviation of 6 mm/h. The rainfall fields were characterized by exponential correlation structure (equation 3.1), with correlation distances of 5, and 50 km. The rainfall fields were applied over the Iowa River basin for durations of 60 min and 240 min. Figure 4.7 shows the peak flow scaling structure for different correlation distances, rainfall durations, and with and without hillslope delay. For the correlation distance of 5 km, the river network effectively integrated the variability in the rainfall fields (Top panels of Figure 4.7). However, when the correlation distance increased to 50 km, larger local variability combined with increased smoothness hampered the efficiency of the river network in integrating the flows (Bottom panels of Figure 4.7).

It is well known that the rainfall intensity distribution is skewed. Therefore our next step was to change the distribution from Gaussian to lognormal. We subjected the Iowa River basin to lognormally distributed, spatially correlated fields of varying correlation distances. The mean of the rainfall field is 5 mm/h and the standard deviation is equal to 8 mm/h. In general, the sensitivity of the peak flow scaling to the lognormal rainfall scenario is similar to that of Gaussian fields. The reduced efficiency of the river network in handling the rainfall with high local variability and the increased correlation distance can be particularly observed for lognormally distributed rainfall fields (Bottom

panels of Figure 4.8). This behavior partially explains the lack of clear scale break for the peak flows corresponding to the NEXRAD radar-rainfall events.

4.5 Sensitivity to Channel Network Extraction

Throughout the thesis, we extracted the river network from the digital elevation models with 30 m resolution. *Mantilla and Gupta* [2005] showed that the network extracted from the 30 m resolution DEM reasonably matched the terrain's actual network. With 30 m resolution DEM, Iowa River basin had approximately 100,000 hillslopes. It is computationally expensive to perform simulations for larger basins such as Mississippi River basin with millions of hillslopes. In this section, we explored the tradeoffs between the channel network extraction and the peak flow scaling structure by repeating the experiments with spatially uniform rainfall on river network extracted from 30 m and 90 m DEMs. As expected, the decrease in DEM resolution from 30 m to 90 m reduced the number of hillslopes from ~ 100,000 to ~29,000. Since the low resolution DEM results in lower drainage density, the Horton orders for the corresponding streams was also affected, particularly for the lower order streams. For example, a third order stream in the network extracted from a 30 m DEM could be first order stream in the network extracted from a 90 m resolution DEM. The peak flow scaling behavior for a spatially uniform rainfall of 25 mm/h, and duration of 60 min and 240 min for two different DEM resolutions is shown in Figure 4.9.

It can be seen that the channel network extraction has a huge impact on the scale-invariant behavior of peak flows. It affects not only the basin response at smaller scales but also towards the larger scales, as noticed from the regression equations and the scaling exponents.

4.6 Summary and Conclusions

In this chapter, we extended the simulation framework described in Chapters 2 and 3 to the Iowa River basin up to Marengo, Iowa. The study area is different from

Whitewater River basin in many aspects. First, it is approximately six times larger than the Whitewater River basin thus allowing us to minimize the sampling effects at large scales and explore the scaling behavior. Secondly, the shape of Iowa River basin is very different from the square shaped White water River basin, and therefore the role of river network would be different. We also studied the effect of hillslope characteristics and channel network extraction on the scale-invariance of peak flows. The results suggested that the variability at the smaller scales, either from rainfall or from the hillslopes, will be averaged out by the river network. However, the efficiency of the river network in attenuating the variability depends on the intensity distribution and smoothness of the rainfall fields. The results also suggested that the DEM resolution and the corresponding channel network plays a major role in the statistical structure of peak flows.

Chapters 2, 3, and 4 foster development of a scaling based predictive framework for peak flows using remotely sensed rainfall products over basins ranging from very small to very large scales. A key question there is, "What is the scale at which remote sensing products provide meaningful predictions?" Our results suggest that the variability contributed by random errors of remote sensing sensors, such as weather radars and satellites, are filtered out by the drainage structure of river basins at some scales. To investigate further into this problem, we need a model for radar-rainfall uncertainties. The next chapter is a step forward in that direction as we present a framework to estimate spatial correlation of radar-rainfall errors.

Table 4.1 Sensitivity of scale break (in km^2) to the rainfall intensity and duration for three different scenarios of hillslope velocity: 1) No hillslope delay (NHD), 2) hillslope velocity of 1.0 m/s, and 3) hillslope velocity of 0.1 m/s.

Intensity	50 mm/h			25 mm/h			5 mm/h		
Duration	NHD	1.0	0.1	NHD	1.0	0.1	NHD	1.0	0.1
240 min	54.73	49.52	30.01	30.01	30.01	20.10	16.45	16.45	11.01
60 min	1.77	1.44	0.13	1.44	1.06	0.11	0.56	0.50	0.06

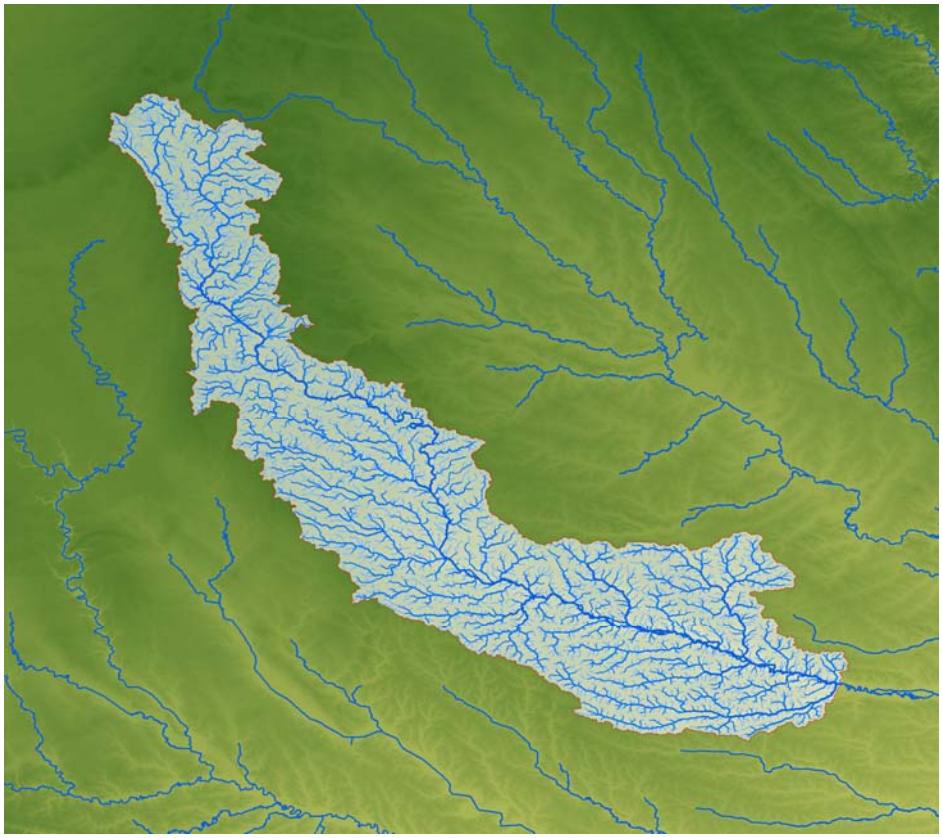

Figure 4.1 Map showing the topography and the river network for the Iowa River basin up to Marengo, Iowa.

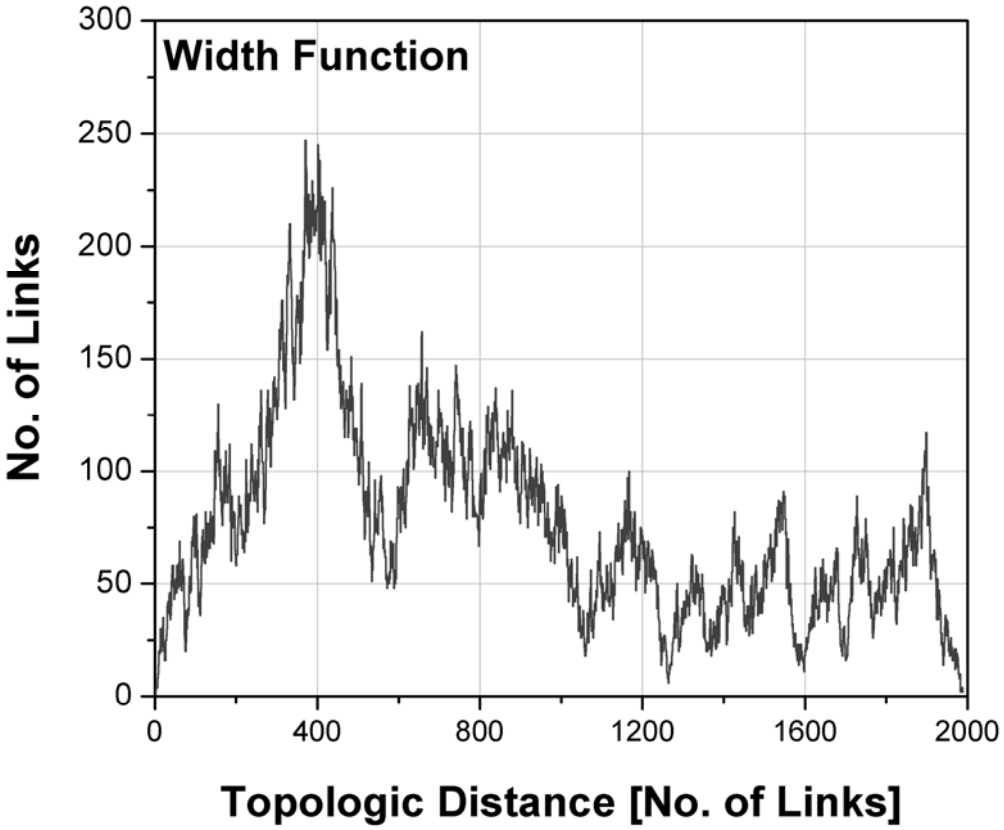

Figure 4.2 Topologic width function for the outlet of the Iowa River basin at Marengo, Iowa

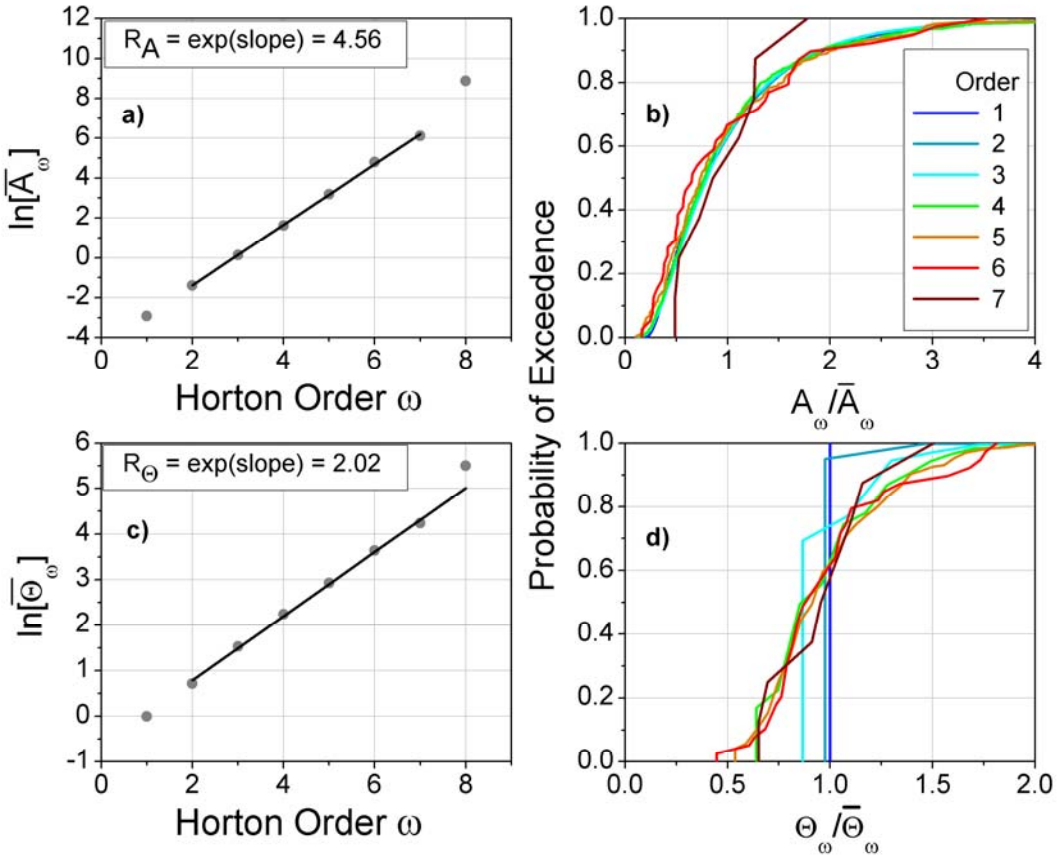

Figure 4.3 Statistical self-similarity of upstream areas and width function maxima in terms of Horton plots (left panels) and rescaled distributions (right panels). The ordinary least square regression is used to obtain the corresponding Horton ratios. The first order and seventh order links are not considered in fitting.

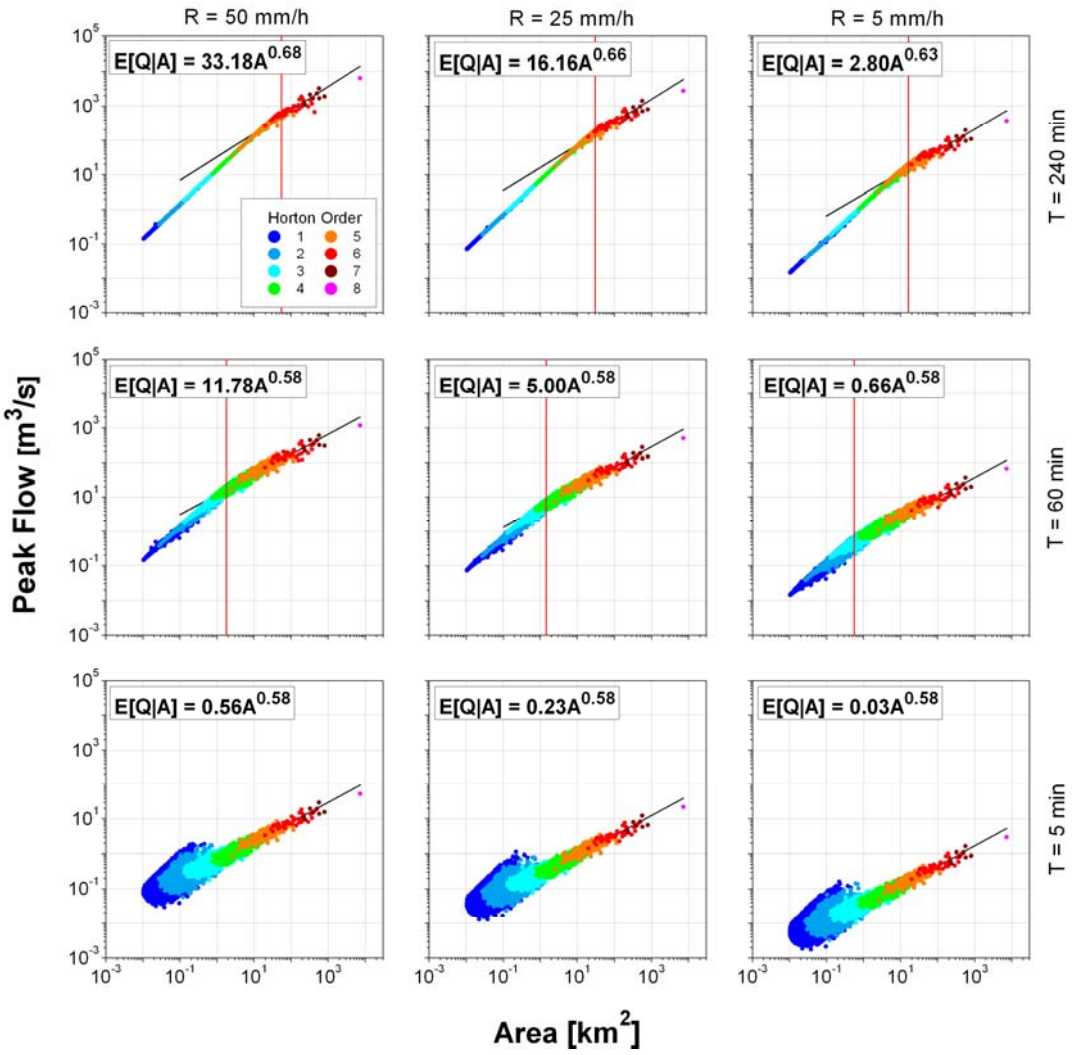

Figure 4.4 Sensitivity of peak flow scaling structure to intensity and duration of spatially uniform rainfall with no hillslope delay (instantaneous flow to the channels), and nonlinear channel routing. The solid black line represents the ordinary least squares fit (equation on each panel) performed in the Hortonian framework. The solid red line indicates the scale break.

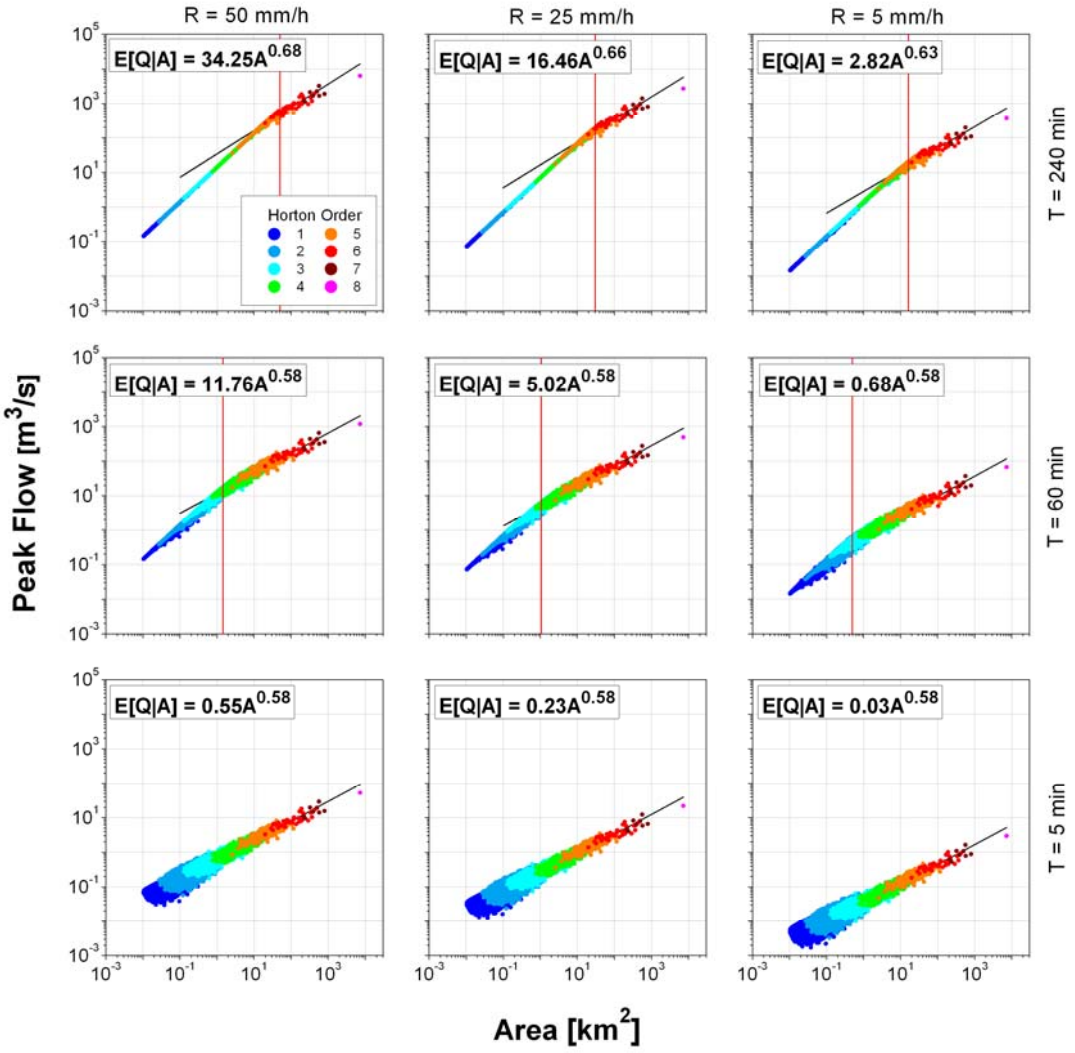

Figure 4.5 Sensitivity of peak flow scaling structure to intensity and duration of spatially uniform rainfall, hillslope velocity of 1.0 m/s, and nonlinear channel routing. The solid black line represents the ordinary least squares fit (equation on each panel) performed in the Hortonian framework. The solid red line indicates the scale break.

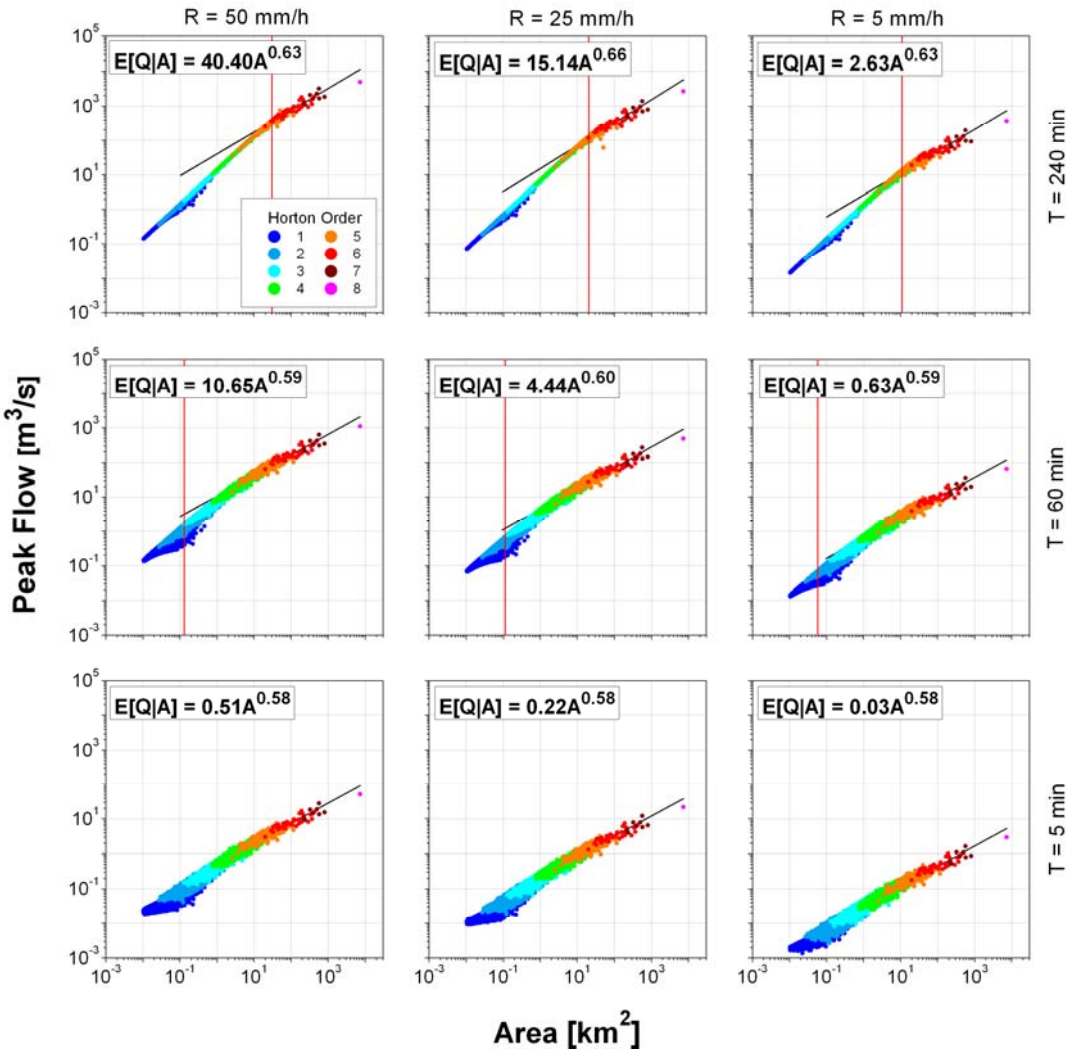

Figure 4.6 Sensitivity of peak flow scaling structure to intensity and duration of spatially uniform rainfall, hillslope velocity of 0.1 m/s, and nonlinear channel routing. The solid black line represents the ordinary least squares fit (equation on each panel) performed in the Hortonian framework. The scale break is indicated with the solid red line.

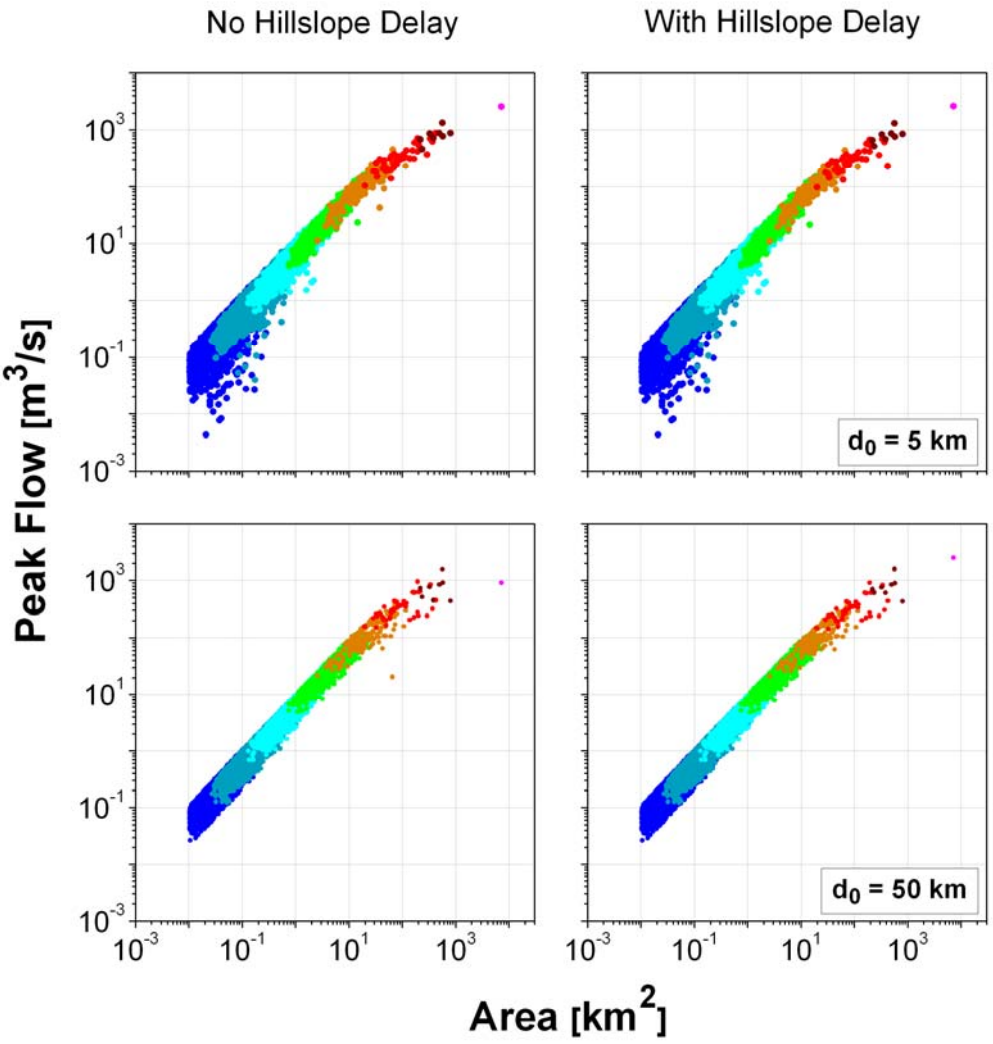

Figure 4.7 Sensitivity of peak flow scaling structure to the smoothness of Gaussian fields. The rainfall field is Gaussian with a mean of 25 mm/h and standard deviation of 6 mm/hr. The correlation function is assumed to be exponential with a correlation distance indicated in the figure. For the right panels, the velocity on hillslopes is equal to 0.1 m/s. For the color scheme, please refer to Figure 4.6.

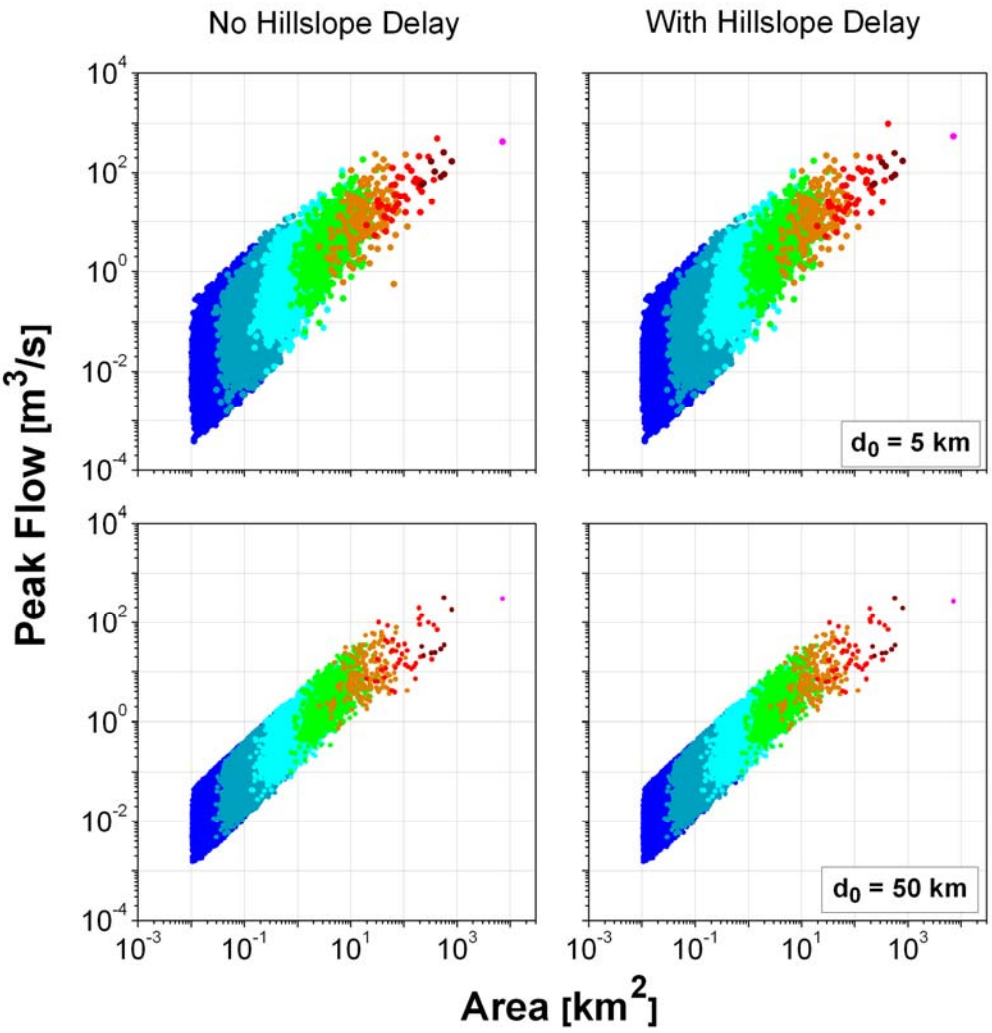

Figure 4.8 Sensitivity of peak flow scaling structure to the lognormal correlated fields. The mean of the rainfall field is 5 mm/h and the standard deviation is 8 mm/hr. The correlation function is assumed to be exponential with a correlation distance indicated in the figure. For the right panels, the velocity on hillslopes is equal to 0.1 m/s. For the color scheme, please refer to the Figure 4.6.

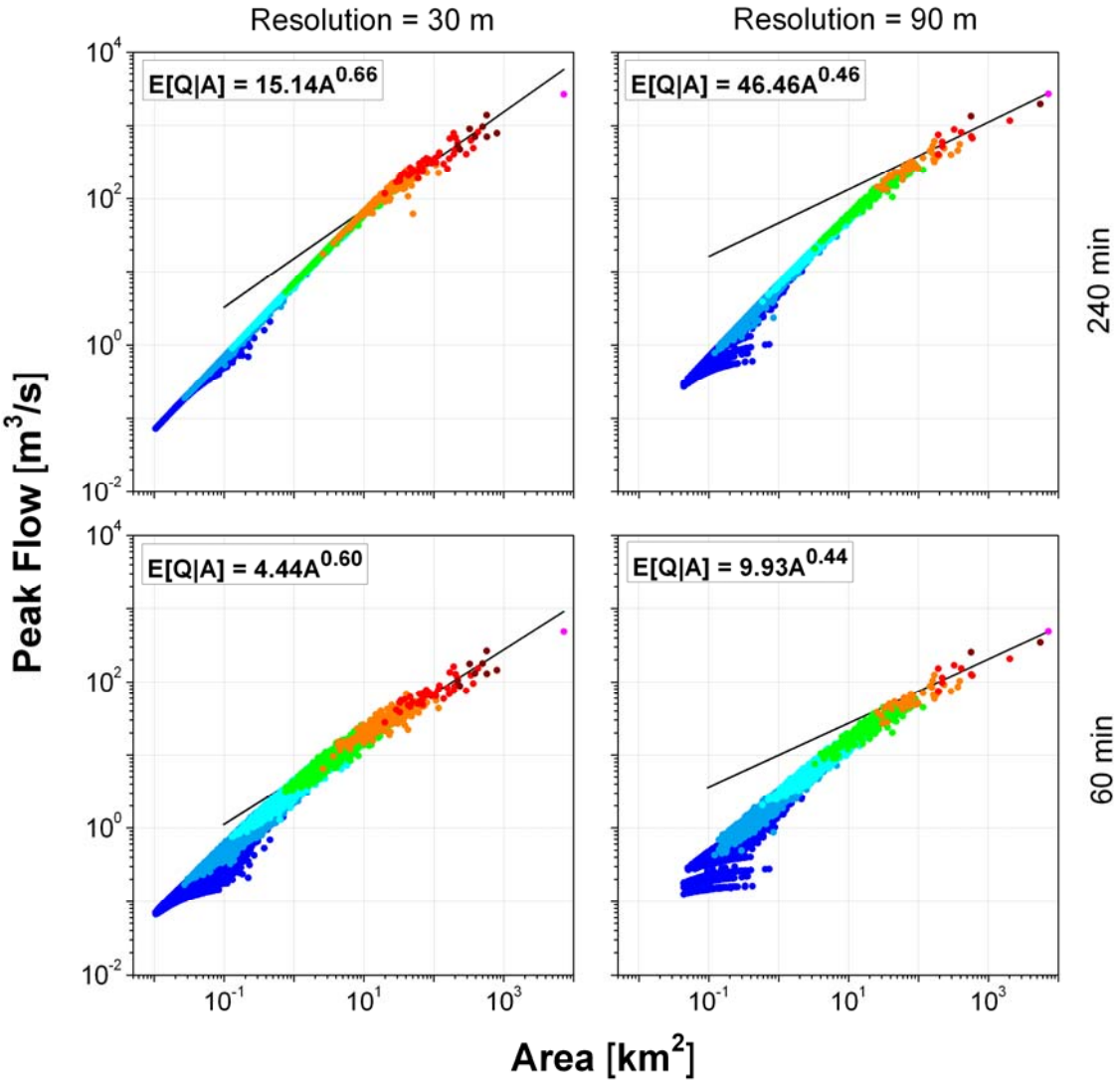

Figure 4.9 Sensitivity of the peak flow scaling structure to the channel network extraction. For the left panels, a 30 m resolution DEM was used to extract the river network, whereas, for the left panels, a 90 m DEM was used. The rainfall field is spatially uniform with an intensity of 25 mm/h for duration mentioned in the figure. The color scheme is same as Figure 4.6.

CHAPTER 5

ESTIMATION OF RADAR-RAINFALL ERROR SPATIAL CORRELATION*

5.1 Introduction

It is widely acknowledged that radar-rainfall estimates, which serve as an input to hydrometeorological and water resource applications, are corrupted by high uncertainties that originate from many sources [e.g., *Zawadzki*, 1984; *Austin*, 1987; *Fabry et al*. 1992; *Kitchen and Jackson*, 1993; *Hunter*, 1996; *Smith et al.*, 1996; *Steiner et al.*, 1999; *Young et al.*, 1999; *Young et al.*, 2000; *Krajewski and Smith*, 2002; *Borga*, 2002]. Given the current state of technology, it is practically impossible to accurately measure the true area-averaged rainfall. Therefore, the quantification of radar-rainfall* uncertainties (including their spatiotemporal structure) must be executed by approximating the true areal rainfall with point rainfall measurements from rain gauges. The problems associated with this approximation are well recognized [e.g., *Zawadzki*, 1975; *Austin*, 1987; *Kitchen and Blackall*, 1992; *Ciach and Krajewski*, 1999a,b; *Villarini and Krajewski*, 2008; *Villarini et al.*, 2008), but are difficult to overcome [e.g., *Ciach et al.* , 2003; *Habib et al.*, 2004; *Ciach and Krajewski*, 2006]. A reliable methodology for the comprehensive characterization of the radar-rainfall uncertainties must start with the identification and estimation of the joint probability distribution of radar-estimated rainfall and the corresponding true rainfall over a broad range of conditions [*Ciach* et al., 2007]. These conditions include different space and time scales, different distances from radar and rainfall regimes, and different radar systems and rainfall estimation algorithms. From the joint distribution, one can derive different synthetic characteristics of the radar

*Adapted from Mandapaka, P.V., W.F. Krajewski, G.J. Ciach, G. Villarini, and J.A. Smith (2009), Estimation of radar-rainfall error spatial correlation, *Advances in Water Resources*, 32, 1020-1030.

rainfall errors, from the simple performance criteria (such as the root mean squared error) to the error distributional properties and its spatiotemporal dependences. Clearly, establishing such a comprehensive methodology poses a challenge. We are confident that this challenge is surmountable, though only through many years of collaborative research. The present study constitutes another small but significant step in this direction.

A complete statistical characterization of radar-rainfall uncertainties must include the biases [e.g., *Smith and Krajewski*, 1991; *Anagnostou et al.*, 1998; *Seo et al.*, 1999; *Ciach et al.*, 2000; *Borga and Tonelli*, 2000], the error variance [e.g., *Ciach and Krajewski*, 1999; *Anagnostou et al.*, 1999; *Chumchean et al.*, 2003], conditional distributions of the errors [*Ciach et al.*, 2007], and a description of the error dependences in space and time. Because correlation functions are simple expressions of these spatiotemporal error dependences, they provide a good starting point.

In their study on probabilistic quantitative precipitation estimation, *Ciach et al.* [2007] developed a model for radar-rainfall errors, in which the relation between true rainfall and radar-rainfall is described by two elements, a deterministic distortion function and a random component. In addition to estimating the frequency distribution of the random component, they showed empirically that this component is correlated in space and time. Other radar-rainfall and satellite-rainfall error models that have been commonly used in the error propagation studies assumed either uncorrelated errors [e.g., *Sharif et al.*, 2002; *Georgakakos and Carpenter*, 2003; *Carpenter and Georgakakos*, 2004] or considered arbitrary models of the error correlation structure [e.g., *Nijssen and Lettenmaier*, 2003; *Gebremichael and Krajewski*, 2004; *Hossain et al.*, 2004; *Carpenter and Georgakakos*, 2006; *Villarini et al.*, 2007a]. However, many researchers realize that more accurate knowledge of the error correlations is crucial to the successful application of any rainfall estimate.

For example, *Nijssen and Lettenmaier* [2003] conducted a Monte Carlo simulation experiment to assess the effects that spatial sampling errors in precipitation have on hydrological predictions. They perturbed gauge-based interpolated rainfall estimates with synthetically generated uncorrelated and correlated error fields. The perturbed rainfall fields were then used as an input to a macroscale hydrologic model. *Nijssen and Lettenmaier* [2003] concluded that spatial correlations in the errors strongly affect predictions of hydrological fluxes and states. *Hossain et al.* [2004], in their study on the effect of satellite-rainfall uncertainties on the flood prediction uncertainty, employed a simple multiplicative error model with temporal error correlation modeled as a lag-one autocorrelation function. While they showed that the effect of temporal error correlation is insignificant for 3-h and 6-h sampling intervals, they indicated that it might be significant at the hourly sampling interval. *Carpenter and Georgakakos* [2006] also studied the sensitivity of ensemble flow predictions to input uncertainties. In their radar-rainfall error model, they assumed a linear spatial correlation function for the errors and concluded that the radar-rainfall error structure affects the mean areal precipitation uncertainties at the sub-catchment and watershed scales. *Villarini et al.* [2007a] studied the effects of radar-rainfall errors on the spatial scaling properties estimated from the radar-based rainfall maps. They perturbed the radar-rainfall fields with lognormally distributed multiplicative errors, assuming different levels of their spatial correlation. The study showed that the error itself has a strong effect on the scaling function of the perturbed radar-rainfall fields. However, they found that increasing the correlation level results in a decrease of the distortions in the scaling function. Based on the aforementioned literature, we argue that it is important to verify whether the radar-rainfall uncertainties are correlated in space and, if they are, to quantify this correlation as a function of the separation distance.

The main objective of this chapter is to present a practical method for estimating the radar-rainfall error spatial correlation (ESC), accounting for the area-point differences

inherent in any radar-gauge comparisons. In their preliminary error correlation estimates, *Ciach et al.* [2007] simply ignored the gauge representativeness errors. The method discussed here is conceptually analogous to the error variance separation (EVS) method proposed by *Ciach and Krajewski* [1999a]. In estimating the radar-rainfall error variance, use of the EVS procedure can account for the gauge representativeness errors. Defining the radar-rainfall error as the difference between the true areal rainfall and corresponding grid-averaged gauge rainfall, the EVS-corrected error variance is obtained by subtracting the area-point difference variance from the variance of the difference between the grid-averaged radar-rainfall and the corresponding rain gauge rainfall. The EVS method is based on the assumption that the radar and gauge errors are uncorrelated. Although in practice this zero-correlation assumption can be not fully satisfied, the EVS-corrected radar-rainfall error variance is considerably more accurate than the uncorrected error variance obtained directly from the radar-gauge pairs [*Ciach et al.*, 2003]. Many studies have employed the EVS method to characterize the variance of the radar and satellite derived rainfall products [e.g., *Young et al.*, 1999; *Krajewski et al.*, 2000; *Habib and Krajewski*, 2002; *Seo and Breidenbach*, 2002; *Gebremichael et al.*, 2003, *Chumchean et al.*, 2003; *Gebremichael and Krajewski*, 2004; *Zhang et al.*, 2007]. In the method for radar-rainfall ESC estimation presented here, we also account for the area-point differences.

The radar-rainfall uncertainties can be represented as either multiplicative or additive errors. The EVS method was originally proposed by *Ciach and Krajewski* [1999a] for the additive errors only. For this error representation, the underlying assumptions of the EVS scheme are systematically discussed in *Ciach et al.* [2003]. For the multiplicative errors, the EVS method can be used in the logarithmic domain [*Anagnostou et al.*, 1999]. However, the possible adverse consequences of the logarithmic transformation remain unverified, which prompted us to follow the additive definition of radar-rainfall errors as in *Ciach and Krajewski* [1999a].

This chapter is organized as follows: in Section 5.2, we derive the expression for the estimation of the radar-rainfall ESC. Section 5.3 describes the procedure to estimate various terms in the expression for ESC. In Section 5.4, we test our method for removing the area-point effects from the ESC estimates using a simulation framework. In Section 5.5, we apply the method to estimate ESC in the National Weather Service's operational radar-rainfall products that are provided by the precipitation processing system (PPS) of the NEXRAD in the form of the hourly digital precipitation arrays (DPA). As a ground reference for these estimates, we use the data from an experimental rain gauge network in Oklahoma known as the ARS Micronet, which is operated by the Agricultural Research Service of the United Stated Department of Agriculture. Section 5.6 summarizes the results and concludes this study.

5.2 ESC Derivation

The formula for error spatial correlation can be derived using an extension of the EVS method since we start by partitioning the variogram of the radar-rainfall. Let the radar estimates over the two pixels 1 and 2 (Figure 5.1) be R_1 and R_2, respectively. The variance of the difference R_1-R_2 can be partitioned to incorporate radar-rainfall error terms as follows:

$$\begin{aligned} Var\{R_1 - R_2\} &= Var\{(R_1 - T_1) - (R_2 - T_2) + (T_1 - T_2)\} \\ &= Var\{R_1 - T_1\} + Var\{R_2 - T_2\} + Var\{T_1 - T_2\} \\ &\quad -2Cov\{R_1 - T_1, R_2 - T_2\} \\ &\quad +2Cov\{R_1 - T_1, T_1 - T_2\} \\ &\quad -2Cov\{R_2 - T_2, T_1 - T_2\} \end{aligned} \tag{5.1}$$

where R_1 and R_2 are radar-rainfall estimates and T_1 and T_2 are true area-averaged rainfall accumulations over the pixels 1 and 2, and operators $Var\{\cdot\}$ and $Cov\{\cdot\}$ are the variance and the correlation of the corresponding random variables. Equation 4.1 contains three variance terms and the three covariance terms. $Var\{R_1 - T_1\}$ and

$Var\{R_2 - T_2\}$ are the variances of radar-rainfall errors over pixels 1 and 2, respectively, and $Var\{T_1 - T_2\}$ is the variogram of the true rainfall. The three covariance terms are:

1. $Cov\{R_1 - T_1, R_2 - T_2\}$ is the covariance between the radar-rainfall errors $R_1 - T_1$ and $R_2 - T_2$, i.e., the error spatial covariance.
2. $Cov\{R_1 - T_1, T_1 - T_2\}$ is the covariance between the radar-rainfall error over grid 1 and the difference $T_1 - T_2$.
3. $Cov\{R_2 - T_2, T_1 - T_2\}$ is the covariance between the radar-rainfall error over grid point 2 and the difference $T_1 - T_2$.

We assume that the covariances (2) and (3) above are equal, and hence the sum $2Cov\{R_1 - T_1, T_1 - T_2\} - 2Cov\{R_2 - T_2, T_1 - T_2\}$ in the equation 4.1 is equal to zero. With this assumption, equation 1 is rearranged to obtain the error spatial covariance as follows:

$$Cov\{R_1 - T_1, R_2 - T_2\} = \frac{1}{2}\left[Var\{R_1 - T_1\} + Var\{R_2 - T_2\}\right] + \frac{1}{2}\left[Var\{T_1 - T_2\} - Var\{R_1 - R_2\}\right] \quad (5.2)$$

The radar-rainfall error spatial covariance can be obtained by adding the following two terms: a) average of error variances at locations 1 and 2 (Figure 5.1) and b) difference of semivariograms of the true areal rainfall and the radar-rainfall. The error correlation can be obtained by normalizing the covariance with the product of the square root of the error variances.

5.3 ESC Estimation

Once we derived the governing formulae, we needed to estimate the terms in equation 5.2. All four terms on the right-hand side of the equation 5.2 for the ESC can be estimated directly from data.

5.3.1 Estimation of Error Variance

We employed the EVS method proposed by *Ciach and Krajewski* [1999a] to estimate the terms $Var\{R_1 - T_1\}$ and $Var\{R_2 - T_2\}$ in equation 4.2. According to this method,

$$Var\{R_i - T_i\} = Var\{R_i - G_i\} - Var\{G_i - T_i\}, \tag{5.3}$$

i.e., the variance of the radar-rainfall error is obtained by separating the gauge-sampling error, $Var\{G_i - T_i\}$, from the overall radar-gauge variability, $Var\{R_i - G_i\}$. The latter can be obtained directly from the data:

$$Var\{R_i - G_i\} = \frac{1}{N}\sum_{t=1}^{N}\left[R_i(t) - G_i(t)\right]^2 \tag{5.4}$$

where $R_i(t)$ and $G_i(t)$, respectively, are radar- and gauge-rainfall estimates for the pixel i at the time step t. Equation 5.3 assumes that the radar-rainfall errors ($R_i - T_i$) are uncorrelated with the rain gauge errors ($G_i - T_i$). This zero-correlation assumption and its validity have been discussed in the Introduction. In equation 5.4, temporal stationarity is assumed only to increase the sample size for the estimation of the $Var\{R_i - G_i\}$. If the rain gauge network is dense enough, the proposed method can be slightly modified and applied at each time step. It should also be noted that if there is more than one gauge in a pixel, then $G_i(t)$ represents average rainfall from all the gauges in the pixel i at the time step t. With the current computational resources, simple averaging of gauge measurements can be replaced with the block kriging. However, we used simple averages for the reasons laid out in the next paragraph.

Assuming the second-order stationarity and isotropy within the areal domain and following *Morrissey et al* [1995], the gauge-sampling error $Var\{G_i - T_i\}$ is obtained as

$$Var\{G_i - T_i\} = Var\{G\} \cdot VRF_i, \tag{5.5}$$

where $Var\{G\}$ is the variance of the point-rainfall and VRF_i is the variance reduction factor for the pixel i containing the rain gauge(s), and it is obtained as follows:

$$VRF_i = \frac{1}{n^2}\sum_{j=1}^{K}\sum_{k=1}^{K}\rho(d_{j,k})\delta(j)\delta(k) - \frac{2}{nK}\sum_{j=1}^{K}\sum_{k=1}^{K}\rho(d_{j,k})\delta(j) + \frac{1}{K} + \frac{2}{K^2}\sum_{j=1}^{K-1}\sum_{k=j+1}^{K}\rho(d_{j,k}), \quad (5.6)$$

The VRF is a statistic that quantifies the spatial sampling error we make in approximating an areal value with an average of a specified number of point measurements in that area [*Morrissey et al.*, 1995]. Such a theoretical framework does not exist if the areal value is estimated from block kriging. For this reason we used simple averages in estimating $G_i(t)$ in the equation 5.4. The sampling domain is divided into *K* grid boxes, over which the rainfall is estimated as the arithmetic mean of *n* gauges. The Kronecker delta function $\delta(j)$ denotes whether box j contains a rain gauge. The term $\rho(d_{j,k})$ represents the rainfall spatial correlation, where $d_{j,k}$ is the distance between boxes j and k. A detailed description of equation 5.6 is given in *Morrissey et al.* [1995] and *Krajewski et al.* [2000].

5.3.2 Estimation of Variogram of True Areal Rainfall

The third term on the right hand side of equation 4.2 denotes the variogram of the true rainfall, and for a second-order stationary process it can be shown that [e.g., *Cressie*, 1993]

$$Var\{T_i - T_j\} = 2\Gamma_T(h) = 2\left[Cov_T(0) - Cov_T(h)\right], \quad (5.7)$$

where $Cov_T(h)$ is the covariance of the true areal rainfall over grids i and j separated by distance h, and $Cov_T(0)$ is the variance of the true rainfall. Section 6.4 of *Vanmarcke* [1983] gives the following expression to obtain the covariance function of the local averages of a homogeneous two-dimensional random field over square areas with sides L, from the corresponding point variance and covariance function:

$$Cov\left\{T_i, T_j\right\} = Cov_T(h) = \frac{Var\{G\}}{4L^4}\left[\sum_{a=0}^{3}\sum_{b=0}^{3}(-1)^a(-1)^b\,\Delta\left(L_{xa}, L_{yb}\right)\right], \quad (5.8)$$

where T_i and T_j respectively are true areal rainfall over pixels i and j (of sides L) separated by distance lag h. To achieve lags smaller than the pixel size, the pixels i and j need to be overlapping, in which case, the equation 5.8 can be simplified to:

$$Cov\{T_i, T_j\} = Cov_T(h) = \frac{Var\{G\}}{4L^4}\begin{bmatrix} \Delta(L_{x0}, L_{y0}) - 2\Delta(L_{x0}, L_{y1}) + \Delta(L_{x0}, L_{y2}) \\ -2\Delta(L_{x1}, L_{y0}) + 4\Delta(L_{x1}, L_{y1}) - 2\Delta(L_{x1}, L_{y2}) \\ +\Delta(L_{x2}, L_{y0}) - 2\Delta(L_{x2}, L_{y1}) + \Delta(L_{x2}, L_{y2}) \end{bmatrix}, \quad (5.9)$$

In the equations 5.8 and 5.9, L_{x0} (L_{y0}) is the distance from the end of the first grid to the beginning of the second grid along x (y) direction; L_{x1} (L_{y1}) is the distance from the beginning of the first grid to the beginning of the second grid along x (y) direction; L_{x2} (L_{y2}) is the distance from the beginning of the first grid to the end of the second grid along x (y) direction; and L_{x3} (L_{y3}) is the distance from the end of the first grid to the end of the second grid along x (y) direction. The $\Delta(\cdot,\cdot)$ function in the equation 5a is obtained as $\Delta(L_x, L_y) = (L_x L_y)^2 \cdot \gamma(L_x, L_y)$, where

$$\gamma(L_x, L_y) = \frac{4}{L_x L_y}\int_0^{L_x}\int_0^{L_y}\left(1 - \frac{l_x}{L_x}\right)\left(1 - \frac{l_y}{L_y}\right)\rho(l)dl_x dl_y, \quad (5.10)$$

where l_x and l_y are the separation distances in x and y directions, with $l = (l_x^2 + l_y^2)^{1/2}$. The variance of the areal values can be obtained from the point variance using the relation:

$$Var\{T\} = Cov_T(0) = Var\{G\} \cdot \gamma(L, L), \quad (5.11)$$

For further information on equations 5.8 through 5.11, the reader is pointed to Section 6.4 of *Vanmarcke* [1983]. The covariance and the variance of the true areal process are then plugged in equation 5.7 to obtain the variogram of the true areal rainfall. Equations 5.6-5.11 demonstrate that the correlation function of the point rainfall (ρ) plays an important role in estimating the ESC. In this study, we employ Pearson's correlation coefficient and justify its use in Section 5.5, where the correlation structure of point rainfall is estimated using rain gauges from the Oklahoma Micronet.

5.3.3 Estimation of Variogram of Radar-Rainfall

The fourth term within the brackets on the right hand side of equation 2 denotes variogram of the radar-rainfall, which we estimate using the classical method-of-moments estimator [e.g., Cressie, 1983].

$$Var\{R_i - R_j\} = 2\Gamma_R(h) = \frac{1}{n(h)}\sum_{i,j}^{n(h)}[R_i - R_j]^2, \tag{5.12}$$

where $n(h)$ is the number of data pairs separated by lag h. We realize that there are estimators which are more robust to the outliers in the data than the method-of-moments estimator. The estimator proposed by *Cressie and Hawkins* [1980] ($\hat{\gamma}_{CH}$) is among the widely used. However, we did not use it in this study as it is valid only when the distribution of differences, $Z(x)-Z(x+h)$, for a particular lag h is Gaussian (Lark, 2000). *Lark* [2000a] compared the robustness and efficiency of $\hat{\gamma}_{CH}$ along with two other robust estimators, which also assume contaminated normal model, i.e., normal distribution in the presence of outliers. The study concluded that all the three robust estimators are influenced by any departure from the assumption of normality of differences [*Lark*, 2000a]. The study also concluded that for the skewed data which has no outliers, the classical method-of-moments estimator performed better than the three robust estimators.

5.3.4 Summary of Assumptions in ESC Methodology

Before we apply the method on the simulated fields and on the radar-rainfall data, we summarize here, all the assumptions involved.

a) The rainfall is second-order stationary in space and time.

b) The rainfall is isotropic. If the rainfall is anisotropic, the spatial correlation function is directional and therefore the proposed methodology has to be applied for each direction resulting in different error spatial correlation functions for different directions.

c) Radar-rainfall estimates and the rain gauge measurements are unbiased.

d) The rain gauge representativeness error is not correlated with radar-rainfall error. The assumptions (c) and (d) are required by the EVS method.

5.4 ESC Testing Using Monte Carlo Simulation

We begin with a Monte Carlo simulation experiment to assess the method for the error spatial correlation and the estimation procedure discussed in the previous section. The simulation framework provides us with full control over the things that real data cannot. In the simulation experiment, we can compare the error correlations retrieved from our methodology with the true error correlations known a priori. A good performance here implies that the assumptions in our methodology are reasonable and gives us the confidence to apply the ESC methodology on the real radar-rainfall data. On a square grid of 100×100 (arbitrary units), high resolution (0.4×0.4) two-dimensional, stationary, Gaussian and lognormal processes with an exponential correlation function are generated using the circulant embedding technique [e.g., *Dietrich and Newsam*, 1993; *Wood and Chan*, 1994]. The exponential correlation function used in the study is same as the one given by equation 3.1 in Section 3.3.

A total of 325 stations with inter-station spacing of 4 units in the x-direction and 8 units in the y-direction (Figure 5.2) sample the process from the simulated high resolution fields, which is similar to the sampling of rainfall process by the rain gauges. The high resolution fields are then averaged to obtain fields at a resolution of 4×4 as it is the most commonly used resolution for radar-rainfall data in many hydrological applications. These average fields are analogous to the true areal rainfall fields at a particular resolution. Hereafter, the high resolution and corresponding average fields are referred to as HR and true areal (TA) processes, respectively.

We do not interpret the simulated Gaussian and lognormal fields as rainfall fields. Our objective in the simulation experiment is to evaluate the methodology on simple case

of Gaussian realizations and on a complex case of lognormal fields. The ESC approach can be applied to any remotely sensed physical process that meets the criteria such as second-order stationarity and isotropy and where there are inherent area-to-point errors in the evaluation. That's the reason we labeled the process from where we sample the "point data" as high resolution (HR) process and the corresponding area averaged process as true areal (TA) process and the error corrupted fields as areal-with-error (AE) process.

The next step in the simulation framework is to generate the error process at a resolution of 4×4, assuming that they are Gaussian with zero mean, unit variance and following an exponential correlation function (equation 5.13) with a correlation distance of 20 distance units. Since the method is derived for additive errors, we add the error fields to the TA fields to obtain the "areal with error" (AE) process. Addition of Gaussian errors to the Gaussian (or lognormal) process results in some negative values in the AE process, which is not realistic for rainfall fields. However, this would not affect our results because our aim in this section is solely to test the method. The above procedure is repeated to obtain multiple realizations of AE fields.

We then applied the method on the simulated fields to estimate the error spatial correlation. Because the spatial correlation of the errors is known, the accuracy of our framework can be assessed by comparing the estimated and true ESC. We also compare them with spatial correlations of the difference between the 325 sampling points and corresponding AE pixels (analogous to radar-gauge difference), which we call "radar-gauge (RG)" error spatial correlation.

5.4.1 Gaussian Realizations

We start by simulating correlated Gaussian realizations with mean zero and variance equal to 1 and we consider two correlation distances of 10.0 and 40.0 (arbitrary distance units). For both cases, the correlation functions of HR, TA, error, and the AE fields are estimated using the 325 sampled points (Figure 5.3). Figure 5.3 illustrates that

the correlation structure of the high resolution process and the error process is quite accurate, which gives us confidence in the simulated fields. The variogram of the TA process obtained using the procedure described in Section 5.3 and that of the AE process obtained using the classical method-of-moments estimator are shown in Figure 5.4 for the two point correlation functions considered. From the TA variograms (particularly for the correlation distance of 10), one sees that the variance (half the sill) is smaller than the variance of the corresponding HR process (which is equal to 1.0). This is expected, as the TA process is obtained as an average of the HR process. From Figure 5.4, it is also apparent that an addition of correlated errors to the TA process increases its variance.

The error correlations obtained using our method is fitted with a three-parameter exponential function using the Levenberg-Marquardt algorithm and is compared with the spatial correlation of RG differences and that of true errors in Figure 5.5. It can be seen from the Figure 5.5 that our method can retrieve the error correlation structure quite accurately for both of the considered correlation functions. Though the correlation structure of RG differences is close to the true one, there is a consistent bias, which is quite evident for a smaller correlation distance of 10.0. This bias is due to the fact that RG differences have inherent area-point errors.

5.4.2 Lognormal Realizations

In testing our method on the simulated lognormal realizations, we considered the same correlation functions that are used in the Gaussian case. In addition to the correlation distances, we also varied the coefficient of variation (CV) of the lognormal realizations. Therefore, for simulations with lognormal realizations, we have four different combinations of variability, as indicated in Table 5.1. The number of realizations used for each case is also shown in Table 5.1.

We obtain the lognormal process by transforming a simulated Gaussian process. However, the transformation results in the alteration of the parameters of the lognormal

realizations. As our aim is to obtain lognormal process with a certain mean, variance, and correlation structure, we selected the parameters of the simulated Gaussian fields such that the transformation will lead to lognormal fields with the desired parameters. The procedure is explained in detail in Appendix A.

The correlation functions of the HR, TA, error, and AE fields estimated using 325 sampling locations are plotted in Figure 5.6. The HR correlation functions estimated from the simulations closely match the theoretical correlation functions. The figure also reveals that as the CV increases and correlation distance decreases, the addition of the error field has little impact on the correlation function of the true areal process. Figure 5.7 illustrates the variograms of TA and the AE process for all combinations of correlation functions and CVs. The behavior of the TA and AE variograms for all four cases is similar to that of the Gaussian simulations with the variance of the TA process being smaller than the corresponding HR process. The ESC technique is applied on all four cases (Table 5.1) and is shown in Figure 5.8 along with the fitted three parameter exponential correlation function and correlation of RG differences. Comparing the correlations of RG differences with the true one, a systematic underestimation can be seen from the Figure 5.8. This bias increases as the lognormal field becomes more variable (a high CV and small correlation distance). As in the Gaussian case, the biased RG correlations result from area-point errors. Unlike in the Gaussian case, the ESC results in a consistent bias for a larger CV and smaller correlation distance (Figure 5.8). Further investigation into case 4 of lognormal simulations revealed that bias in the estimation of variance reduction factor is the main reason for the systematic underestimation in ESC and, hereafter, all the results in this section are for case 4.

As mentioned in Section 5.3, the VRF is a theoretical framework to quantify the spatial sampling error involved in approximating an areal value with an average of point measurements within the area. For the VRF to be strictly applicable in the simulation framework, the areal values need to be obtained by averaging infinite number of point

values. For the Gaussian simulations, this is not an issue as the TA value (4.0×4.0) - obtained as an average of 100 pixels of HR process (0.4×0.4) - is reasonably closer to the true areal process. Case 4 in the lognormal simulations, with a high CV and small correlation distance, is highly variable, and the 100 HR pixels are inadequate to obtain the TA process. Therefore, the VRF estimated as in equation 5.6 overcorrects the area-point discrepancies leading to underestimation of error correlation (Figure 5.8d). For the lognormal simulations, we repeat case 4 (Table 5.1 and Figure 5.8d) with the TA process obtained by averaging the HR process simulated on 0.1×0.1 grids and Figure 5.9 shows that the bias in ESC reduces with an increase in the resolution. Therefore, our method's systematic underestimation of error spatial correlation for lognormal realizations is largely an artifact of the simulation experiment, and the ESC method performs well in a simulation framework.

5.5 Application to the Oklahoma ARS Micronet Dataset

After testing the method in the simulation framework, we applied it to estimate the ESC of NEXRAD hourly digital product (DPA) from the Oklahoma City NEXRAD (KTLX) radar site (Fulton et al., 1998). We used a high density and high quality rain gauge network, the Oklahoma Micronet, established by the Agricultural Research Service in Little Washita experimental watershed, Oklahoma [e.g., *Allen and Naney*, 1991; *Young et al.*, 2000; *Ciach et al.*, 2003]. Figure 5.10 shows the network and the hydrologic rainfall analysis project (HRAP) radar grid, which is a quasi-rectangular grid with the size of the cell ranging from 3.5 km in the southern U.S. to 4.5 km in the northern U.S. [e.g., *Reed and Maidment*, 1999]. The rain gauge network covering an area of approximately 1200 km^2, consists of 41 gauges on a fairly regular grid with intergauge distances ranging from around 3 to 40 km and the average nearest-neighbor distance of approximately 5 km. In our analysis, we considered a square radar domain with sides of

60 km, covering the rain gauge network, for a 36-month period (April-September data for six years 1998-2003), thereby reducing any seasonal effects.

The first step in applying the ESC method is to estimate the point correlation function of rainfall. Though Pearson's estimator is widely used to estimate the correlation, it has been shown in the literature that for highly skewed distributions, it results in biased estimates [e.g., *Hutchinson*, 1997; *Habib et al.*, 2001]. Assuming that the data come from the mixed lognormal distribution, *Habib et al.* [2001] proposed an alternative unbiased approach to estimate the correlation. To check if the lognormality assumption is valid for the Oklahoma Micronet, the rainfall data measured by the rain gauges is binned into intervals of 0.5, and a mixed lognormal distribution is fitted to this binned data. The gauge-rainfall had a lighter tail with respect to a mixed lognormal distribution. Therefore, we stick to the Pearson's estimator to obtain the point correlation function of the rainfall. Figure 5.11 portrays the correlation function of the rainfall measured by the rain gauges and the fitted three-parameter exponential function (Equation 3.1 with an additional nugget parameter). We used this parametric form of the correlation structure to obtain the error variance and the variogram of the true areal rainfall (equations 5.3 through 5.11). The radar pixel is nine orders of magnitude larger than the sampling area of rain gauge. Therefore, artifacts introduced by equation 5.6 in the estimation of VRF and subsequently the spatial correlation of errors are negligible.

The next step in the ESC method is the estimation of the variogram of the radar-rainfall. Its estimation is performed by using a classical method-of-moments estimator (equation 5.12) and binning the moments into 4 km distance classes. All the pixels covering 60×60 km^2 are used in the estimation process. The size of the estimation domain and number of realizations play an important role in the estimation of variance and covariance leading to their severe underestimation. We did a systematic sensitivity study on this aspect and found that the bias decreases rapidly with the number of independent realizations. We explain the results of that study in detail in Appendix A.

For example, if the estimation domain is half the decorrelation distance, we found that as few as 50 independent fields would suffice to give us an unbiased estimate of variance and covariance. Since we used six years of hourly data, underestimation of variance and covariance is not an issue.

Figure 5.12 shows the estimated and fitted radar estimates along with the variogram of the true areal rainfall obtained using equation 5.7. The error variances and the variograms of the true areal and radar-rainfall are then used in the ESC method (equation 5.2) to estimate the spatial covariance of errors, which when normalized with error variances, results in an error spatial correlation structure (Figure 5.13). The spatial correlation of the RG differences is also plotted in the same figure for comparison. From Figure 5.13, one notices that our method results in tighter (less scatter) correlation estimates than those obtained directly from the radar-gauge pairs that neglect the area-point errors. The decreased scatter is due to the use of parametric point correlation functions in equations 4.3 through 4.11. However, for both cases, we can infer that radar-rainfall errors are significantly correlated with a correlation distance of approximately 20 km (Figure 5.13). This is an important result as in most of the error propagation studies, the errors are assumed to be uncorrelated.

Further, one can observe in Figure 5.13 that ignoring the area-point differences results in underestimation of the ESC at the shorter distances. Though this bias is not significant at the daily and, to an extent, hourly scales for stratiform rainfall, it might be significant at the sub-hourly time scales and for the tropical regime as the rainfall becomes more variable. Before we conclude, we would like to mention that the overall mean of the Micronet gauge estimates over 36 months is equal to 0.083 mm and that of collocated radar pixels for the same time period is 0.116 mm. The bias can be treated in an additive or multiplicative manner. The former approach would result in unrealistic negative rainfall values for the radar pixels. When the latter approach is used, the variogram of the DPA estimates would be almost the same as the variogram of true areal

rainfall. Our methodology (equation 5.2) applied in such a situation leads to the error covariance being constant and equal to the error variance. The result is in contrast to the RG error correlations estimated from multiplicatively corrected radar fields as they are decreasing with distance. Therefore, we applied the ESC approach directly on the uncorrected radar data. To make sure that the proposed methodology is robust to handle such bias we ran a simulation experiment similar to Section 5.4 but with uncorrelated errors and a bias of 0.5 to the AE fields. Despite the bias in the AE fields, the proposed methodology correctly identified the zero error correlations.

5.6 Summary and Conclusions

Defining the radar-rainfall error as the difference between the radar-estimated rainfall and the corresponding true areal rainfall, we propose a method to estimate the error spatial correlation (ESC) that accounts for the rain gauge representativeness errors. The required information on the area-point difference structure was obtained from relatively dense rain gauge networks. Although this study considered only the additive error definition, it can be applied to the multiplicative errors by transforming the variables into the logarithmic domain. Conceptually, our ESC estimation method is an extension of the error variance separation method proposed by *Ciach and Krajewski* [1999]. After formal derivation of the method, it was tested on simulated Gaussian and lognormal fields with known ESC. It performed very well in estimating the ESC for Gaussian fields and for lognormal fields with lower coefficients of variation (CV). However, as the CV of the lognormal fields increased, our test results tended to underestimate the ESC. The fact that this specific underestimation decreases systematically with increasing the resolution of the simulated fields shows that this effect is an artifact of the simulation caused by the finite resolution of the simulated fields. The ESC obtained using our method was also compared with the ESC obtained directly from the simulated "radar-gauge" differences. This comparison showed that ignoring the area-

point errors can result in considerable underestimation of the error spatial correlation structure, especially at small separation distances.

After the computer simulation tests, the ESC method was applied to the DPA products that are the standard outcomes of the PPS in the NEXRAD. This application was based on six years of warm season (April to September) KTLX radar data, and the corresponding data from a relatively dense experimental rain gauge network (Oklahoma ARS Micronet). These results show that the radar-rainfall errors are spatially correlated and that their correlation distance is approximately 20 km. Although this application has been limited to the central Oklahoma region, it offers insight into the spatial correlation of DPA radar-rainfall errors in any region where the rainfall is mostly of stratiform and mixed form. The results in this study can be helpful for new radar-rainfall error propagation studies that account for the ESC in the radar-rainfall data. It is worth mentioning that the radar errors are dependent on location and range from the radar. However we are applying the ESC method within the 60×60 km^2 pixel assuming stationarity and isotropy. If we can have dense networks such as Micronet in different zones (distances from radar), we can estimate the error spatial correlation for each zone and for each direction using the proposed methodology. We would also like to mention that the proposed method is not conditioned on the intensity and type of precipitation.

The ESC method presented here can be extended by regarding the ESC as a dynamically varying random function for which state estimation procedures are derived. The simplest way of implementing such an extension is by treating the correlation distance of the radar-rainfall error field as a random process in time. One can include this state variable as an element of the ESC estimation problem where the updating of the state is based on the currently observed radar-rainfall fields and the corresponding rain gauge data. This extension would eliminate the need for the data to be stratified into several "seasons." It would also reflect the fact that even during the peak of the warm season, the large extratropical cyclones with prevailing areas of stratiform rainfall regime

can be mixed with highly variable convective cells embedded within the stratiform rainfall fields.

Table 5.1* Parameters used in the four simulation scenarios with lognormal HR field.

σ_G^2	μ_{LN}	σ_{LN}^2	CV	Correlation distance		n
				40	10	
0.5	1.28	1.07	0.80	CASE 1	CASE 2	5000
1.0	1.65	4.67	1.31	CASE 3	CASE 4	25000

*The error field (Err) is Gaussian with zero mean and standard deviation equal to 1.0 and with the correlation distance equal to 20. The distance units are related to the size of the simulation area that is assumed to be 100 by 100 arbitrary units. The number of realizations is shown in the last column.

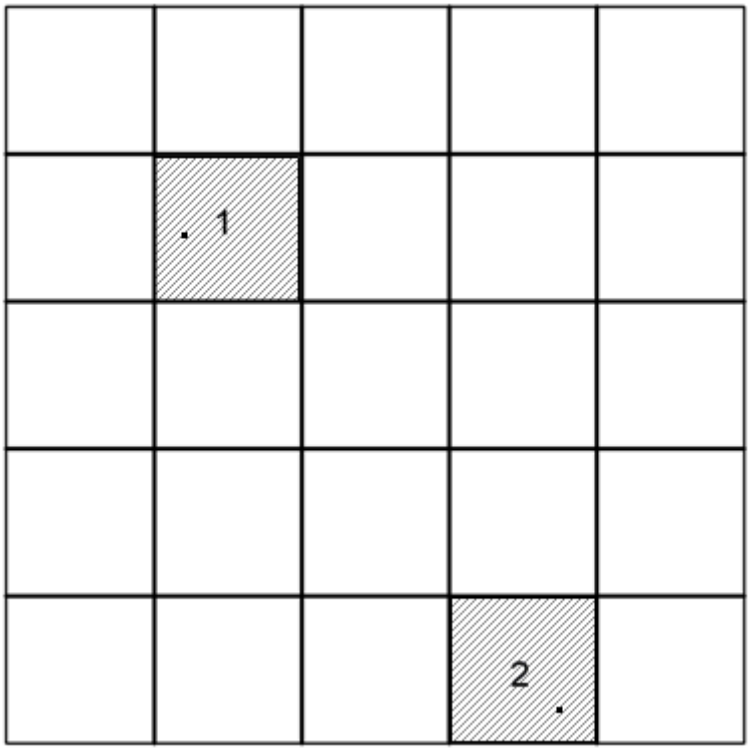

Figure 5.1 A schematic example of radar-rainfall grid with two pixels containing rain gauges (the hatched squares). The gauges are located randomly within the radar pixels.

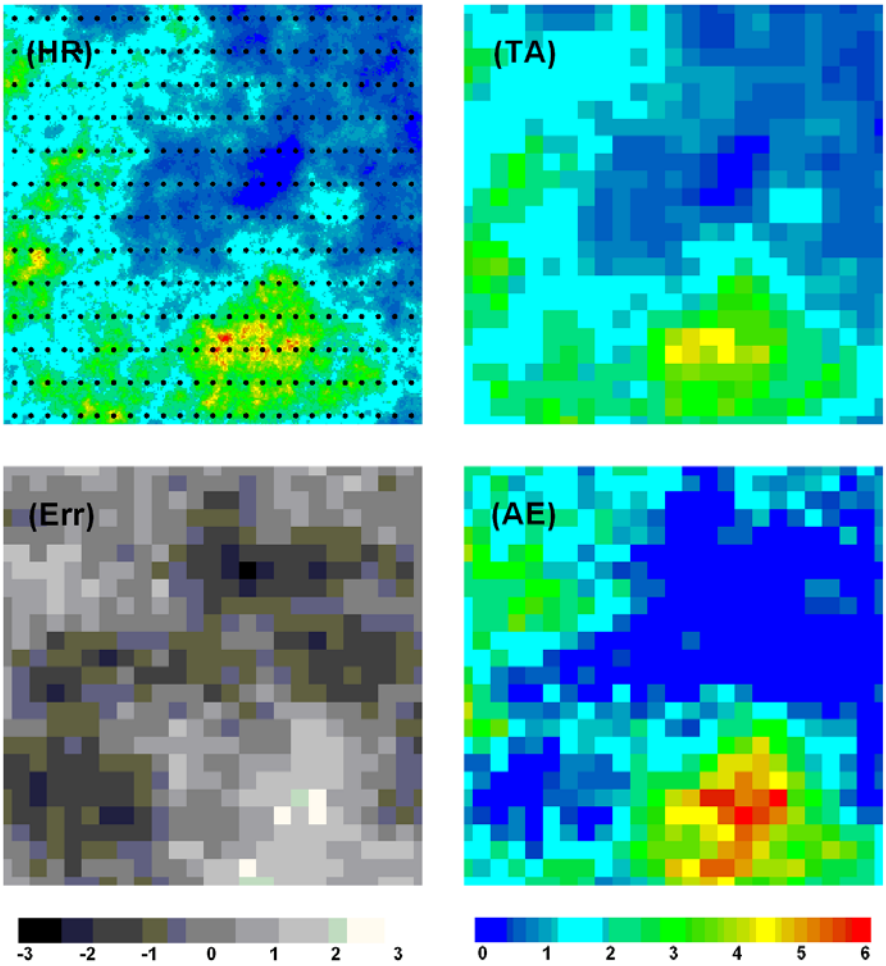

Figure 5.2 Example realization of four random fields: high resolution field (**HR**), true area averaged field (**TA**), error field (**Err**), and the areal with error (**AE**) field. The correlation distance of **HR** is d_{HR}=40, and its marginal distribution is lognormal with the mean equal to 1.65 and the coefficient of variation equal to 1.31. The **HR** panel also shows the sampling gauge network consisting of 325 (25 by 13) gauges. The error field (**Err**) is Gaussian with zero mean and a standard deviation equal to 1.0 and with the correlation distance of d_{Err}=20. The distance units are related to the size of the simulation area that is assumed to be 100 by 100 arbitrary units. The resolution of the **HR** field is equal to 0.4, and it is 4.0 for the remaining fields.

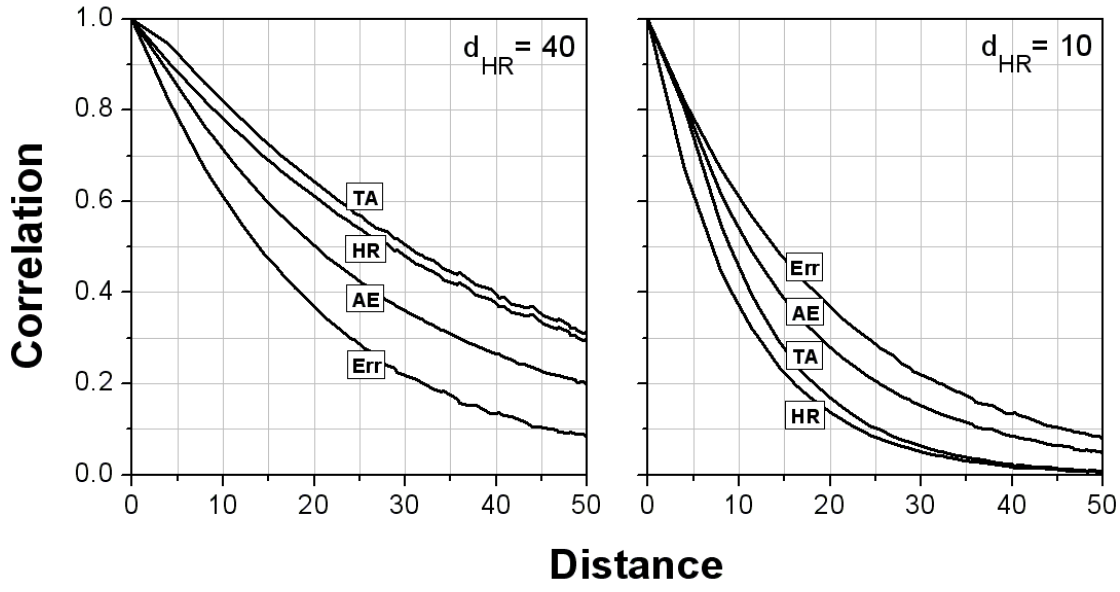

Figure 5.3 Correlation functions of the simulated high resolution (**HR**) Gaussian process, true areal (**TA**), error (**Err**), and areal with error (**AE**) processes estimated using the sampling network shown in Figure 2.2. The distance units and the parameters of the **Err** field are the same as in Figure 5.2.

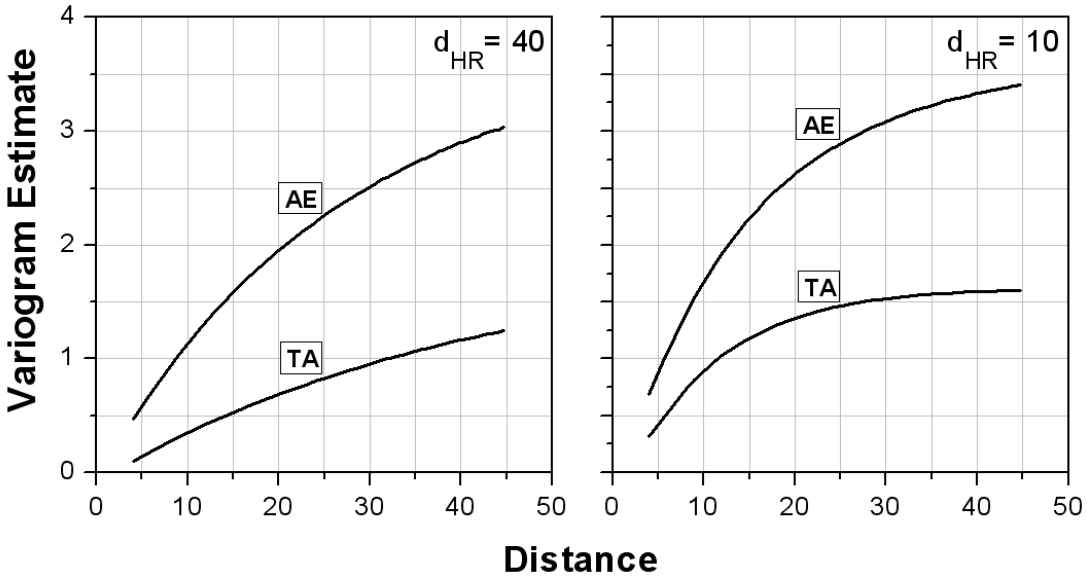

Figure 5.4 True areal (**TA**) and areal with errors (**AE**) variograms of the simulated Gaussian processes. The **AE** variogram is estimated using all the pixels available, and the **TA** variogram is obtained by integrating the variogram of the **HR** field. The distance units and the parameters of the **Err** field are the same as in Figure 5.2.

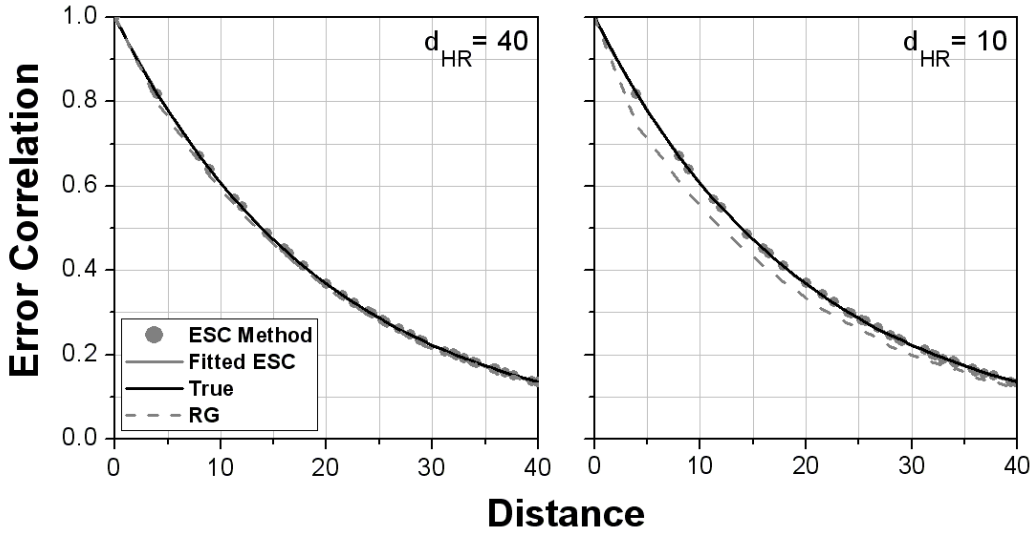

Figure 5.5 Comparison of three functions: 1) the estimated error spatial correlation (ESC) for Gaussian realizations corrected using our method, 2) the true ESC (with a correlation distance of 20) and 3) the spatial correlation structure of radar-gauge (RG) differences. The distance units and the parameters of the **Err** field are the same as in Figure 5.2.

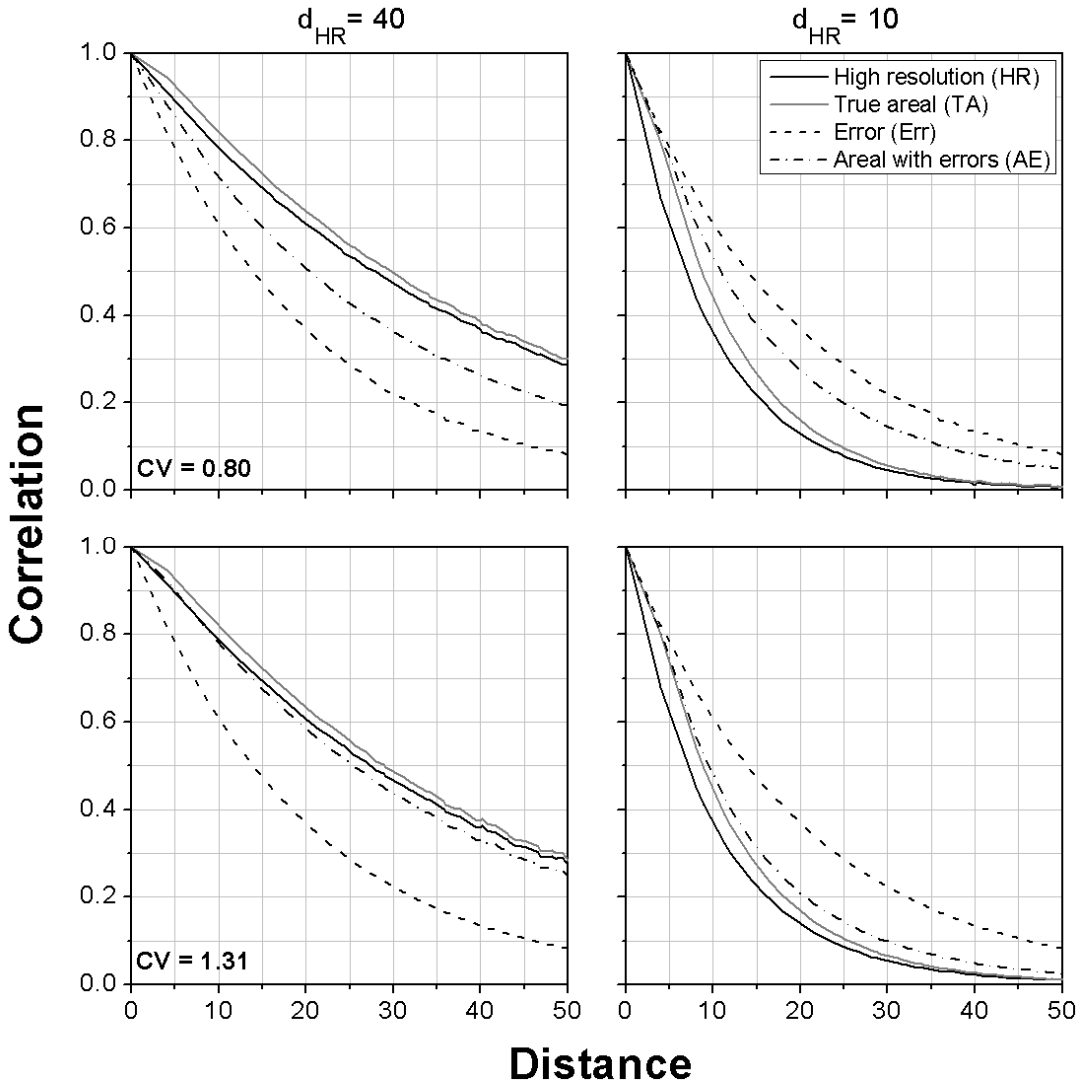

Figure 5.6 Correlation functions of the lognormal **HR** fields simulated with the parameters given in Table 2.1 and estimated using the sampling network shown in Figure 2.2. The distance units and the parameters of the **Err** field are the same as in Figure 5.2.

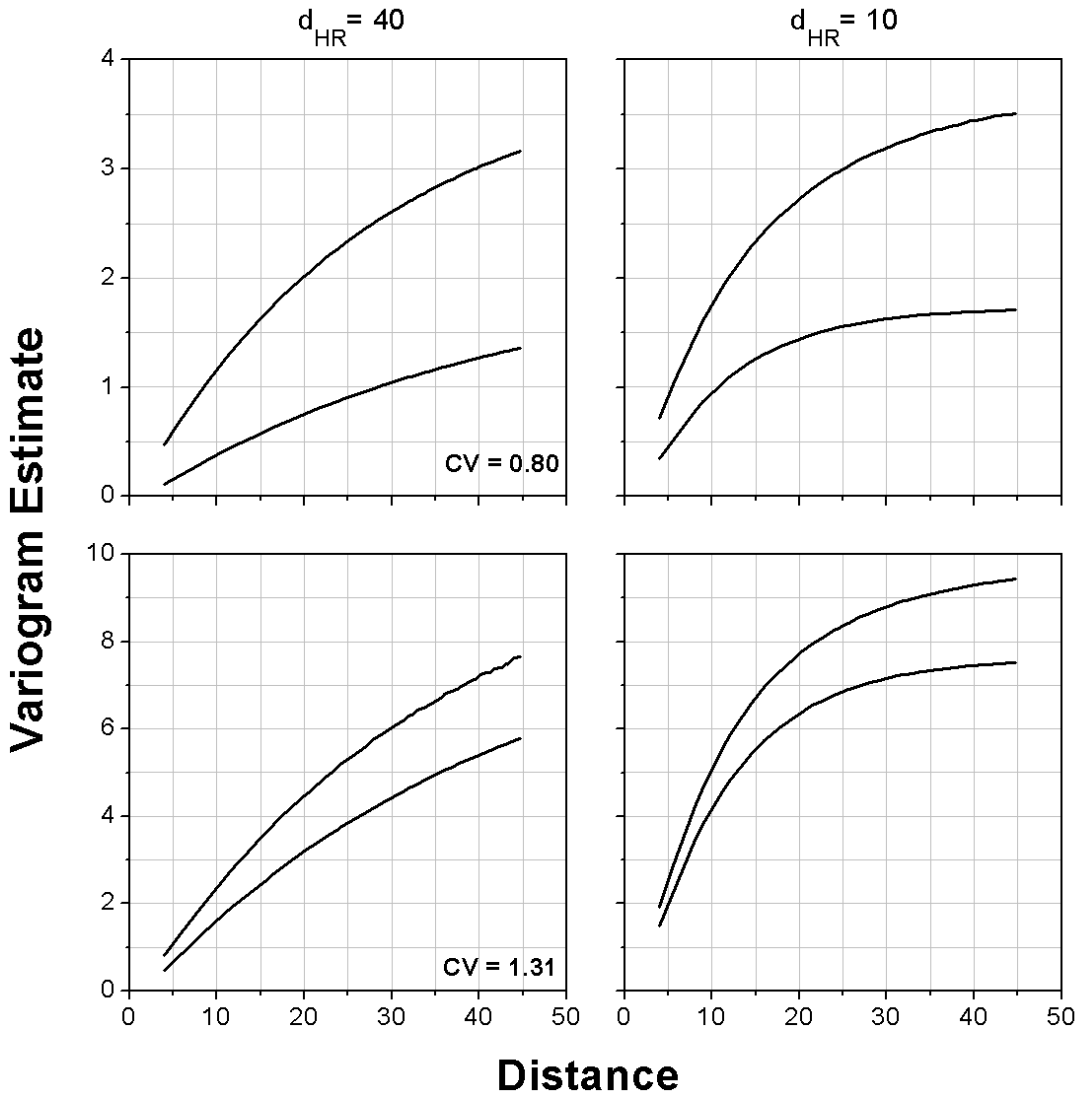

Figure 5.7 True areal (**TA**) and areal with errors (**AE**) variograms of the simulated lognormal processes of Table 2.1. In each panel, upper and lower curves are **AE** and **TA** variograms, respectively. The **AE** variogram is estimated using all the pixels available, and the **TA** variogram is obtained by integrating the point variogram. The distance units and the parameters of the **Err** field are the same as in Figure 5.2.

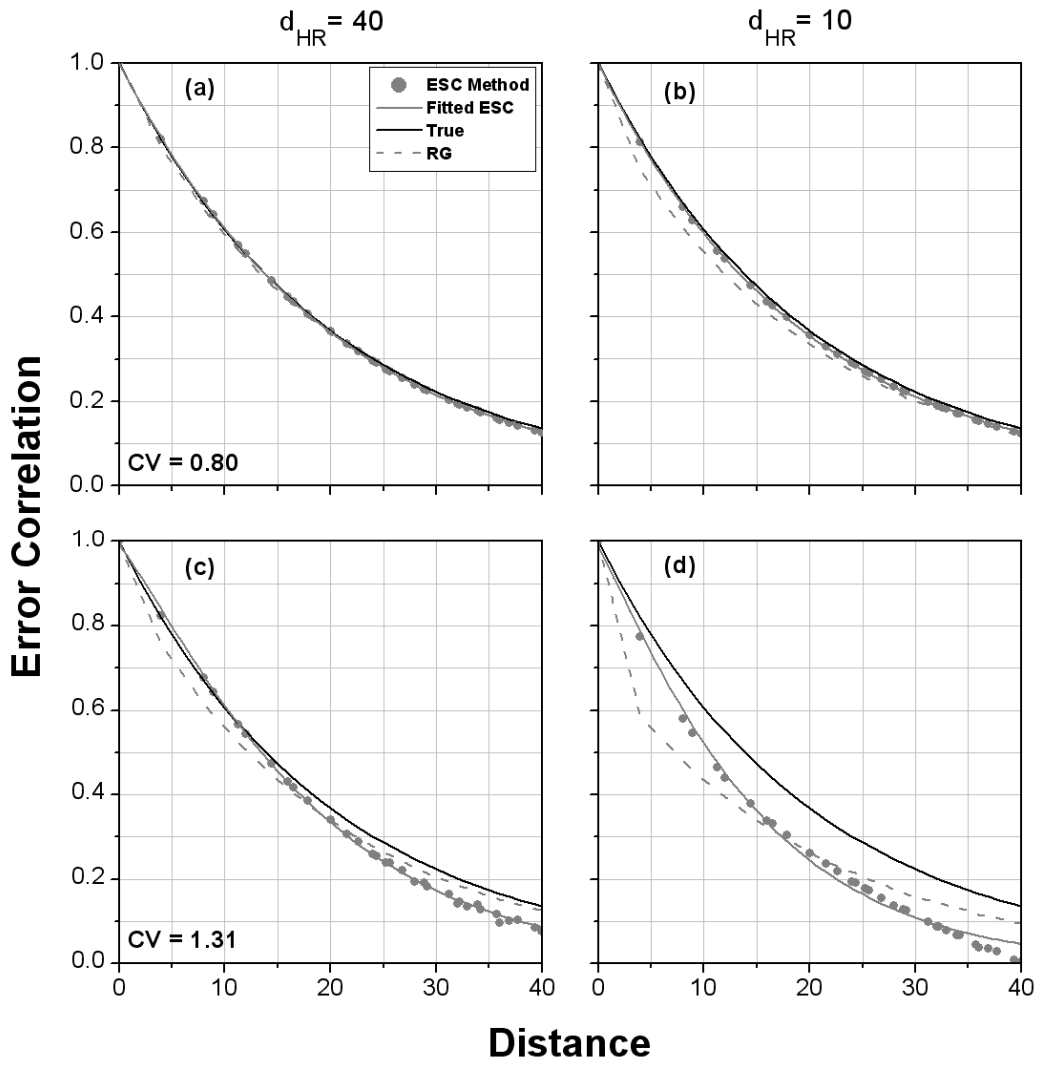

Figure 5.8 Comparison of error spatial correlation (ESC) for lognormal realizations retrieved using our correction method with the true ESC and with the spatial correlation structure of radar-gauge (RG) differences. The distance units and the parameters of the **Err** field are the same as in Figure 5.2.

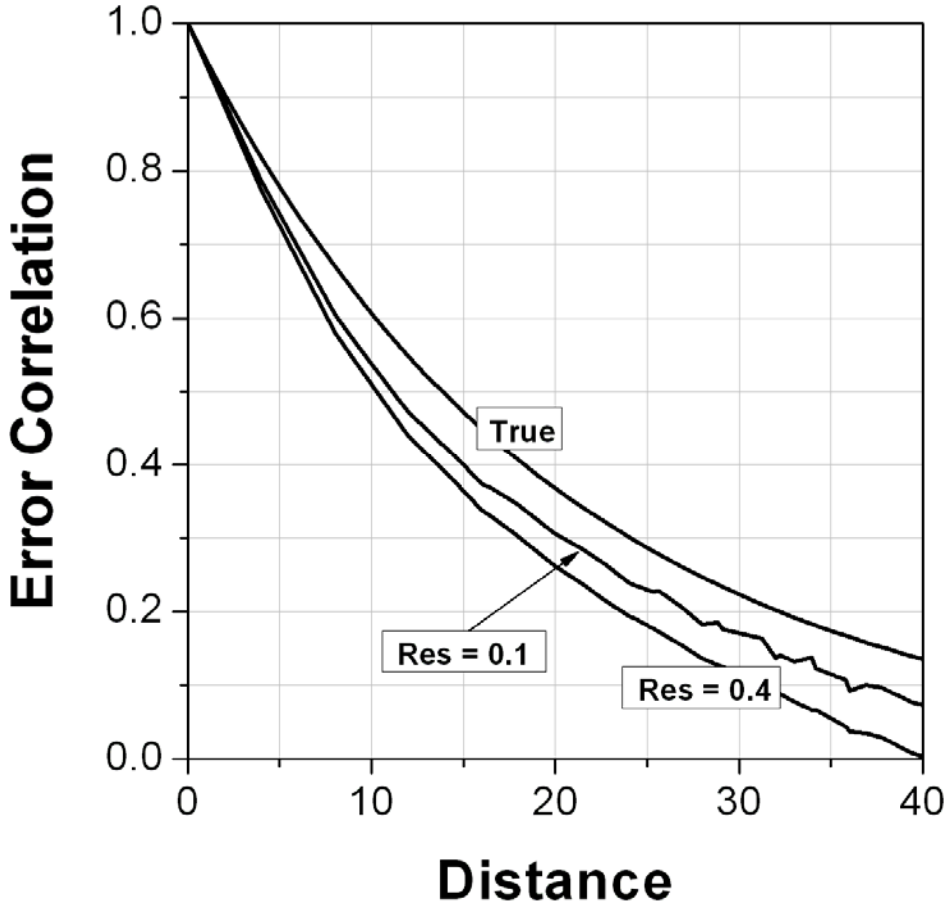

Figure 5.9 Sensitivity of the ESC method to the resolution (Res) of the simulation grid. The distance units and the parameters of the **Err** field are the same as in Figure 5.2.

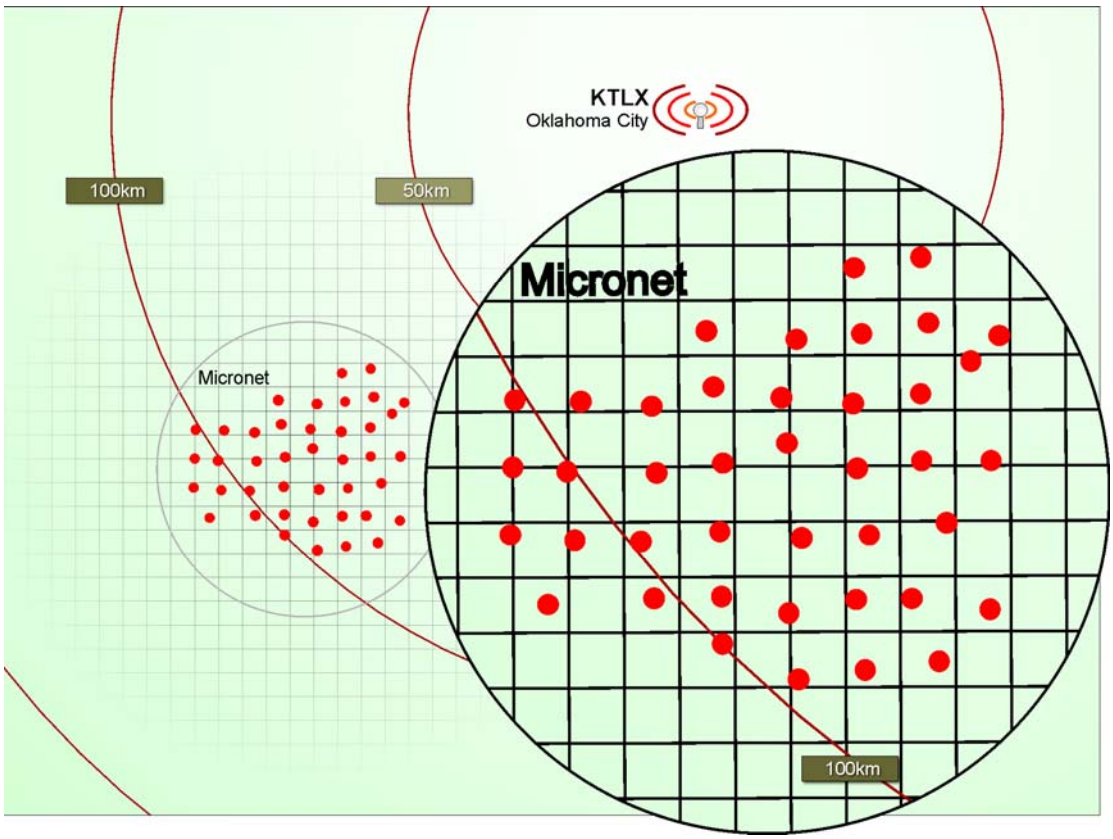

Figure 5.10 Map showing the location of the Oklahoma City NEXRAD (KTLX) radar site, Oklahoma Micronet, and HRAP grid.

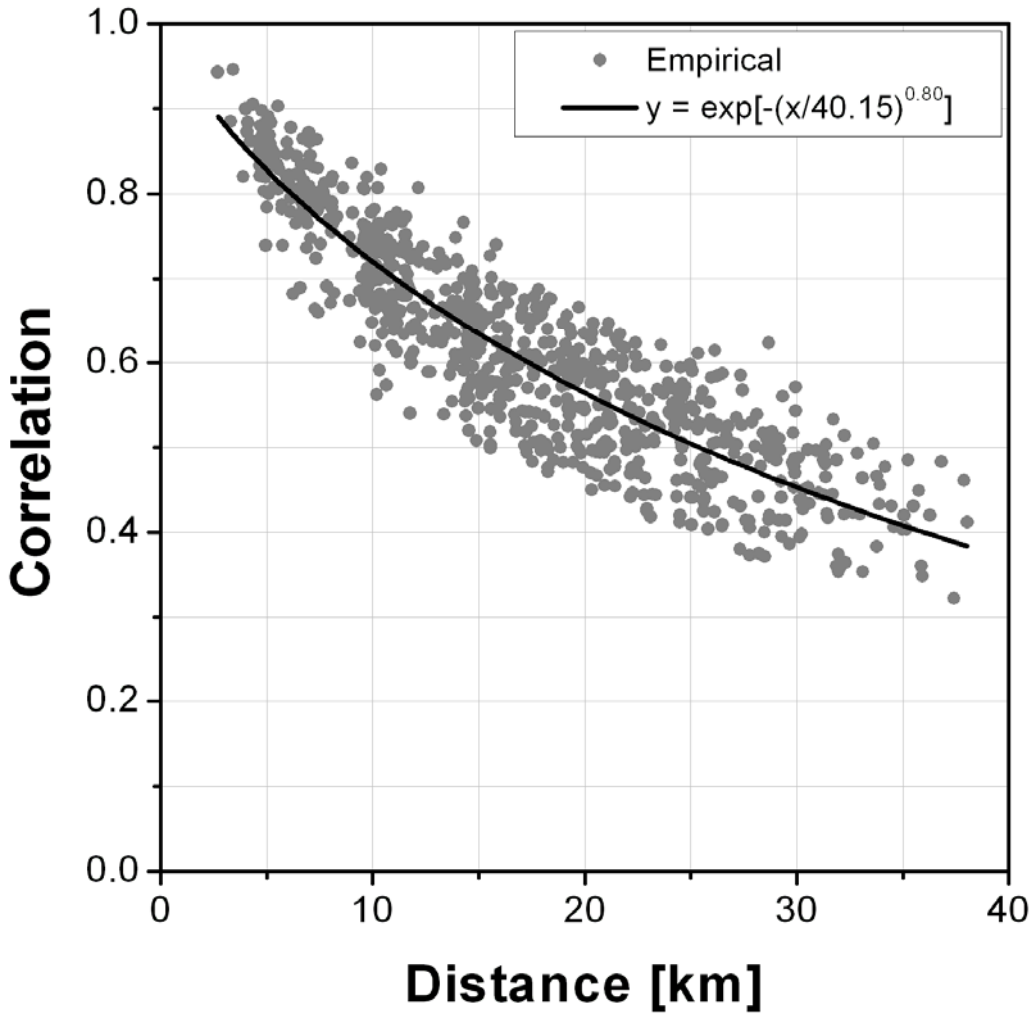

Figure 5.11 Pearson's correlation function of gauge-rainfall obtained using six years of warm season (April to September) rainfall data from Oklahoma Micronet.

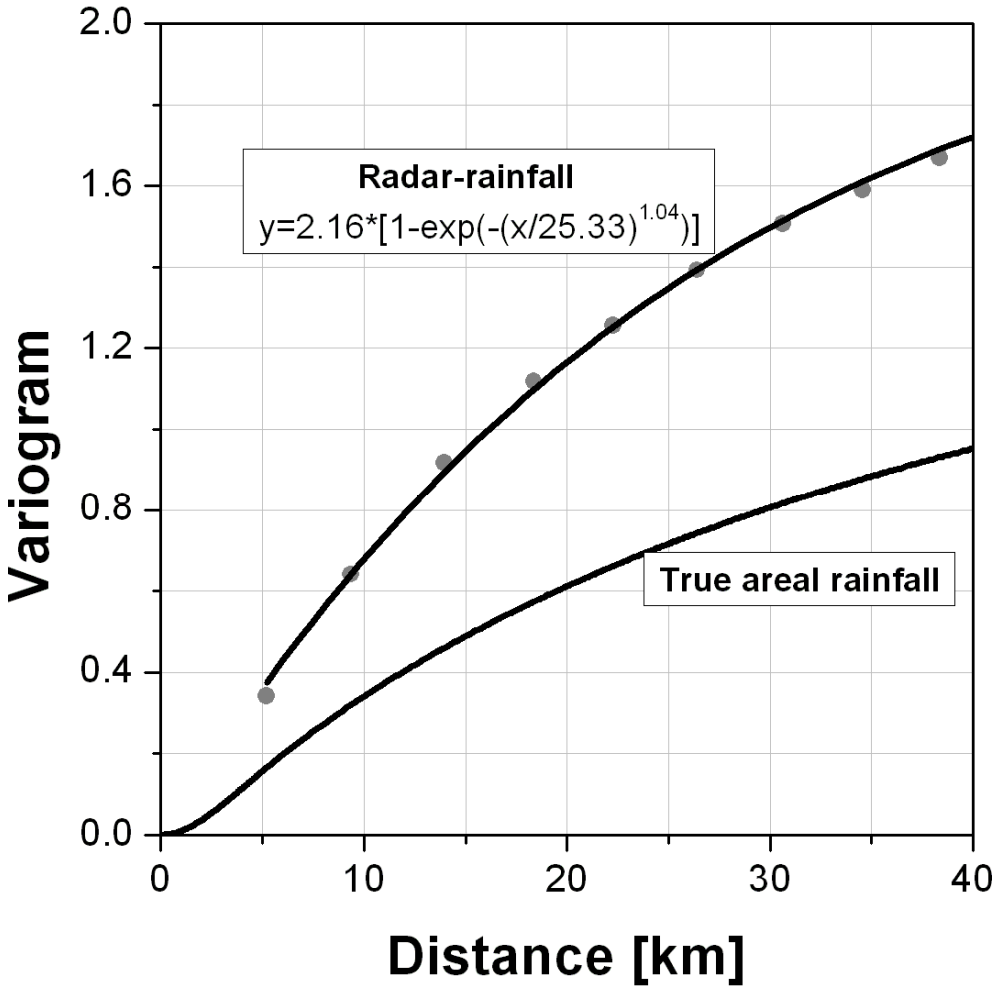

Figure 5.12 Radar-rainfall variogram for the National Weather Service's hourly digital product over the Oklahoma Micronet. The variogram of true areal rainfall is obtained by integrating the gauge-rainfall correlation function over the 4 km pixel.

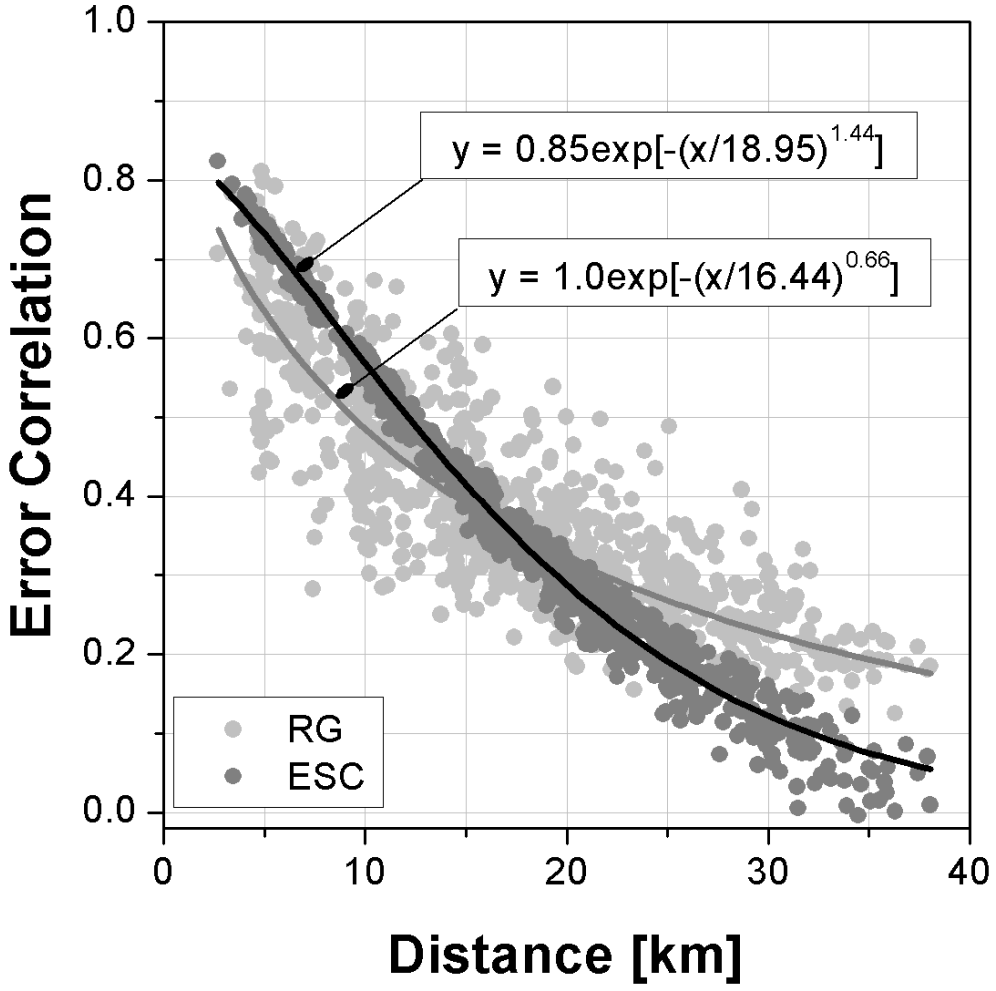

Figure 5.13 Comparison of error spatial correlation of the National Weather Service's hourly digital product with the spatial correlation structure of radar-gauge (RG) differences.

CHAPTER 6

EFFECT OF RADAR-RAINFALL UNCERTAINTIES ON THE SPACE-TIME CHARACERIZATION OF RAINFALL EVENTS

6.1 Introduction

In chapters 2, 3, and 4 we investigated the statistical structure of the peak flows employing simulation scenarios with only the natural variability of the rainfall. However, it is equally important to quantify the effect of rainfall estimation errors [e.g., *Austin*, 1987; *Krajewski and Smith*, 2002; *Ciach et al.*, 2007] on the scaling of peak flows. In this chapter we investigate the effect of radar-rainfall estimation errors on the statistical characterization of rainfall fields. The chapter serves as an intermediate step between characterizing the rainfall estimation error and propagating them through the distributed hydrologic model to investigate their impact on the peak flow scaling structure.

Several studies have characterized the spatial variability of rainfall by employing a variety of techniques ranging from correlation functions and variograms [e.g., *Sumner*, 1982; *Nicholson*, 1986; *Berndtsson*, 1988; *Bacchi and Kottegoda*, 1995; *Ricciardulli and Sardeshmukh*, 2002; *Krajewski et al*., 2003; *Gebremichael and Krajewski*, 2004; *Ciach and Krajewski*, 2006; *Villarini et al.*, 2008] to multiscaling analysis tools such as moment scaling and structure functions [e.g., *Schertzer and Lovejoy*, 1987; *Tessier et al.*, 1993a; *Gupta and Waymire*, 1993; *Menabde et al.*, 1997; *Nykannen and Harris*, 2003; *Lovejoy and Schertzer*, 2006; *Gebremichael et al.,* 2008; *Lovejoy et al.*, 2008; *Morales and Poveda*, 2009; *Mandapaka et al.*, 2009b]. The spatial scales in the aforementioned studies ranged from a few meters to continental scales [e.g., *Lovejoy and Schertzer*, 2006; *Gebremichael et al.*, 2008; *Lovejoy et al.*, 2008; *Mandapaka et al.*, 2009b].

While some early studies employed rain gauge networks to characterize the spatial variability of rainfall, use of remotely sensed data from radars and satellites has increased due to their wide spatial coverage. However, it is well known that rainfall

products based on remotely sensed data contain random and systematic errors from various sources [e.g., *Bell et al.*, 1990; *Bell and Kundu*, 1996; *Smith et al.*, 1996; *Krajewski and Smith*, 2002; *McCollum et al.*, 2002; *Gebremichael et al.*, 2005; *Smith et al.*, 2006; *Ciach et al.,* 2007; *Germann et al.*, 2009; *Mandapaka et al.*, 2009a; *Villarini et al.*, 2009b; *Villarini and Krajewski*, 2009a]. Due to the lack of information on the rainfall estimation error structure, the majority of studies on the spatial characterization of rainfall have not accounted for the errors. Very few works [e.g., *Krajewski et al.*, 1996; *Villarini et al.*, 2007a; *Villarini et al.*, 2007b; *Villarini and Krajewski*, 2009b] have investigated the impact of rainfall estimation errors on the spatial characterization of rainfall.

Krajewski et al. [1996] investigated the impacts of radar beam averaging, the reflectivity-rainfall relationship, and polar-Cartesian grid transformation on the spatial characteristics of radar-rainfall (RR) fields by employing a space-time rainfall model capable of reproducing the observed spatial characteristics to generate an ensemble of rainfall fields. Since these synthetic fields are devoid of any observation errors, they can be considered "error-free rainfall" fields. The radar observation process was simulated by converting rainfall into reflectivity values, imposing Gaussian error fields with zero mean and 1 dBZ standard deviation, converting the reflectivity back to rainfall, and performing a coordinate transformation. The spatial characteristics of the resulting "radar-rainfall" fields were compared with those of "error-free" fields to assess the impact of the radar observation process. They reported that the errors have a significant impact on the spatial characteristics, leading to the underestimation of the coefficient of variation and an overestimation of the spatial correlations. They also showed that the radar observation process results in the underestimation of a random cascade scaling parameter.

Villarini et al. [2007a] analyzed the effect of systematic and random errors on the spatial multifractal properties of rainfall. The systematic effects that they investigated

observed included zero-rain threshold, distance from the radar (range effect), and the parameters of the Z-R equation ($Z = aR^b$, where Z is the radar-reflectivity and R is the rainfall in mm/h) on the scaling behavior of the estimated statistical moments. They reported that the zero-rain threshold and distance from the radar have negligible effects on the estimated scaling functions. However, they also suggested that the range effect could be significant if a larger distance range were investigated in their study. The scaling of statistical moments was most sensitive to the exponent b in the Z-R relation. When the moments against the scale parameter in the log-log domain, they found that the regression line fitted to the moments was steeper for the lower values of the exponent b. To investigate the impact of random errors, they assumed that the RR fields were error-free and convoluted them with uncorrelated and correlated lognormal error fields. They showed that the presence of random errors would lead to the overestimation of the moment scaling functions. *Villarini et al.* [2007b] showed how non-meteorological returns (ground clutter) in RR estimates could affect the estimation of the scaling function.

Although the approach of imposing error fields on the "true" rainfall fields to study the impact of errors is widely used in error propagation studies [e.g., *Sharif et al.*, 2002; *Carpenter and Georgakakos*, 2006; *Vivoni et al.*, 2007], the main limitation in the earlier two studies [*Krajewski et al.*, 1996; *Villarini et al.*, 2007a] was the arbitrary assumptions regarding the statistical structure of the RR error fields.

Villarini and Krajewski [2009b] used a generator of probable true rainfall fields developed by *Villarini et al.* [2009a] (based on the data-driven RR error model in *Ciach et al.* [2007]) to study the impact of rainfall estimation errors on the generalized structure function of rainfall events. For a given RR field, they generated an ensemble of probable rainfall fields and showed that the presence of RR estimation errors results in the overestimation of the structure functions for all of the 15 events considered in their study.

Here we extend the analysis of *Villarini and Krajewski* [2009b] by investigating the effect of errors on the estimates of the spatial structure of rainfall fields. Specifically, we quantify the impact of RR errors on the spatial correlation function, power spectrum, moment scaling function, and the breakdown coefficients. We employed the error model developed by *Ciach et al.* [2007] and the generator developed by *Villarini et al.* [2009a].

This chapter is organized as follows: In Section 6.2, we describe the RR data used in this study. A short description of the RR error model and the probable rainfall generator is presented in Section 6.3. Section 6.4 briefly describes the analysis tools employed to characterize the rainfall events. The results are discussed in Section 6.5, followed by conclusions in Section 6.6.

6.2 Radar-Rainfall Data

Customized high resolution RR data with a spatial resolution of 1×1 km2 and a temporal resolution of 15 minutes were obtained from the Pseudo Precipitation Processing System (PPPS) and Hi-Fi algorithms of the Hydro-NEXRAD system of data distribution [e.g., *Krajewski et al.*, 2009] at The University of Iowa. Krajewski et al. [2009] provide an overview of the Hydro-NEXRAD system and the algorithms used to create the RR products. Radar-rainfall products obtained from the Hydro-NEXRAD system have been used in previous studies published in the literature [e.g., *Ntelekos et al.*, 2008; *Ntelekos et al.*, 2009; *Villarini and Krajewski,* 2009d]. Here, we give a brief description of the PPPS and Hi-Fi algorithms.

PPPS is the Hydro-NEXRAD implementation of the National Weather Service's (NWS) Precipitation Processing System algorithm [e.g., *Fulton et al.*, 1998] that enables us to obtain PPPS RR products at higher spatial and temporal resolutions (1×1 km^2 and 15 min) than the operational products. We refer to it as pseudo-PPS since, despite using the same logic and major algorithmic steps, it does not reproduce exactly the official

NWS products that account for predefined site-specific information such as clutter and terrain occultation maps.

In the Hi-Fi products, corrections are performed to mitigate the errors due to anomalous propagation, range effect, and storm advection. The procedure for constructing the hybrid scan is different for the PPPS and Hi-Fi products. While PPPS takes reflectivity values from the angle that corresponds to 1.0 km above radar altitude, Hi-Fi uses kernel smoothing to alleviate a discontinuity problem in the rainfall maps and to suppress no-rain echoes around the radar [e.g., *Seo et al.*, 2009]. The zero-rainfall threshold for the reflectivity values is also different for the PPPS and Hi-Fi rainfall products. It is equal to 18 dBZ for the former and 10 dBZ for the latter. These factors mentioned above may result in differences in rainfall amounts at the level of final products. In addition to looking at the impact of radar-rainfall uncertainties, we also compare the spatial characteristics of these two RR products.

We selected ten rainfall events measured by Weather Surveillance Radar-88 Doppler (WSR-88D) [e.g., *Crum and Alberty*, 1993; *Klazura and Imy*, 1993] radar in Wichita, Kansas (KICT) and obtained the high resolution rainfall products using the aforementioned algorithms for a square domain of 256×256 km^2 with the radar at its center. Table 1 lists the time of occurrence, duration, and the storm total accumulation for the PPPS and Hi-Fi products of the ten rainfall events. We have selected only warm season (May – August) events to avoid seasonal effects. The events KICT-05 and KICT-06 are shorter duration events lasting for less than a day, while KICT-08 and KICT-09 are longer events lasting for more than 4 days (Table 6.1). The storm total varies from about 7 mm to 90 mm. Table 1 illustrates that the storm accumulation for the products based on two algorithms does not vary greatly.

6.3 Radar-Rainfall Error Model

Ciach et al. [2007] developed a product-error-driven model for RR errors, in which the relation between true areal rainfall and radar-rainfall was described by two components: a systematic distortion function and a stochastic component. Both components were conditioned on the RR values. While the systematic function accounts for biases conditional on the RR values, the stochastic component accounts for the remaining random errors. The results in *Ciach et al.* [2007] were based on a large sample (six years) of hourly accumulation fields (Digital Precipitation Arrays; Fulton et al., 1998) from the Oklahoma City radar, averaged over 4-km pixels, and generated with the Precipitation Processing System (PPS; *Fulton et al.* 1998). Radar-rainfall estimates were complemented with rain gauge measurements, which were used as an approximation of the true ground rainfall. *Ciach et al.* [2007] showed how the systematic distortion function could be approximated by a power law function, while the random component was parameterized by a Gaussian distribution, with mean equal to 1, standard deviation that was a hyperbolic function of rainfall, and with significant correlation both in space and time.

The results in *Ciach et al.* [2007] were then used by *Villarini et al.* [2009a] to develop a generator of probable true rainfall fields conditioned on hourly radar-rainfall maps. As discussed in *Villarini et al.* [2009a], the generator accounts for the conditional and unconditional biases, non-stationarity in variance, and spatial correlation of the random component but not for the temporal dependencies. In this study, we assumed that the parameters of the systematic and random components for the Oklahoma City radar [*Ciach et al.*, 2007] can be used to describe the uncertainties in the RR products generated from the KICT radar. There is no guarantee that the parameters for the Oklahoma City radar are valid for the KICT radar, even though there is evidence that the overall model structure should be valid [*Villarini and Krajewski*, 2009c]. The transferability of the results in *Ciach et al.* [2007] to other radars should be investigated

in future studies. Moreover, the differences in the parameters of the error model obtained using PPS or Pseudo-PPS are not significant [*Villarini and Krajewski*, 2009d].

We performed our analyses for the products aggregated to 4×4 km2 and hourly scales since these are the scales used in *Ciach et al.* [2007]. The time series of hourly accumulations for the ten events are shown in Figure 1.

6.4 Analysis Tools

This section briefly describes the analysis techniques as applied to the selected rainfall events. We start with the description of how the spatial correlation functions were estimated for each rainfall event and then proceed to an estimation of scaling analysis tools such as the power spectrum, the moment scaling function, and the distribution of breakdown coefficients. For all of the analysis tools, we assumed that the RR accumulation fields are temporally independent. We checked this assumption by aggregating individual accumulation fields to various spatial scales and estimating the temporal correlation for each spatial aggregation level. The temporal correlations dropped rapidly, making this assumption reasonable.

6.4.1 Spatial Correlation Function

The correlation is a normalized measure of the linear association between the two random vectors. We used the Pearson's product moment estimator to obtain the correlation of the process $\{Z(\mathrm{u}) : \mathrm{u} \in \mathrm{R}^2\}$ for all u_i and u_j in an $n \times n$ grid such that $\{u_i - u_j = d;\ i,j = 1,2,\ldots n^2\}$

$$\hat{r}(d_t) = \frac{\langle Z(\mathrm{u}_i) Z(\mathrm{u}_j) \rangle - \langle Z(\mathrm{u}_i) \rangle \langle Z(\mathrm{u}_j) \rangle}{\sqrt{\langle Z(\mathrm{u}_i)^2 \rangle - \langle Z(\mathrm{u}_i) \rangle^2} \cdot \sqrt{\langle Z(\mathrm{u}_j)^2 \rangle - \langle Z(\mathrm{u}_j) \rangle^2}} \tag{6.1}$$

where $\langle \cdot \rangle$ represents the expectation operator and $\rho(d)$ is the correlation for a distance lag d. As the accumulation fields were assumed to be independent in time, the averages in equation 1 were estimated by pooling together the pixel values that correspond to the

distance lag *d* in all of the rainfall fields. In this study, *d* varied from 0 to approximately 48 km. The estimated correlations, when plotted against the distance lag, represent the spatial correlation function. If the correlation drops rapidly with distance, then the process is considered highly variable in space. The correlation functions were parameterized by fitting a two-parameter power exponential function (equation 3.1 with $\theta_0 = 1.0$) using Levenberg-Marquardt algorithm. The two parameters are the correlation distance and the shape factor, which are described in detail in Section 3.3.

6.4.2 Power Spectrum

A physical process is said to be scale-invariant or scaling if large scale and small scale structures are related by a scale-changing operation that involves only the scale ratio and an exponent [e.g., *Schertzer and Lovejoy*, 1987]. If different exponents are required to describe the scaling behavior of different moments, then the process is said to be multiscaling. In addition to understanding the rainfall process across multiple scales, other attractive features of the multiscaling framework are that parsimonious models can be developed to generate synthetic rainfall fields at a given resolution and statistical downscaling techniques can be developed to obtain rainfall fields at much higher resolution.

The Fourier power spectrum is one of the most widely used tools to detect the presence of scale-invariance in rainfall. The power spectrum of the rainfall event was obtained as follows:

1) Each rainfall accumulation image is Fourier transformed and the amplitudes are modulus squared to obtain the 2D power spectrum.
2) The two-dimensional power spectrum, which is conjugate symmetric, is folded about the Nyquist frequency.
3) Assuming isotropy, the folded spectrum is radially averaged about the corner.

4) The individual spectra estimated at each time instant were then averaged to obtain the average spatial spectrum.

A process is said to be scale-invariant if the power spectrum displays log-log linearity (power law) within a finite range of frequencies. If *E(f)* is the power for the frequency *f*, then the scale-invariant field will have a power spectrum of the form:

$$E(f) = f^{-\beta} \tag{6.2}$$

where β is the negative slope of the spectrum in the log-log domain. The power spectrum slope is an indicator of the spatial organization of the field. The higher the value of β, the smoother and more organized the rainfall field [e.g., *Purdy et al.*, 2001; *Nykanen and Harris*, 2003]. It should be noted that for the fields with a fixed resolution, the power spectrum is distributed in uniform frequency bins. When such spectrum is plotted in a double logarithmic plot, most of the spectrum is concentrated towards higher frequencies. To avoid excess weighting on the higher frequencies, we estimated β by performing ordinary least squares regression on the octave binned power spectrum [e.g., *Harris et al.*, 1997] in the double-logarithmic domain. The log-log linearity was checked based on the R^2 value in linear regression.

6.4.3 Spatial Moment Scaling Analysis

The next tool we used in this study is the moment scaling analysis to investigate the presence of multifractality. For a multifractal process, it has been shown [e.g., *Menabde et al.*, 1997] that the spectral slope β is always less than the dimension (D) of the field. If $\beta > D$, which is often the case with geophysical phenomena including rainfall, moment scaling analysis has to be performed on either the fractionally differentiated field [e.g., *Schertzer and Lovejoy*, 1987; *Nykanen and Harris*, 2003] or on the small-scale fluctuations (gradient) of the original field [e.g., *Tessier et al., 1993a*; *Menabde et al.*, 1997].

In this study, we adopted the latter approach and obtained small-scale fluctuations as the difference between the value at the given pixel and the mean of the four nearest neighbors (three for edge pixels and two for the ones on the corner). For each rainfall event, the gradient fields of resolution r (= 4 km) and size L (=256 km) were then averaged over a range of scales $l(n)$. The value of n varies from 0 to 6, with $l(0)$ equal to 4 km and $l(6)$ equal to 256 km. The averaged fields are referred to as φ_λ, where λ is the ratio of the size of the field L to the averaging scale $l(n)$. Assuming that the fluctuation fields within the rainfall event were time-independent, the average values in the φ_λ fields at all of the time steps were pooled together, and the trace moments $M_q(\lambda)$ of various moment orders q were estimated for each scale ratio λ. The higher the value of λ, the larger the sample size available to estimate the moments.

The gradient field is multifractal if there is a scaling relationship of the form

$$M_q(\lambda) = \left\langle \varphi_\lambda(x,t)^q \right\rangle \sim \lambda^{-K(q)} \tag{6.3}$$

where, $K(q)$ (the slope of $M_q(\lambda)$ versus λ in the log-log domain) is a nonlinear function of the moment order q. Theoretically, $K(q)$ is required for q values ranging from 0 to ∞ to fully characterize the multifractality in the rainfall fluctuation fields. However, *Tessier et al.* [1993a] proposed a universal multifractal (UM) model for $K(q)$ based on multiplicative cascades consisting of parameters α and C_1.

$$K(q) = \frac{C_1}{\alpha - 1}\left(q^\alpha - q\right) \qquad 0 \le \alpha < 1 \quad \text{and} \quad 1 < \alpha \le 2 \tag{6.4}$$

$$K(q) = C_1 q \log q \qquad \alpha = 1 \tag{6.5}$$

The parameter α is the Levy-stable (or multifractality) index that characterizes the spikiness and indicates the probability distribution from which the weights are generated in the cascading process. The case $0 < \alpha < 2$ ( ) corresponds to log(Levy) multifractals, and if $\alpha = 1$ the multifractal process is log(Cauchy) [e.g., Tessier et al., 1993a]. The case with $\alpha = 2$ corresponds to lognormal multifractals. C_1 is the intermittency parameter that

characterizes the sparseness of the mean. It should be noted that the above model (equations 6.5 and 6.6) is for the multiscaling conservative cascades for which $\beta < D$. However, rainfall fields often display a non-conservative, multi-affine nature with a power spectrum slope greater than the dimension of the process. The degree of non-conservation is quantified in terms of the Hurst exponent H, estimated as

$$H = \frac{[\beta - D + k(2)]}{2} \tag{6.6}$$

where $K(2)$ is the scaling exponent corresponding to the second moment order. The Hurst exponent is also an indicator of the smoothness of the field [e.g., *Harris et al.*, 2003].

In this study, the parameters C_1 and α were obtained using the double trace moments (DTM) technique [e.g., *Tessier et al.*, 1993a]. In the DTM technique, we take various powers η of the rainfall fluctuations at their highest resolution, average the powered fluctuations to various spatial scales with scale ratio λ and estimate the statistical moments (referred to as double trace moments) of various orders q.

$$M_{\eta,q}(\lambda) = \left\langle \left(\varphi^{\eta}(x,y)\right)_{\lambda}^{q} \right\rangle \tag{6.7}$$

In the case of universality, the double trace moments will depend on the scale ratio as [e.g., *Tessier et al.,* 1993a]:

$$M_{\eta,q}(\lambda) \sim \lambda^{K(q,\eta)} \tag{6.8}$$

and

$$K(q,\eta) = \eta^{\alpha} \cdot K(q) \tag{6.9}$$

Therefore, α is the slope of $K(q,\eta)$ versus η in a double logarithmic plot for a fixed q. The value of C_1 can be obtained by plugging α in equation 6.5 for a fixed q. For a detailed description of the DTM technique, the reader is referred to *Tessier et al.* [1993a]. Several studies have applied the UM model over the last 15 years to describe a variety of

geophysical phenomena such as wind and atmospheric temperature [e.g., *Schmitt et al.*, 1994; *Lazarev* et al., 1994; *Schmitt et al.*, 1996], rain and clouds [*Tessier et al.*, 1993a; *Hubert et al.*, 1993; *Naud et al.*, 1996; *Marsan et al.*, 1996], ocean surface [*Tessier et al.*, 1993b], hydraulic conductivity [e.g., *Liu and Molz*, 1997], and topography [e.g., *Lavallee et al.*, 1993; *Tchiguirinskaia et al.*, 2000; *Gagnon et al.*, 2006].

6.4.4 Breakdown Coefficients

There are many ways to perform the multiscaling analysis. One such approach - the moment scaling analysis – was discussed in the previous subsection. Another technique considered in this study is the behavior of the breakdown coefficients, which was originally applied in the field of hydrodynamical turbulence (*Novikov* [1990] and references therein). Unlike other analysis methods used in this study, the theory of breakdown coefficients has not been extensively employed in the rainfall literature. *Menabde et al.* [1997] demonstrated the self-similarity of the logarithm of breakdown coefficients of rainfall fluctuations and proposed a self-similar random cascade model for rainfall. *Harris et al.* [1998] proposed parameter estimation techniques for the distribution of breakdown coefficients. They also demonstrated that the scaling behavior of the breakdown coefficients and the trace moments in the moment scaling analysis were quite close for large sample sizes. In this study, we followed the methodology of *Menabde et al.* [1997] and *Harris et al.* [1998] to estimate the breakdown coefficients.

The rainfall gradient fields were integrated at various scales *l(n)*, and the logarithm of breakdown coefficients were estimated as follows.

$$X_n = -\ln\left(\frac{\psi_{l(n)}}{\psi_{l(n+1)}}\right) \tag{6.10}$$

where, $\psi_{l(n)}$ is the pixel value for the integrated field at scale *l(n)*, X_n is the breakdown coefficient at scale *l(n)*, and the value of *n* ranges from 0 to 3. Larger scales were not considered as there were insufficient data to obtain the probability distribution. The pixel

value at scale *l(n+1)* will always be greater than the four individual pixels at the scale *l(n)*. Therefore, the breakdown coefficients will always be non-negative. As the rainfall fields were assumed to be independent in time, the breakdown coefficients were pooled together for all the fields within the rainfall event, and the probability distribution was obtained. We employed breakdown coefficients only to demonstrate the statistical self-similarity of the rainfall gradient fields. We did not attempt to link the characteristics of the probability distribution (such as shape and peak) to the spatial structure (such as smoothness).

6.5 Results and Discussion

We obtained each of the aforementioned spatial characterization estimates for the PPPS and Hi-Fi radar-rainfall products as well as for the rainfall events generated using the rainfall generator (Section 6.3) and conditioned on the PPPS storm data (for each event, we generated an ensemble of 200 probable true events). First, we compared the estimates for both products to see if corrections involved during the generation of Hi-Fi products altered the estimated spatial structure. Then, we compared the estimates for the PPPS product with those of probable rainfall events to assess the impact that the rainfall estimation errors in the PPPS product had on the spatial characterization of rainfall events. We obtained each of the aforementioned spatial characterization estimates for the PPPS and Hi-Fi radar-rainfall products as well as for the rainfall events generated using the rainfall generator (Section 6.3) and conditioned on the PPPS storm data (for each event, we generated an ensemble of 200 probable true events). First, we compared the estimates for both products to see if corrections involved during the generation of Hi-Fi products altered the estimated spatial structure. Then, we compared the estimates for the PPPS product with those of probable rainfall events to assess the impact that the rainfall estimation errors in the PPPS product had on the spatial characterization of rainfall events.

6.5.1 Effect on Spatial Correlation Function

The spatial correlations were estimated for the radar-rainfall data using Pearson's product moment estimator (equation 6.1). However, we found that correlation estimates for both algorithms were quite close, meaning that the spatial correlation structure was not sensitive to the corrections performed during the creation of Hi-Fi rainfall products. The dark solid line in Figure 6.2 represents the correlation estimates plotted against the distance lag for the rainfall products obtained using the Pseudo-PPS algorithm. For the sake of clarity, we did not plot the corresponding estimates for the Hi-Fi product. The correlation estimates were fitted using the two-parameter exponential function (equation 6.2), and the parameters (correlation distance and shape parameter) are given in Table 6.2 for both radar-rainfall algorithms. The correlation distance varied from a minimum value of around 18 km for the storm KICT-09 to a maximum value of about 48 km for the storm KICT-10 (Table 6.2). The shape parameter ranged between 0.74 and 1.20. If the shape parameter is smaller than 1.0, the decay in correlations is faster than exponential, while if it is greater than 1.0, the small-scale correlation decays at a much slower rate, leading to smoother fields. From the values of correlation distances and shape parameters, we can infer that the storms KICT-03, KICT-06, and KICT-09 are more variable than the others. Similarly, the storms KICT-04, KICT-05, and KICT-10 are smoother than the other events considered.

The gray lines in Figure 6.2 represent the spatial correlation functions estimated from each of the 200 probable rainfall events. The effect of RR errors varies with each storm. The bias in the estimation of the spatial correlation function was clearly evident for storms KICT-05 and KICT-10, whereas the bias was not noticeable for the storms KICT-06 and KICT-09 (Figure 6.2). The mean, 5th, and 95th percentiles for the parameters of the exponential correlation function (equation 3.1) fitted to the gray lines in Figure 6.2 are listed in Table 2. The correlation distance for the PPPS product is greater than the average correlation distance from the ensemble of probable rainfall events for all

the storms (Table 6.2). Nonetheless, the correlation distance always lies within the 90% confidence interval obtained from the probable rainfall events (Table 6.2), indicating that there is not enough statistical evidence to say that they are different at the 10% significance level. The impact of RR errors was more pronounced for the shape parameter. The estimates of the shape parameter for the PPPS product were always greater than the average shape parameter from the probable rainfall fields (Table 6.2). Moreover, they are always larger than the 95th percentile. Therefore, the presence of errors induces spurious correlations in the RR fields at smaller scales.

6.5.2 Effect on Power Spectrum

Although the spatial correlation function provides information on the variability of the rainfall fields, it does not tell us how the small scale and large scale structures in rainfall interact with each other. This information can be obtained by estimating the power spectrum. We estimated the power spectrum for all of the ten rainfall events using the algorithm discussed in Section 6.4.2. Similar to the correlation estimates, the power spectrum was also unaffected by the corrections performed in the Hi-Fi rainfall products. Figure 6.3 (dark lines) shows the power spectrum for the rainfall products generated using the PPPS algorithm. Except for the slight departure towards the lower frequencies (for example, event KICT-02 in Figure 6.3), caused mainly by sampling effects, the power spectrum displayed log-log linearity for most of the frequency domain. The negative slope β was then estimated for RR events using the octave binning technique discussed in Section 6.4.2. Since linear regression was performed on the octave binned spectrum, the sampling issue does not significantly affect the estimation of the slope β. The R^2 value in the linear regression was always greater than 0.97, confirming log-log linearity of the spectrum. The spectral exponents for all of the events are listed in Table 6.3. The values of β ranged from 2.05 to 2.88, with lower values for the storms KICT-06 and KICT-09 and higher values for KICT-04 and KICT-10. The highest value of β for

KICT-10 indicates a smoother and more organized rainfall event, which is in agreement with the results from the correlation analysis that reported high values of correlation distance and shape parameter for that particular event.

The power spectra estimated for an ensemble of probable rainfall fields (gray lines in Figure 6.3) also displayed log-log linearity for most of the frequency domain. Comparing the spectrum obtained from PPPS RR products with those of the probable rainfall events (Figure 6.3), the errors have a larger impact towards higher frequencies (small scales). The effect of errors is to smooth the rainfall fields at small scales, thereby decreasing the contribution of higher frequencies and increasing the value of β. The β value for the PPPS RR product was greater than the average β from the probable rainfall events for all of the storms (Table 6.3). Except for the storms KICT-06 and KICT-09, the β value falls outside the 90% confidence interval obtained for the probable rainfall events (Table 6.3). The events KICT-06 and KICT-09 also happen to be the most variable of all the events considered in the study (Tables 6.2 and 6.3). Even in the correlation analysis, the bias in shape parameter is smaller for storms KICT-06 and KICT-09. With the exception of these two storms, the RR errors result in the overestimation of the power spectrum slope (the value is larger than the 95th percentile). The small scale smoothing effect of the RR errors is consistent with the results from the correlation analysis. In Figure 6.4, we compared the results from correlation and power spectrum analysis for the events KICT-06 and KICT-08. KICT-06 is shown because it is one of the shortest duration events recording the smallest rainfall accumulation, whereas KICT-08 was the longest event recording one of the highest rainfall accumulations. The results are similar for other events.

6.5.3 Effect on Moment Scaling Function

The value of β from the power spectrum analysis was always greater than the dimension of the field for all the rainfall events. Therefore, as mentioned in Section

6.4.3, moment scaling analysis has to be performed on the gradient of rainfall fields. We obtained the gradient fields following the methodology proposed by *Menabde et al.* [1997], which was briefly described in Section 6.4.3. The gradient fields were averaged to various spatial scales, and the trace moments (equation 6.3) were estimated for moment orders ranging from 0.1 to 4.0. The trace moments of various orders estimated for all the rainfall events are shown in Figure 6.5. The slopes of the moments were obtained by performing ordinary least squares linear regression. However, to avoid sampling effects, the trace moment value corresponding to the smallest scale ($\lambda = 1$) was not included in the regression analysis. The R^2 value in the regression was always greater than 0.98, confirming the log-log linearity of moments. For all the rainfall events, the slopes varied with the moment order, indicating that the rainfall gradients were not simple scaling (Figure 6.5). The dark lines in Figure 6.6 represent the moment scaling or *K(q)* functions (slopes against the moment order) for the PPPS RR product. From the nonlinearity of the *K(q)* function (Figure 6.6), we can infer that the rainfall gradient fields were multiscaling.

We parameterized the *K(q)* functions with the universal multifractal model (equations 6.4 and 6.5) using the double trace moment technique described in Section 6.4.3. In Figure 6.7, the DTM analysis was shown for the PPPS products of the rainfall events KICT-06 and KICT-08. The top panels of Figure 6.7 show the double trace moments (estimated using equation 6.7) plotted against the scale ratio λ for different moment orders η and for a fixed q value of 1.6. The double trace moments displayed log-log linearity with the scale ratio λ for various η values (top panels of Figure 6.7). We obtained the slopes *K(q,η)* in equation 6.8 by performing ordinary least squares regression, excluding the moment value corresponding to the lowest value of λ. The slope of *K(q,η)* plotted against η in the double logarithmic plot (bottom panels of Figure 6.7) gives the value of the multifractality index α. The intermittency parameter C_1 was obtained by using equations 5 and 10 for a fixed q of 1.6. In this study, we carried out

the DTM analysis for different q (equal to 1.2, 1.6, 2.0, and 2.4) and estimated the average UM model parameters α and C_1.

Figure 6.8 shows the moment scaling exponents (from Figure 6.6) and the DTM-fitted moment scaling functions for the PPPS RR products of all ten rainfall events. The average parameters for all ten rainfall events and for both the algorithms (PPPS and Hi-Fi) are listed in Table 6.4. Similar to the correlation and power spectrum analysis, the UM model parameters were not significantly different for the PPPS and Hi-Fi products (Table 6.4). The intermittency parameter for the PPPS product ranged from 0.21 for the event KICT-10 to 0.40 for KICT-09, whereas the multifractality index varied from 0.99 for KICT-01 to 1.39 for KICT-08 (Table 6.4). From the low value of C_1 and the relatively high value of α for the storm KICT-10 (Table 6.4), we can say that the storm is more space-filling and has relatively fewer spikes than the other storms; the result is consistent with correlation and power spectrum analysis. Similarly, the high value of C_1 and relatively low value of α for the KICT-09 suggest that the storm is spikier and less space-filling than other storms.

We repeated the trace moment estimation and obtained the slopes for all of the 200 probable rainfall events (gray lines in Figure 6.6). The effect of RR errors varied for each storm, with noticeable bias for the storms KICT-01, KICT-08, and KICT-10 and small bias for the storms KICT-02, KICT-04, and KICT-07 (Figure 6.6). We performed the DTM analysis for all 200 probable rainfall fields and compared the UM model parameters with those of the PPPS product to assess the impact of RR errors (Table 6.4). In general (except C_1 for KICT-02 and α for KICT-07), the UM model parameters for the PPPS product are greater than the corresponding average values estimated from probable rainfall events (Table 6.4). The higher value of C_1 for the PPPS product meant an increase in the intermittency or a decrease in the smoothness of the field, which is contradictory to the result obtained from the correlation and power spectrum analysis. However, it can be explained by focusing on the behavior of α. Table 6.4 illustrates that

the value of α for the PPPS product is greater than the average α from probable rainfall events, which means fewer spikes and smoother fields. The net effect is the smoothing of the field and is therefore consistent with the correlation and power spectrum analysis.

The explanation provided in the previous paragraph must be taken with caution as C_1 and α for the PPPS product lie within the 90% confidence interval for some of the selected storms. For example, although C_1 and α for storms KICT-03, and KICT-06 are greater than corresponding average values from probable rainfall events, they lie within the 5th and 95th percentile of the distribution obtained from the probable rainfall events (Table 6.4). At the 10% significance level, there is not enough evidence that the UM parameters are overestimated for KICT-03 and KICT-06 We then estimated the Hurst exponent H using equation 6.6 for the RR products as well as the probable rainfall events (Table 6.5). The Hurst exponent for the PPPS products is larger than the 95th percentile for all of the storms. As the Hurst exponent characterizes the degree of smoothness, the effect of RR errors is to increase the smoothness of the rainfall fields. Since the power spectrum slope is used in the computation of the Hurst exponent, any impact of errors on β will propagate into the estimation of the Hurst exponent. The histograms of the UM model parameters and the Hurst exponent for events KICT-06 and KICT-08 are shown in Figure 6.9.

6.5.4 Effect on Breakdown Coefficients

As mentioned in Section 6.4.4, we employed breakdown coefficients mainly to demonstrate the self-similarity of the gradient of the rainfall fields. We estimated the breakdown coefficients using equation 6.10 for all ten rainfall events. Figure 6.10 shows the probability distribution of the negative logarithm of breakdown coefficients at four successive scales as well as the average probability distribution for each rainfall event. From Figure 6.10, we can infer with reasonable confidence that the breakdown coefficients at successive scales are from the same probability distribution, confirming

that the rainfall gradient fields are self-similar. We refrained from fitting the probability distribution of breakdown coefficients with a parametric form as it is not a trivial task to generate four random weights at each cascade step from a particular function keeping their sum equal to one [e.g., *Menabde et al.*, 1997]. *Menabde et al.* [1997] and *Harris et al.* [1998] proposed iterative techniques to obtain the weights whose distribution follows the empirical distribution of breakdown coefficients. However, we did not employ those techniques in this study and limit ourselves to a preliminary assessment of the impact of the errors.

We then estimated the breakdown coefficients for the 200 probable rainfall events. Figure 6.10 shows the distribution of breakdown coefficients for the probable rainfall events along with that of the PPPS RR product. Unlike other analysis techniques used in this study, the effect of RR errors on the distribution of breakdown coefficients is not consistent for all rainfall events. For example, RR errors led to underestimation of the peak of the distribution for KICT-07 and KICT-10 and overestimation for the events KICT-08 and KICT-09. This inconsistent behavior could be attributed to the storm-to-storm variability. However, more research is needed to relate the biases to the specific characteristics of the storms.

6.6 Summary and Conclusions

Remotely sensed rainfall products, which are widely used for the spatial and temporal characterization of rainfall events, are affected by errors from various sources. Therefore, it is imperative to quantify the effect of errors on the estimated rainfall characteristics. In this study, we investigated the impact of radar-rainfall (RR) estimation errors on the spatial characterization of RR products for warm season rainfall events over Wichita, Kansas. For each storm, we generated 200 probable rainfall events using the rainfall generator developed by *Villarini et al.* [2009a] that was based on a recent RR error model of *Ciach et al.* [2007]. We assessed the effect of errors on different spatial

characterization tools such as the spatial correlation function, power spectrum, moment scaling function, and the breakdown coefficients. Besides performing qualitative analysis, we also quantified the impact of errors by parameterizing the aforementioned functions with models widely used in the rainfall literature.

From the spatial correlation analysis, we found that the shape parameter, which characterizes the behavior of the spatial correlation function at the small separation lags, was significantly overestimated for the RR products of all of the storms. In general, there was a tendency for the correlation distance (defined as the distance at which the correlation drops to 1/e) to be also overestimated for the RR products. The power spectrum analysis revealed that the presence of RR errors will smooth the fields at higher frequencies, leading to the overestimation of the negative power spectrum slope for eight out of ten rainfall events. The correlation and power spectrum analysis also showed that the effect of RR errors was the smallest for the more variable events. Moment scaling analysis was then carried out to see the effect of errors on the scaling of statistical moments of various orders. The moment scaling functions were fitted with a universal multifractal model using the double trace moment (DTM) technique. The results from the DTM analysis also suggested smoothing of rainfall fields by the RR errors. However, the results were not as conclusive as the correlation and power spectrum analysis for some of the rainfall events. The Hurst exponent, which characterizes the degree of non-conservation or smoothness in the rainfall fields, was also significantly overestimated for all of the storm events. We then estimated the breakdown coefficients for all of the storms as well as the probable rainfall events. We used breakdown coefficient distributions only to demonstrate the self-similarity of the rainfall fields. Radar-rainfall errors altered the shape of the breakdown coefficient distributions. However, we did not attempt to relate the bias to the synoptic characteristics of the precipitation systems.

Overall, the results from our study indicate that RR errors induce spurious correlations in the RR products at the smaller scales. The results are in agreement with

previous studies, which utilized arbitrary error models. Although the impact is significant for characterizing the spatial structure of RR products, the real impact of the errors will depend on where these products are applied. In particular, propagation of these errors through hydrologic models would provide significant insight into the impact of radar-rainfall uncertainties on streamflow predictions.

As mentioned above, we have focused on the spatial characterization of rainfall events. The investigation of the effect of radar-rainfall errors on the space-time characteristics of rainfall would require including the temporal dependencies of the random errors into the generator developed by *Villarini et al.* [2009a], which is beyond the scope of this study.

Table 6.1* List of the selected rainfall events indicating the beginning, end, duration, and the overall accumulation.

Storm ID	Begin mm/dd/yyyy	End mm/dd/yyyy	N [hr]	Accumulation [mm]	
				PPPS	Hi-Fi
KICT-01	06/30/2004 15:15	07/03/2004 03:00	60	32.58	30.92
KICT-02	07/26/2005 02:15	07/27/2005 10:00	32	7.97	7.13
KICT-03	08/12/2005 15:15	08/14/2005 12:00	45	36.89	36.78
KICT-04	08/24/2005 02:15	08/25/2005 21:00	43	58.36	53.95
KICT-05	06/21/2006 20:15	06/22/2006 16:00	20	22.45	21.31
KICT-06	07/26/2006 20:15	07/27/2006 19:00	23	11.31	11.30
KICT-07	05/05/2007 20:15	05/07/2007 11:00	39	94.32	86.18
KICT-08	06/26/2007 12:15	07/01/2007 00:00	108	84.37	80.02
KICT-09	07/27/2007 15:15	07/31/2007 17:00	98	26.68	26.10
KICT-10	05/07/2008 00:15	05/08/2008 03:00	27	45.11	42.42

*The accumulation value is shown for the rainfall events obtained from Pseudo-PPS (PPPS) and Hi-Fi algorithms. The size of the domain is 256 × 256 km^2 with the KICT radar at its center. The spatial resolution of the data is 4 km.

Table 6.2* Parameters of spatial correlation function estimated from the radar-rainfall data obtained from Pseudo-PPS (PPPS) and Hi-Fi algorithms.

Storm ID	Correlation Distance [km]					Shape Factor				
	PPPS	Hi-Fi	Probable Rainfall			PPPS	Hi-Fi	Probable Rainfall		
			$Q_{0.05}$	Mean	$Q_{0.95}$			$Q_{0.05}$	Mean	$Q_{0.95}$
KICT-01	28.61	28.23	20.81	25.49	29.62	1.07	1.07	0.83	0.90	0.96
KICT-02	27.90	28.20	18.76	23.52	27.95	1.18	1.19	0.79	0.90	1.01
KICT-03	20.44	20.39	16.00	18.93	21.78	0.93	0.92	0.78	0.82	0.87
KICT-04	37.59	36.53	29.66	34.54	39.54	1.09	1.10	0.83	0.89	0.94
KICT-05	47.24	46.40	35.84	43.84	52.16	1.04	1.06	0.77	0.84	0.93
KICT-06	21.55	22.71	14.57	19.66	25.13	0.74	0.75	0.58	0.65	0.72
KICT-07	33.25	32.60	29.45	32.90	36.10	1.16	1.16	0.86	0.93	0.99
KICT-08	37.53	38.02	24.09	30.84	38.03	0.90	0.91	0.66	0.71	0.77
KICT-09	18.39	18.10	13.93	17.38	21.10	0.80	0.78	0.70	0.73	0.77
KICT-10	48.03	50.73	38.27	46.12	52.21	1.11	1.13	0.74	0.82	0.90

*The table also shows the mean, 5th, and 95th percentiles of the corresponding parameters for the probable rainfall fields. All of the parameters are estimated for hourly accumulations averaged over a 4-km grid.

Table 6.3* The power spectrum slope estimated using ordinary least squares regression for the radar-rainfall data obtained from Pseudo-PPS (PPPS) and Hi-Fi algorithms.

Storm ID	Power Spectrum Exponent β				
	PPPS	Hi-Fi	Probable Rainfall		
			$Q_{0.05}$	Mean	$Q_{0.95}$
KICT-01	2.53	2.53	2.15	2.29	2.41
KICT-02	2.71	2.73	2.17	2.35	2.51
KICT-03	2.25	2.25	2.01	2.11	2.21
KICT-04	2.76	2.77	2.33	2.46	2.57
KICT-05	2.71	2.72	2.30	2.44	2.58
KICT-06	2.05	2.10	1.80	1.94	2.09
KICT-07	2.75	2.75	2.37	2.48	2.58
KICT-08	2.57	2.60	2.10	2.25	2.40
KICT-09	2.07	2.05	1.86	1.99	2.09
KICT-10	2.88	2.94	2.36	2.49	2.61

*The table also shows the mean, 5th, and 95th percentiles of the corresponding parameters for the probable rainfall fields. All of the parameters are estimated for hourly accumulations averaged over a 4-km grid.

Table 6.4* Parameters of the universal multifractal model estimated using the Double Trace Moment technique from the radar-rainfall data obtained from Pseudo-PPS (PPPS) and Hi-Fi algorithms.

Storm ID	Intermittency Parameter C_1					Multifractality Exponent α				
	PPPS	Hi-Fi	Probable Rainfall			PPPS	Hi-Fi	Probable Rainfall		
			$Q_{0.05}$	Mean	$Q_{0.95}$			$Q_{0.05}$	Mean	$Q_{0.95}$
KICT-01	0.36	0.37	0.31	0.32	0.35	0.99	0.99	0.80	0.89	0.96
KICT-02	0.27	0.28	0.25	0.28	0.32	1.09	1.07	0.86	0.98	1.08
KICT-03	0.32	0.32	0.28	0.31	0.34	1.07	1.05	0.93	1.02	1.09
KICT-04	0.25	0.26	0.22	0.23	0.25	1.02	1.03	0.94	1.01	1.06
KICT-05	0.25	0.25	0.20	0.21	0.23	1.18	1.17	1.07	1.15	1.22
KICT-06	0.32	0.33	0.26	0.29	0.33	1.12	1.11	1.00	1.11	1.20
KICT-07	0.24	0.26	0.21	0.22	0.23	1.12	1.11	1.13	1.19	1.25
KICT-08	0.27	0.28	0.21	0.23	0.27	1.39	1.39	1.03	1.15	1.25
KICT-09	0.40	0.42	0.34	0.37	0.40	1.07	1.02	0.94	1.03	1.11
KICT-10	0.21	0.21	0.16	0.18	0.19	1.33	1.31	1.13	1.24	1.33

*The table also shows the mean, 5th, and 95th percentiles of the corresponding parameters for the probable rainfall fields. All of the parameters are estimated for hourly accumulations with the spatial resolution 4-km.

Table 6.5* Hurst exponent estimated using equation 6.6 for the radar-rainfall data obtained from Pseudo-PPS (PPPS) and Hi-Fi algorithms.

Storm ID	Hurst Exponent *H*				
	PPPS	Hi-Fi	Probable Rainfall		
			$Q_{0.05}$	Mean	$Q_{0.95}$
KICT-01	0.50	0.51	0.28	0.35	0.40
KICT-02	0.56	0.57	0.28	0.36	0.43
KICT-03	0.35	0.35	0.21	0.26	0.30
KICT-04	0.55	0.56	0.32	0.38	0.43
KICT-05	0.54	0.55	0.31	0.37	0.43
KICT-06	0.27	0.30	0.11	0.17	0.23
KICT-07	0.55	0.56	0.35	0.40	0.44
KICT-08	0.50	0.52	0.22	0.28	0.35
KICT-09	0.33	0.32	0.19	0.24	0.29
KICT-10	0.61	0.63	0.31	0.37	0.43

*The table also shows the mean, 5th, and 95th percentiles of the corresponding parameters for the probable rainfall fields. All the parameters are estimated for hourly accumulations with the spatial resolution 4-km.

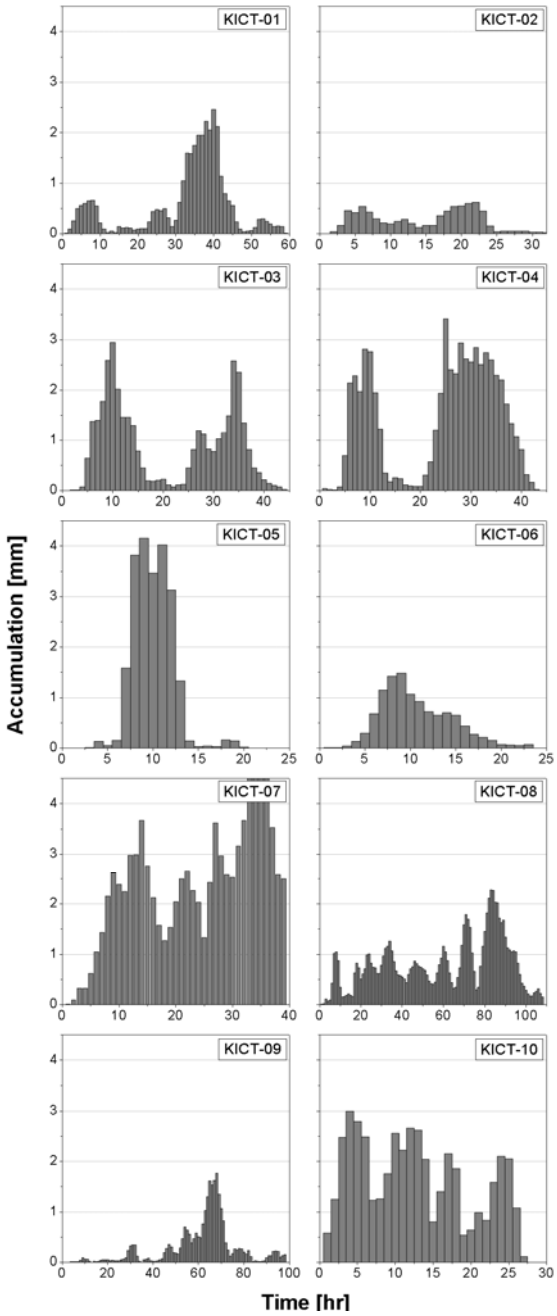

Figure 6.1 Time series of hourly accumulations for the selected storms. The radar-rainfall fields are generated using the Pseudo-PPS algorithms. The storm IDs are indicated on each panel. The maximum accumulation for KICT-07 is 5.5 mm. However, we truncated the axis at 4.5 mm for clarity.

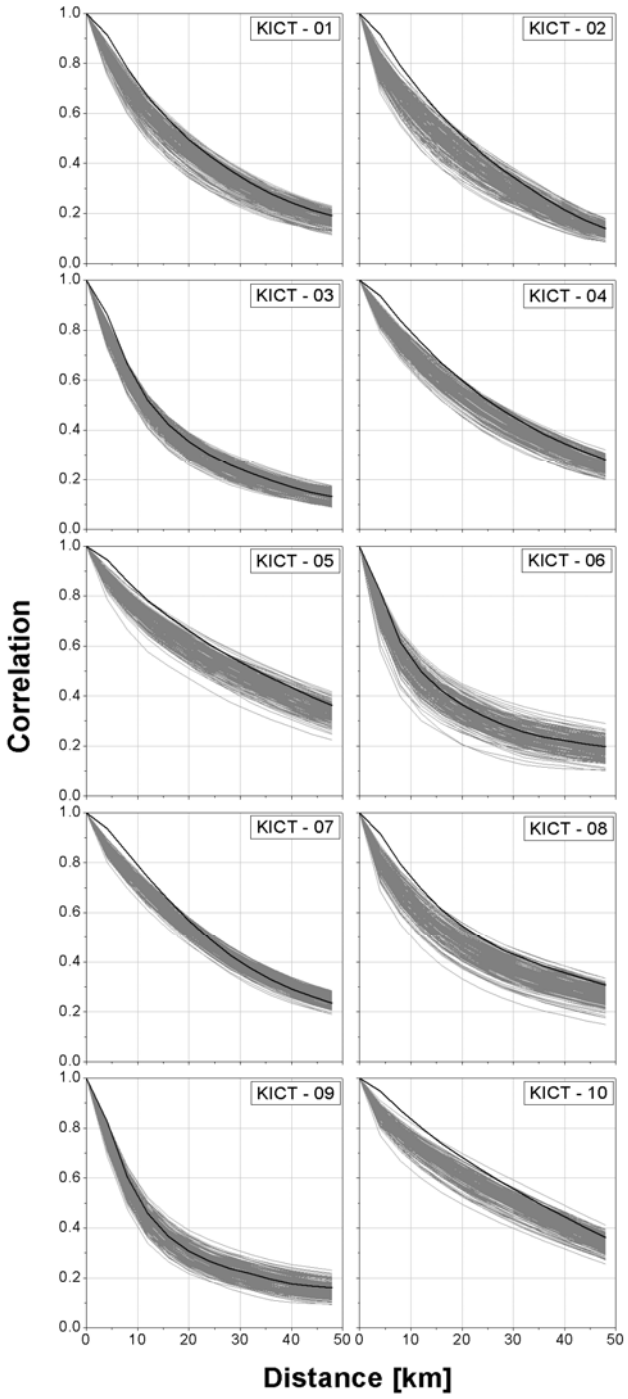

Figure 6.2 Effect of radar-rainfall errors on the spatial correlation function. The solid dark line represents the correlation function obtained from the Pseudo-PPS rainfall product, whereas the gray lines are the correlation functions for the ensemble of probable rainfall fields generated using *Villarini et al.* [2009a].

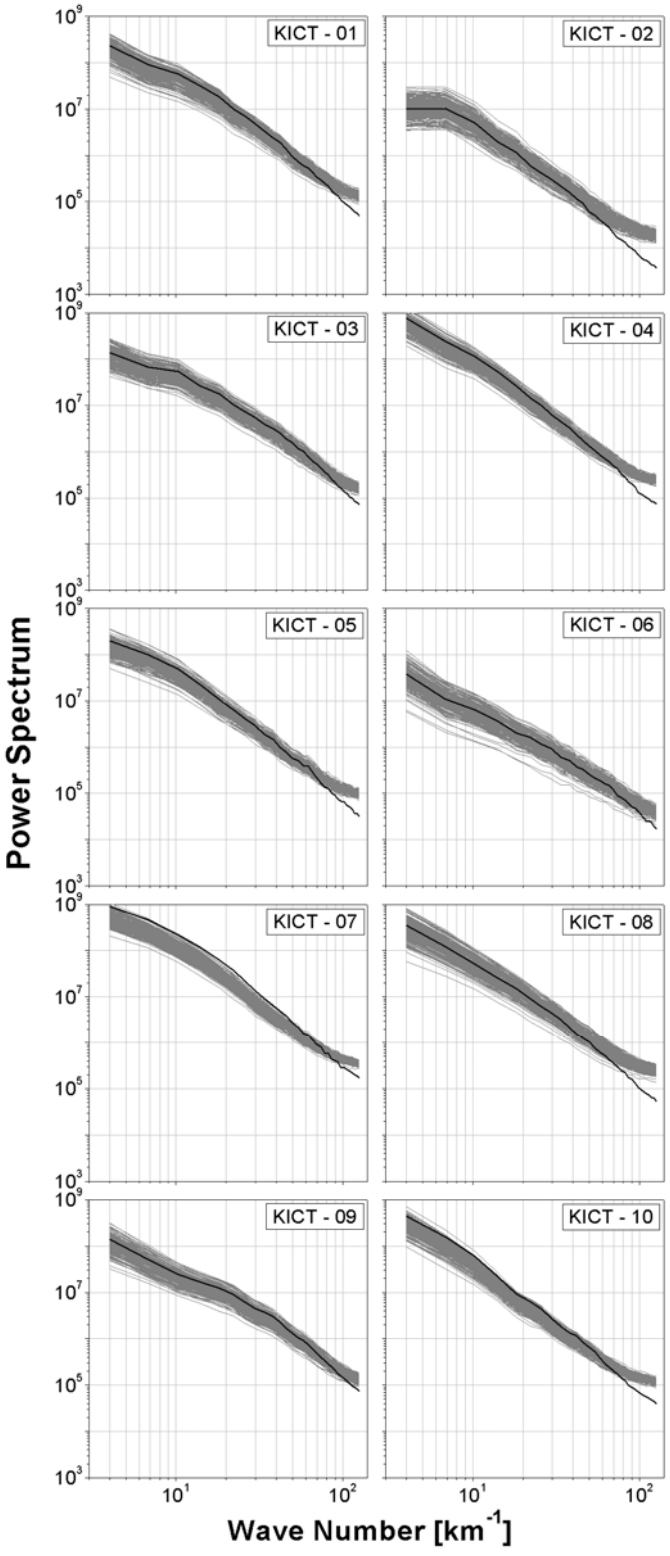

Figure 6.3 Same as Figure 6.2 but for the power spectrum.

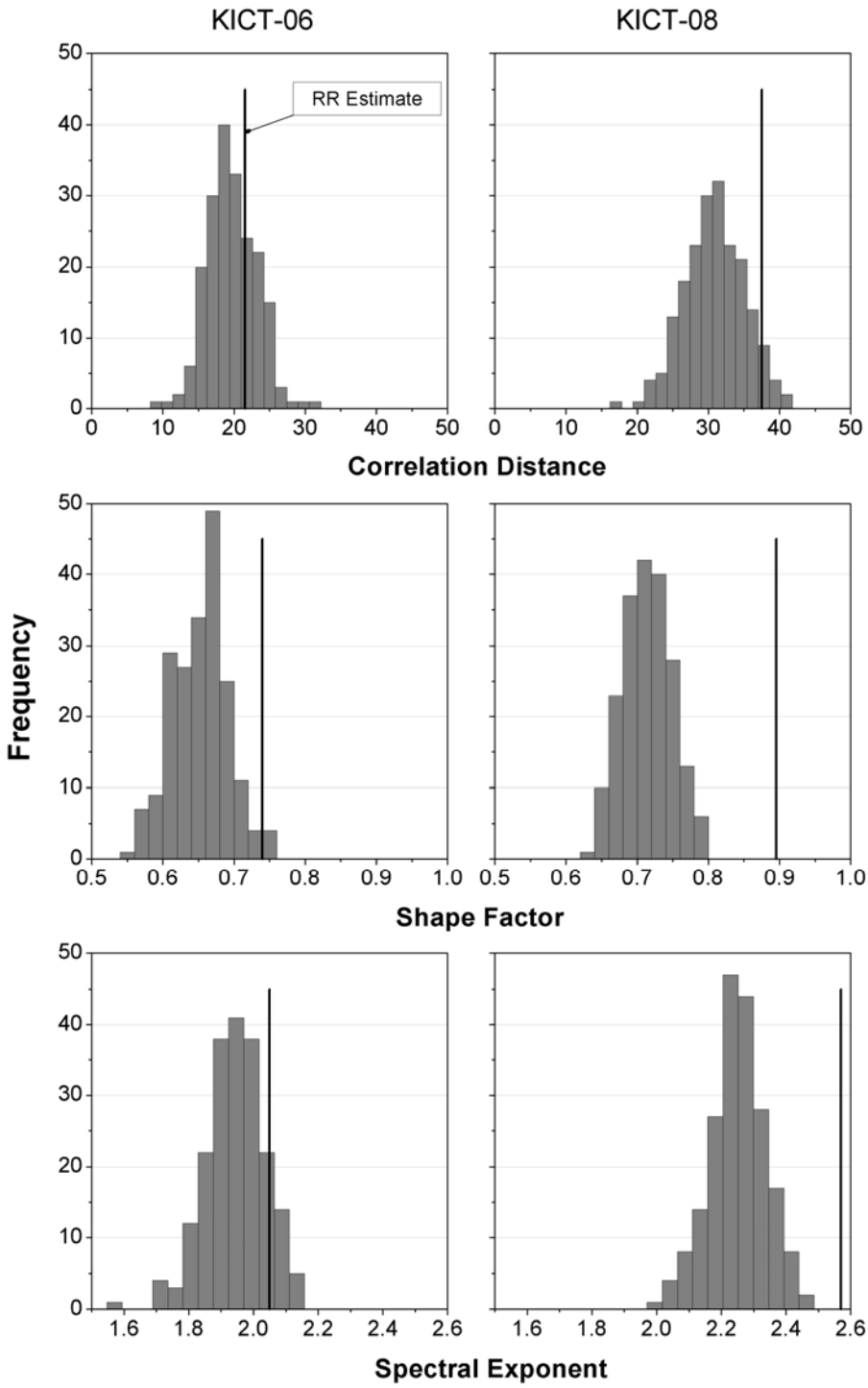

Figure 6.4 Histograms of the parameters of the spatial correlation function (top four panels) and slope of the power spectrum for two rainfall events (KICT-06 and KICT-08) after accounting for radar-rainfall estimation errors. The solid line represents the corresponding estimate obtained from radar-rainfall.

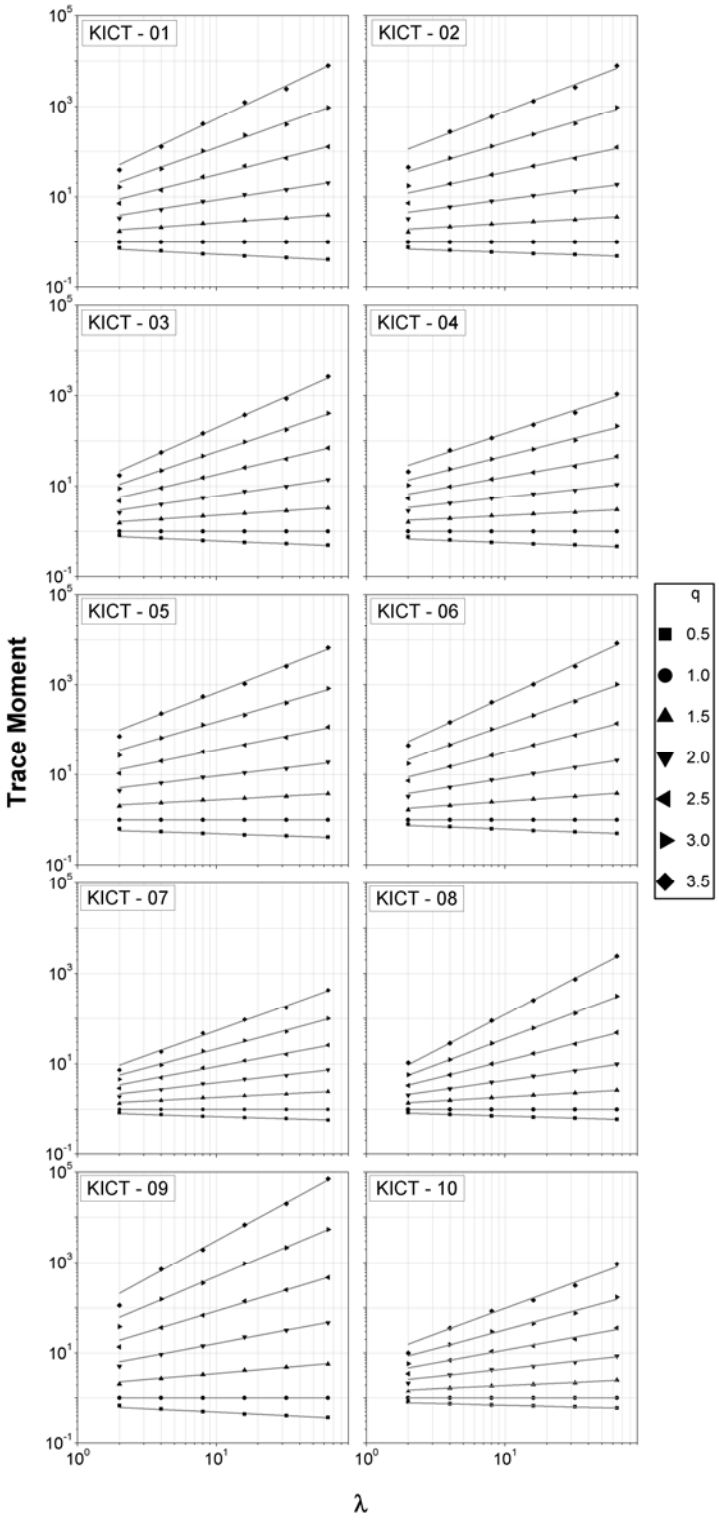

Figure 6.5 Scaling of trace moments of different moment orders q for the PPPS products of ten selected rainfall events in Table 1. The solid lines represent the ordinary least square regression fits.

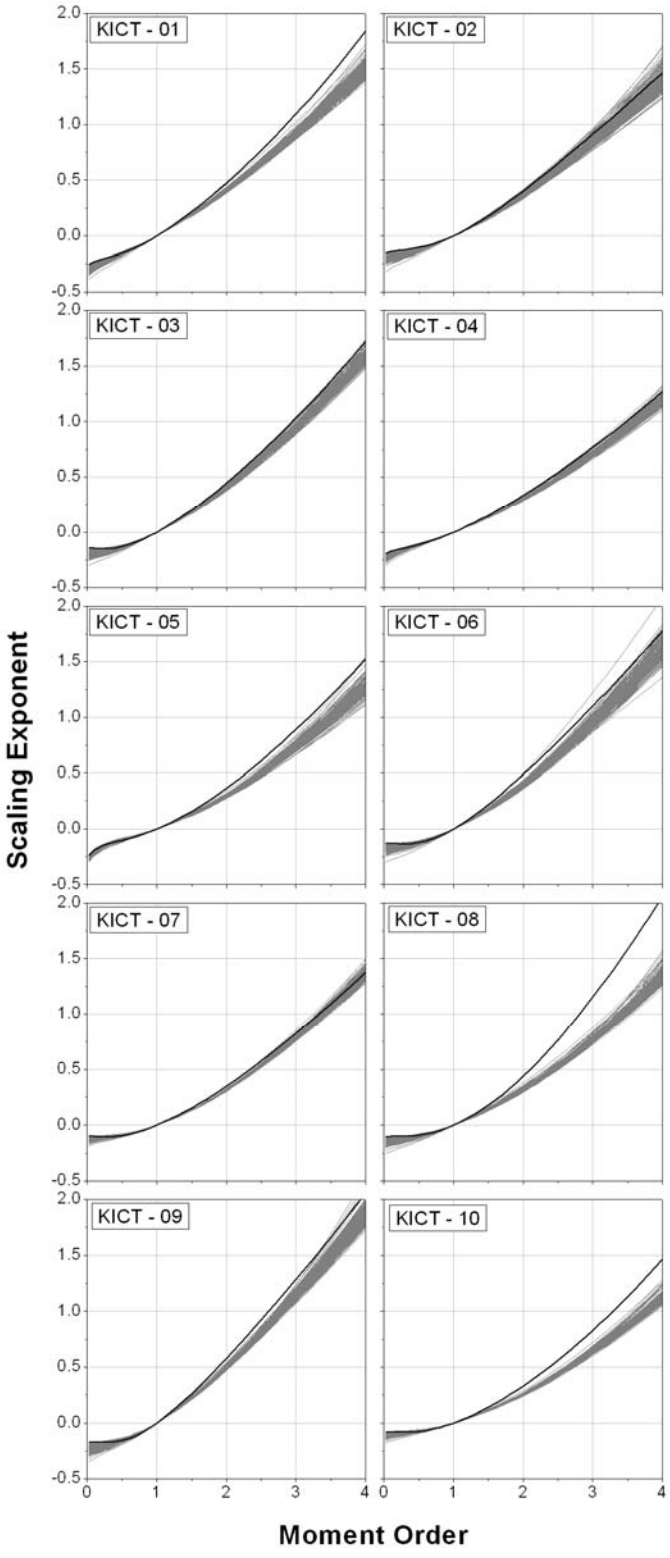

Figure 6.6 Same as Figure 6.2 but for the moment scaling exponents.

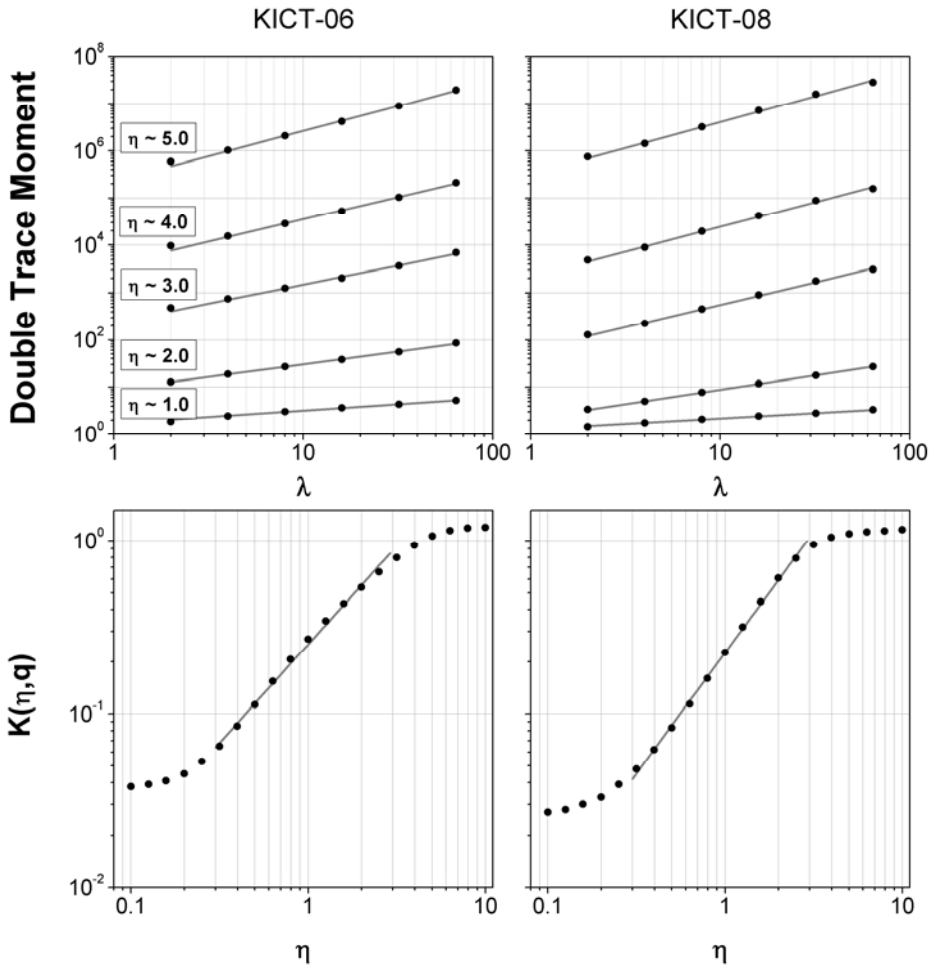

Figure 6.7 Top panels: Scaling of normalized double trace moments (DTM) with respect to the scale ratio for the two rainfall events. The solid lines show the ordinary least squares regression fit, where the first point is not considered in regression. Bottom panels: Scaling of the DTM exponents with respect to the moment order η. The solid lines show the ordinary least squares regression fit, where the first five and last six points are not considered in regression.

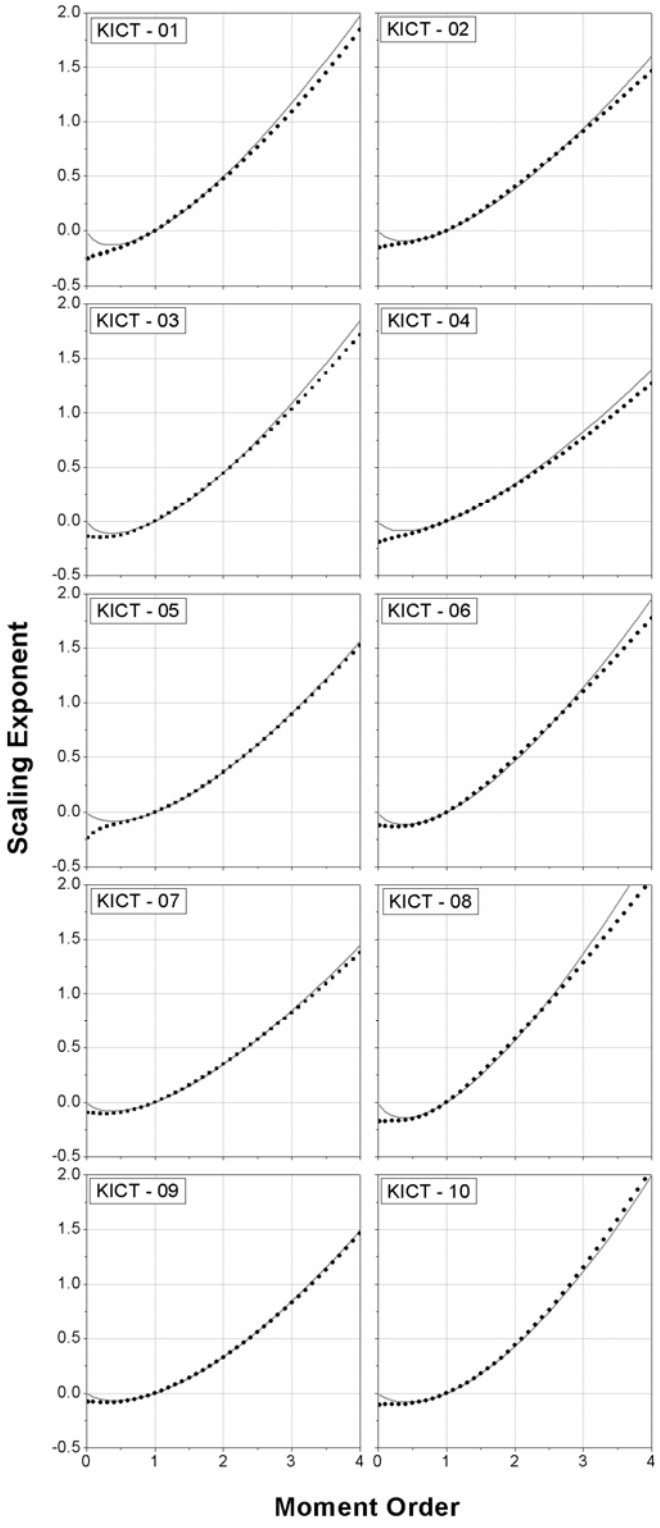

Figure 6.8 Moment scaling functions (solid lines) fitted to the scaling exponents (dots) using the Double Trace Moment technique.

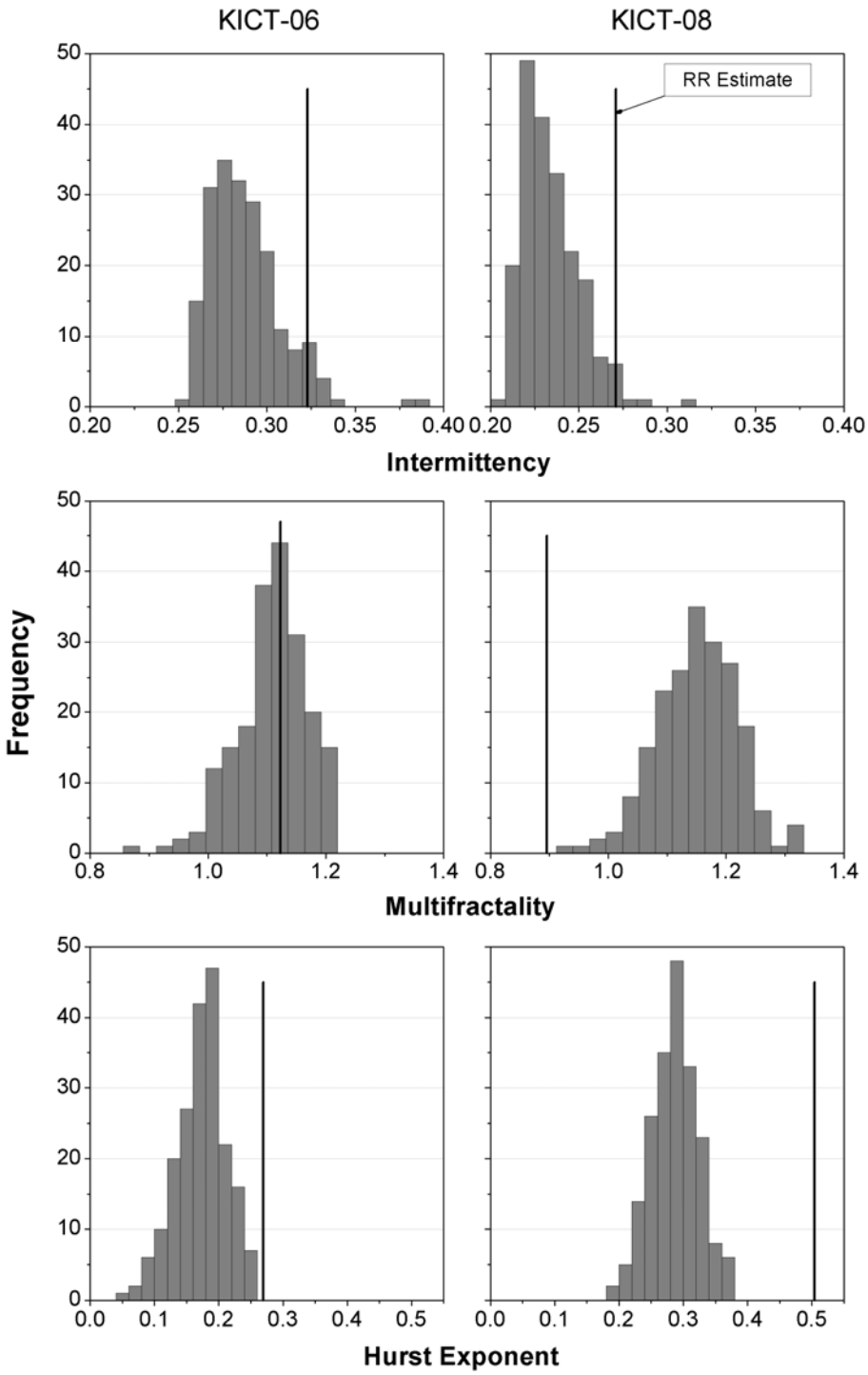

Figure 6.9 Histograms of the universal multifractal model parameters (top four panels) and the Hurst exponent for the two rainfall events (KICT-06 and KICT-08) after accounting for radar-rainfall estimation errors. The solid line represents the corresponding estimate obtained from radar-rainfall data.

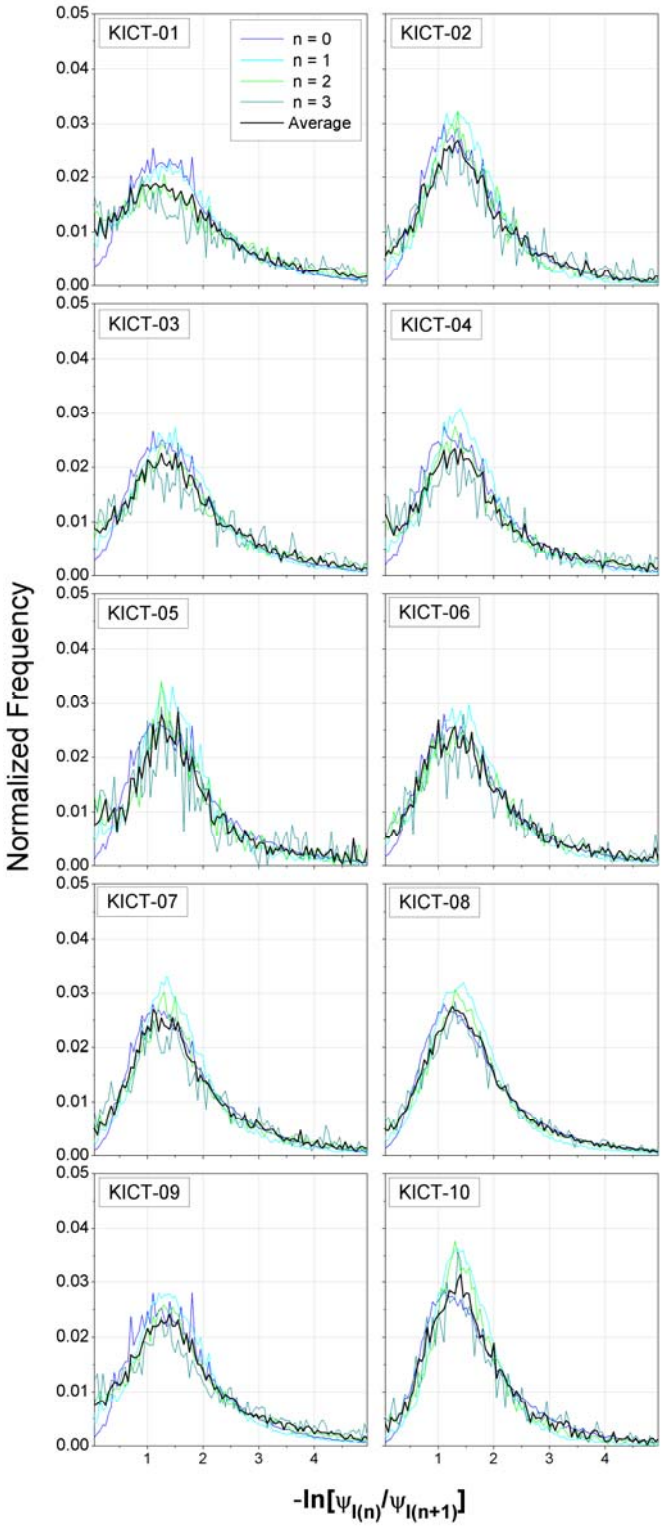

Figure 6.10 Probability distribution of the logarithm of breakdown coefficients for gradient fields of the selected ten rainfall events at different spatial scales *l(n)*.

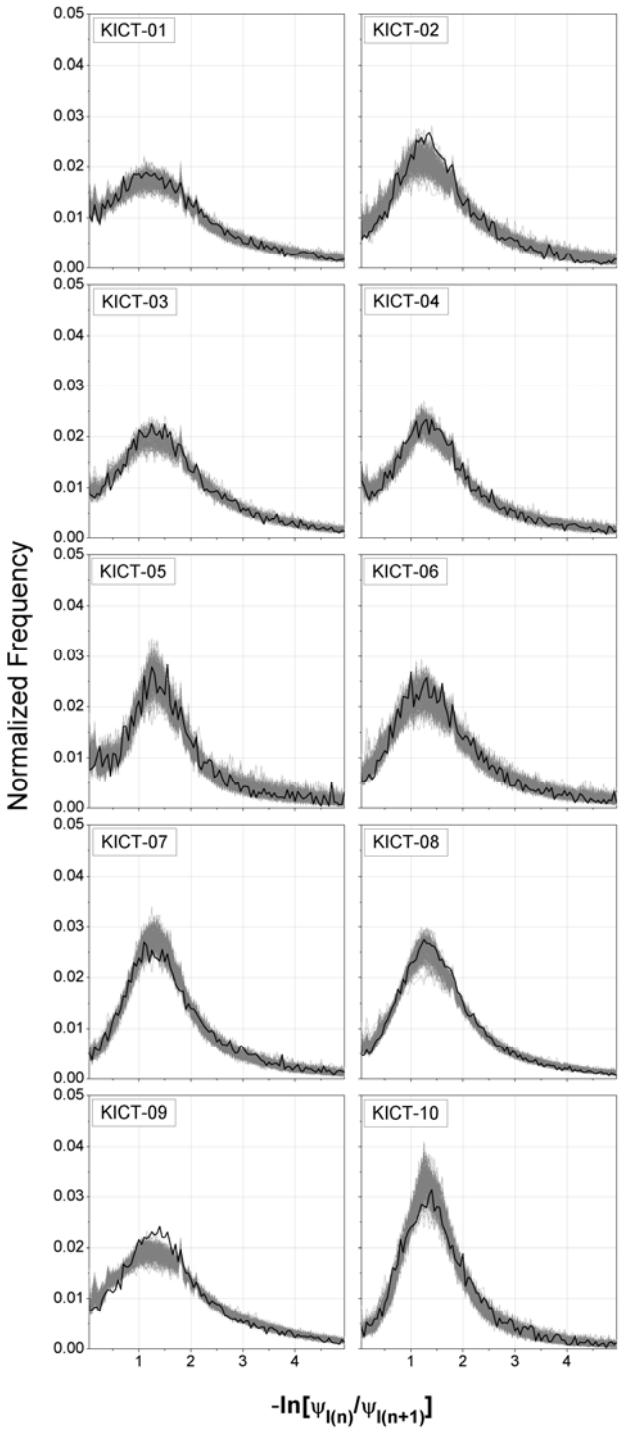

Figure 6.11 Impact of radar-rainfall estimation errors on the average probability distribution of the logarithm of breakdown coefficients for the gradient fields. The dark solid line and the gray lines represent the scaling exponents estimated for radar-rainfall and probable rainfall fields (after accounting for radar-rainfall errors), respectively.

CHAPTER 7

EFFECT OF RADAR-RAINFALL UNCERTAINTIES ON THE PEAK FLOW SCALING EXPONENTS: PRELIMINARY RESULTS

7.1 Introduction

The variability observed in the radar-rainfall fields is a combination of natural variability in rainfall and the variability induced by the radar-rainfall errors. In the previous chapter, we investigated the role of radar-rainfall uncertainties on the spatial characterization of rainfall events. This chapter takes a step forward by propagating the errors through the distributed hydrologic model and quantifies the effect of uncertainties on the statistical structure of peak flows.

7.2 Methodology

We selected a radar-rainfall event over the Whitewater River basin and obtained the peak flows for the event from the hydrologic model CUENCAS. We then generated 100 probable rainfall events employing *Ciach et al.* [2007] and *Villarini et al.* [2009a]. These probable rainfall fields are conditioned on the selected radar-rainfall event but account for the radar-rainfall uncertainties. It should be noted that the uncertainty model [*Ciach et al.*, 2007] and rainfall generator [*Villarini et al.,* 2009a] were developed for the Oklahoma region for the radar-rainfall products at hourly time scales and spatial resolution of 4 km. In this study, we not only transferred the error model and its parameters to a different geographical region, but also assumed the parameters to be valid for spatial resolution of 1 km and temporal scale of 15 min. We selected Whitewater River basin as the Monte Carlo simulation framework is computationally expensive for the larger basins such as Iowa River or Cedar River basins. We estimated the peak flow scaling exponents for the NEXRAD radar-rainfall event and for the probable rainfall events. The effect of radar-rainfall uncertainties can be quantified by comparing the scaling exponents for real event and the error-free probable rainfall events. Quantifying

the impact of radar-rainfall uncertainties on peak flows by comparing the peak flows from radar-rainfall fields with those of probable rainfall fields

7.3 Results and Discussion

We selected rainfall event that occurred from 26th to 27th July 2006 and lasted for ~ 23 hours. We assumed that the runoff generation is Hortonian, and flow to the channels is instantaneous. Figure 7.1 shows the peak flow scaling structure obtained for the rainfall measured by the KICT, Wichita, NEXRAD radar, and one of the probable rainfall events generated using *Villarini et al.* [2009a]. The scaling structure is typical of the realistic rainfall fields, i.e. large variability at smaller scales, absence of sharp scale break, and less variability towards the large scale basins (Figure 7.1). We then estimated the scaling exponents and coefficients in Hortonian framework. It can be seen from the Figure 7.1 that the scaling exponents and the intercepts are different for radar-rainfall and probable rainfall. However, it should be noted that here we are only comparing with one of the 100 realizations of probable rainfall events. We estimated the scaling exponents for the peak flows corresponding to all 100 probable rainfall events. Figure 7.2 compares the histograms of the intercepts and scaling exponents for the 100 probable rainfall fields with those of the radar-rainfall event. Radar-rainfall errors do not seem to significantly affect the scaling exponents and the intercepts (Figure 7.2). However, this is only a preliminary result and must be taken with caution before generalizing it to larger basins.

It should be noted that in this chapter we presented the framework and some early results. Research needs to be carried out particularly in the area of radar-rainfall error structure at very high space-time resolutions for different radars, before we perform a thorough investigation of the effect of radar-rainfall errors on the peak flow scaling structure.

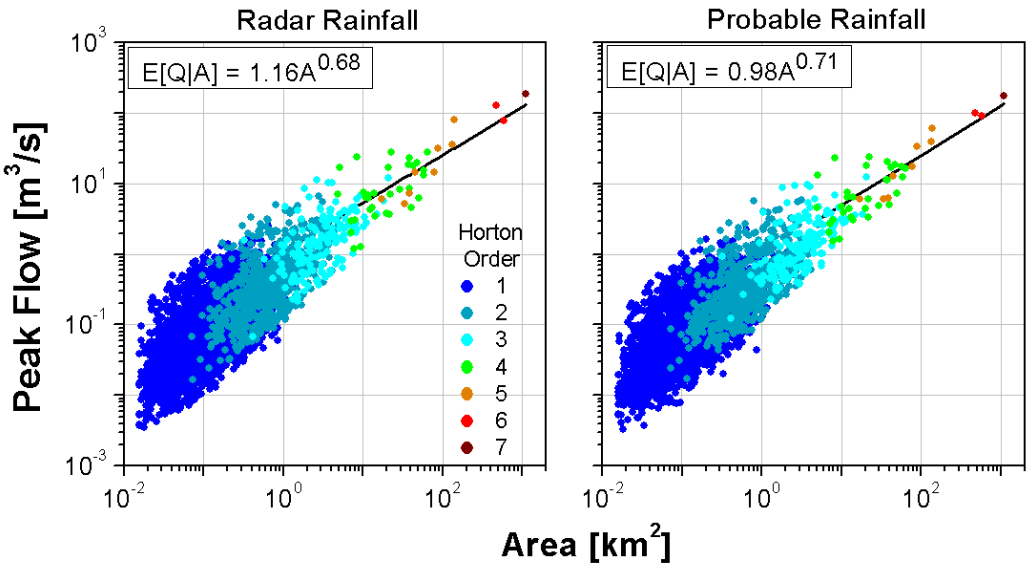

Figure 7.1 Scaling of peak flows from the Whitewater River basin for (left panel) the rainfall event measured by KICT weather radar at Wichita, Kansas, and (right panel) the one of the probable rainfall events generated using *Villarini et al*. [2009a].

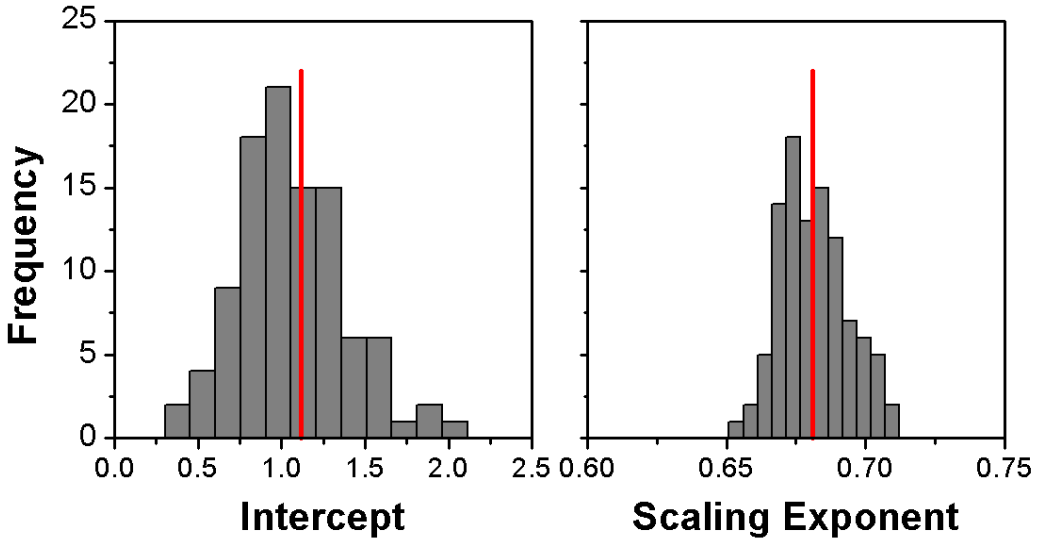

Figure 7.2 Comparison of the regression equations (intercept and exponent) for the peak flow scaling behavior corresponding to the radar-rainfall events measured by the KICT weather radar at Wichita, Kansas with the histograms of intercepts and scaling exponents obtained for probable rainfall events generated using *Villarini et al.* [2009a].

CHAPTER 8

SUMMARY AND CONCLUSIONS

The complex interaction of the rainfall with the landscape together with the lack of measurement techniques to obtain spatially distributed parameters makes the problem of flood prediction quite challenging. Scale-invariance behavior of peak flows offers us an alternative framework to predict the peak flows. However, a physical understanding of mechanisms behind the observed scaling behavior is necessary before we implement scaling-based hydrologic models in a predictive framework. In the last two decades, several studies have focused on scale-invariant behavior of rainfall and the physical mechanisms behind such power-laws. While some of them were data-based, others were based on simulations under idealized conditions such as spatially uniform or cascade-based rainfall on deterministic or random self-similar river networks. *Furey and Gupta* [2007] presented a five-step framework for diagnosing the observed scale-invariance in single-event peak flows. However, such a framework can be applied only for a basins having large number of rainfall and streamflow gauges operating concurrently, which is the case only for small research based watersheds such as Goodwin Creek Basin in Arizona. For this reason, we resorted to simulation framework starting from idealized rainfall scenarios to complex space-time rainfall events on river basins in Iowa and Kansas with maximum area of ~7250 km^2. Instead of a grid-based structure, we employed hillslope-link decomposition of the landscape in our distributed hydrologic model.

It is important to understand the relative roles of all the factors such as rainfall, runoff generation on hillslopes, channel network topology, and channel routing on the scaling exponents. Space-time rainfall variability itself is a very complex feature especially with the presence of estimation errors and therefore, we studied only the effect of rainfall on scaling exponents. One could always make the study more complex by

including different hillslope runoff generation mechanisms, channel network topologies and hydraulic geometries, but such an analysis would not provide an insight into relative roles of each component on the scaling exponent. For this reason, we assumed Hortonian runoff mechanism with instantaneous flow into the channels and restricted this study to two watersheds in Iowa and Kansas.

We focused on three aspects of the peak flow scaling structure for all the scenarios: scatter, scale break and the scaling exponents. The results showed that the peak flow scaling exponents for all the scenarios considered in this study are greater than the width function scaling exponent. In Chapter 2, we studied the effect of basic characteristics of rainfall such as storm intensity, duration and advection on the scaling structure of peak flows. For a fixed intensity, the scaling exponent increased with an increase in the rainfall duration, and for a fixed duration, the scaling exponent decreased with intensity, but only for nonlinear channel routing scenario. We also found that as the storm advection velocity increased, the scaling exponent decreased exponentially.

We then introduced variability in the rainfall fields in Chapter 3 and investigated its effect on the peak flow scaling structure. The main effect of storm variability, as characterized by variance, spatial correlation and the spatial intermittency, was to increase the scatter in the scaling structure of the peak flows. However, at larger spatial scales, the effect of variability decreases, as seen from peak discharges at the outlet of the basin. The analyses in Chapters 2 and 3 were carried out on ~1200 km^2 Whitewater River basin in Kansas. In Chapter 4, we repeated our simulation experiments but on ~7250 km^2 Iowa River basin up to Marengo, Iowa. The results for the Whitewater and Iowa River basins strongly indicated that under idealized conditions of runoff generation, flow routing on hillslopes and in channels, the smaller scale basin response was dominated by the rainfall intensity (and spatial distribution), while the hydrologic response of larger scale basins was driven by rainfall volume, river network topology and flow dynamics. In general, we conclude that the effect of rainfall variability on the peak

flow scaling behavior is larger for smaller scale basins. As we moved towards the larger basins, the river network effectively integrated the variability noticed at smaller scales. However, the efficiency of the river network in averaging the variability also depends on the rainfall intensity distribution, localization and correlation structure.

In Chapter 4, we also relaxed the assumption of instantaneous flow to the channel to study the effect of hillslopes on the scaling structure of peak flows. We incorporated hillslope velocity into the CUENCAS hydrologic model. Three different hillslope velocities were considered: 1) velocity corresponding to instantaneous flow to the channel, 2) velocity of 1.0 m/s, and 3) velocity of 0.1 m/s. Results suggested that the hillslope velocities strongly affect the scatter in the peak flow scaling structure at small scales, and the location of the scale break. However, their impact on the scaling exponents was not found to be significant. Besides the role of hillslopes, we also analyzed the effect of channel network extraction on the peak flow scaling behavior. We considered digital elevation models with a resolution of 30 m and 90 m, extracted the channel networks, and imposed spatially uniform rainfall for different durations. The results showed that channel network extraction had huge impact on the peak flows across the full range of scales considered. It not only affected the scatter at the smaller scales but also significantly altered the scaling exponents.

Throughout the thesis, we modeled rainfall space-time variability in terms of stationary random fields with a certain intermittency and correlation structure in space and time. Although the emphasis of this approach is more on generation of space-time random fields and lacks a direct link to the physical aspects of the rainfall process, one can qualitatively relate the statistical parameters to the meteorological aspects such as back-building thunderstorms, squall lines and convective systems. For instance, the size of the convective system with respect to the size of the basin is an important characteristic that controls the basin response. This information is embedded in the correlation distance and shape factor in equation 4. There are other ways of

characterizing the space-time variability of rainfall fields such as modeling of rainfall based on spatial cluster processes for rain cells, dynamic modeling of rainfall based on partial differential equations for mass and momentum conservation and scaling-based modeling of rainfall space-time structure. It would be interesting to employ these models and relate the parameters of rainfall models to the characteristics of peak flow scaling structure.

Anthropogenic alteration of the landscape will have an impact on the extracted drainage network and subsequently on the aggregation of the flows affecting the larger scale basin response. The hydraulic geometry of the channels is another key factor which is strongly related to the channel-floodplain interactions and influences the storage zones, movement of the flood waves and the travel time within the channel. Therefore, it is expected to have an influence on the peak flow scaling structure. The CUENCAS hydrologic modeling framework allows for the inclusion of this feature by modifying the local velocity law. However, investigation of this factor is beyond the scope of this work, and our strategy is to avoid any aspects in the dynamics that can obscure the effect of rainfall variability.

Throughout this study, we have estimated the Horton ratios by ordinary least squares regression of logarithms of arithmetic averages with the Horton order (See Figure 2). Based on a simulation study, *Furey and Troutman* [2008] suggested the use of individual quantities (for example, areas or peak flows) instead of arithmetic averages. Since the main focus of the study is on the role that rainfall variability plays in the scaling structure of peak flows, our use of arithmetic averages instead of individual quantities in the Horton analysis would not affect the results. Also, in this study, we have assumed that the runoff generation is Hortonian with no infiltration so that we can focus on the role that rainfall plays in the statistical structure of peak flows. The key issue is how to specify the infiltration threshold for each of 100,000 hillslopes in CUENCAS, which differ due to spatial variability in soil and vegetation properties. This problem of

"dynamic parametric complexity" is a major research problem [e.g., *Gupta*, 2004; *Furey and Gupta*, 2007] and not addressed in this study.

To investigate the effect of radar-rainfall errors on the peak flows, we need characteristics of RR errors. In chapter 5, we developed a framework to estimate the RR error spatial correlation structure (ESC). We applied the framework to estimate the ESC of hourly NEXRAD rainfall products over Oklahoma. The errors were found to be significantly correlated with a correlation distance of about 20 km. In chapter 6, we investigated the impact of errors on the spatial characterization of rainfall. The results indicated that the errors induce spurious correlations in the radar-rainfall fields, particularly for small scales. We then propagated the rainfall estimation uncertainties in a Monte Carlo framework through our distributed hydrologic framework to study their impact on the peak flow scaling behavior. Preliminary results indicated that although radar-rainfall errors affect spatial characterization of rainfall events, their impact on the peak flows was not significant.

Most of the hydrologic models in the past were developed with a short term goal of predicting the hydrographs in the basin. Therefore, many hydrological modeling studies were carried out without strong theoretical basis and relying heavily on calibration. And they did predict the hydrographs with reasonable accuracy. In an operational setup, the forecasters have a responsibility to forecast the outflows from a basin. Therefore they are compelled to employ such models. However, there is a strong need for parallel research to thoroughly understand the hydrological processes across wide range of scales, and develop a robust theoretical framework which is scale-independent. We believe that a framework based on hillslope-link landscape decomposition and conservation of mass and momentum in the basin is a way forward instead of ad hoc grid based or sub-basin based landscape partitioning. We understand that this approach does suffer from the same fundamental problem as the traditional grid based hydrologic models, which is lack of clear understanding of the hillslope hydrology.

However, it is physically meaningful to incorporate the new developments in the field of hillslope hydrology into the hillslope-link based framework. In a systematic and diagnostic manner, we analyzed the role of rainfall variability and the river network in the statistical self-similarity of peak flows. By incorporating latest developments in hillslope hydrology, and moving closer to the reality, a comprehensive geophysical theory for floods can be developed.

APPENDIX A

SIMULATION BASED QUANTIFICATION OF THE EFFECTS OF SAMPLING ON THE ESTIMATION OF SPATIAL CORRELATION STRUCTURE

A.1 Introduction

Spatial phenomena (e.g., soil moisture, hydraulic conductivity or rainfall) are often modeled as realizations of a stationary stochastic process, where their spatial variability is characterized in terms of second-order statistics such as the semivariogram and the correlation structure [e.g., *Lebel et al.*, 1987; *Bras and Rodriguez-Iturbe*, 1993; *Mohanty et al.*, 1994; *Baachi and Kottegoda*, 1995; *Western et al.*, 1999; *Lark*, 2000b; *Germann and Joss*, 2001; *Gebremichael and Krajewski*, 2004; *Ciach and Krajewski*, 2006; *Villarini et al.*, 2008]. The semivariogram is a measure of dissimilarity in the process and is defined for an intrinsically stationary process. A stochastic process $Z(\mathrm{u})$ ($\mathrm{u} \equiv (x,y) \in \Omega$) is considered intrinsically stationary and isotropic if the mean is constant, and the semivariogram, defined as

$$\frac{1}{2} \cdot Var\{Z(\mathrm{u}_1) - Z(\mathrm{u}_2)\} = \gamma\{\mathrm{u}_1 - \mathrm{u}_2\} = \gamma(h) \tag{A.1}$$

is a function of distance lag $h = \mathrm{u}_1 - \mathrm{u}_2$. On the other hand, the correlation function is a measure of similarity in the process and is defined for a second-order stationary process, where the mean and variance are constant and covariance is a function of distance lag

$$Cov\{Z(\mathrm{u}_1), Z(\mathrm{u}_2)\} = C\{\mathrm{u}_1 - \mathrm{u}_2\} = C(h) \tag{A.2}$$

A second-order stationary process will always be intrinsically stationary, but the converse need not always be true. The advantages of the semivariogram over correlation structure are that the former does not require the estimation of the mean or an assumption about population variance [e.g., *Cressie*, 1993]. However, the correlation structure has been widely used to characterize the spatial dependency and is required by many

applications [e.g., *Bras and Rodriguez-Iturbe*, 1993; *Morrissey et al.*, 1995; *Ciach and Krajewski*, 1999; *Habib et al.*, 2004]. The two most important properties that can be obtained from a correlation function are the correlation length and the behavior of the function at the origin. The former is a measure of the distance over which the process is correlated; it can be on the order of hundreds of meters for processes such as soil moisture or subsurface properties and on the order of hundreds of kilometers for atmospheric processes [e.g., *Rea and Knight*, 1998; *Western et al.*, 1999; *Michalak et al.*, 2004]. The behavior of the function at the origin provides information about measurement errors and the smoothness of the process.

The correlation structure of a process has traditionally been estimated from data collected from a network of sampling stations. With the advent of radar and satellite remote sensing products, the sampling network consists of a regular two-dimensional grid of pixels. In both cases, the network available for estimation is governed by geographic, economic and technological criteria. Also, in such fields as rainfall and soil moisture, multiple realizations of the process are available to estimate the correlation structure, whereas in others (e.g., ground water hydrology), only a single or few realizations are available. In the following two paragraphs, we briefly review the literature pertaining to sampling problems in the estimation of the correlation structure and the semivariogram.

Several studies have described the problems associated with estimation of the covariogram or correlation function from a network of sampling points [e.g., *Journel and Huijbregts*, 1978; *Krajewski and Duffy*, 1988; *Cressie and Grondona*, 1992; *Dubin*, 1994]. *Journel and Huijbregts* [1978] suggested that a minimum of 30 pairs be used at each lag to estimate the covariogram and variogram. They also show that non-centered covariogram estimates are unbiased, and the bias in the centered covariogram estimates vanishes when the dimension of the field is much larger than the correlation range. *Krajewski and Duffy* [1988] conducted a simulation experiment to investigate the effect of the number of sampling points and the ratio of correlation length to the domain size on

the estimated mean and variance of the Gaussian random fields characterized by exponential, Bessel and Gaussian correlation functions. They showed that when the ratio is greater than 0.20, the bias in the estimated variance is large even when a large number of sampling points (1000) are used. However, the bias drops rapidly when the ratio is less than 0.20 for even a small number of sampling points (25 -100). For a one-dimensional linear model with a polynomial trend, *Cressie and Grondona* [1992] derived analytical expressions for the bias in the covariogram and variogram estimates for situations where both measures exist. They concluded that bias in the variogram is an order of magnitude less than the corresponding covariogram estimate. *Dubin* [1994] investigated the performance of three different methods to estimate the parameters of the correlation function: 1) Ordinary least squares (OLS) fit to the experimental correlogram, 2) Maximum likelihood (ML) method and 3) Generalized least squares (GLS) method, using simulated two-dimensional fields (with a size of 100 units) with correlation lengths of 10 and 20 units. The study concluded that both the OLS and GLS techniques underestimate the parameters, and the bias is higher for the higher degree of spatial dependence. It further concluded that the ML method performs best in estimating the parameters and is independent of the degree of spatial dependence of the underlying process.

The effects of sampling on the estimation of the semivariogram were reported by many studies [e.g., *Journel and Huijbregts*, 1978; *Warrick and Myers* 1987; *Russo and Jury*, 1987; *Morris*, 1991; *Zimmerman and Zimmerman*, 1991; *Cressie and Grondona*, 1992; *Lark*, 2000; *Zheng and Silliman*, 2000]. *Warrick and Myers* [1987], based on a series of constraints, developed a method for the reliable estimation of the semivariogram by optimizing the number of points and their locations in a network. Based on a simulation study, *Russo and Jury* [1987] suggested that the number of points should exceed 100 and that the minimum distance between the sampled points should be smaller than half the range of the underlying process for the estimated correlation length to

approach the correlation length of the phenomenon under study. *Russo and Jury* [1987] also reported that as the correlation range increases with respect to the size of the domain, the bias in its estimation decreases while the estimation variance increases. *Morris* [1991] argued that for correlated data, the number of pairs is not a satisfactory measure of estimation accuracy, and he derived an expression for the maximum equivalent uncorrelated pairs separated by any given distance in a Gaussian stationary process characterized by a non-increasing convex correlation function. *Zimmerman and Zimmerman* [1991] conducted a simulation study to compare semivariogram estimators for different semivariogram models with varying degrees of spatial dependence. They concluded that the ML estimator is superior to others, particularly for the exponential semivariogram (with a correlation range spanning from 0.1 to 0.9 units on a 4-unit simulated field) and that the estimates are severely biased when the spatial dependence is strong. Based on a simulation study involving a 300-unit two-dimensional field with a correlation range spanning from 2 to 6 units, *Lark* [2000] compared the ML estimators with an OLS fit to the method-of-moments estimator and reported that the performance of the ML estimator is significantly better when the spatial dependence in the data is weaker. *Zheng and Silliman* [2000] analytically demonstrated that the variance associated with the estimates of the semivariogram depends not only on the sampling network but also on the correlation among the data pairs.

The essence of the aforementioned studies is that for the estimation error to be within certain limits, an estimator of spatial dependence requires a certain minimum effective number of independent pairs at each distance lag, which depends on the strength of spatial dependence of the process. For the correlation function, which is the main focus of the study, the spatial dependence, and therefore the effective number of pairs, is controlled by the correlation length and the behavior of the function at the origin. If only one realization of the process is available, the number of pairs is limited by the density and spatial extent of the sampling network. In the case of remote sensing (2D fields),

where multiple independent realizations are possible, the resolution, size of the domain and number of fields govern the number of independent pairs.

Based on the above discussion, we classify the sampling effects on the estimation of correlation structure into three cases: a) effect of size of the domain used, b) effect of smoothness of the process and c) effect of the number of realizations. The main objective of this simulation study is to systematically quantify the above three effects on the estimation of correlation structures typically observed in water resources and related fields. The simulated stationary random fields are characterized by a two-parameter power-exponential correlation structure of the form:

$$\rho(d) = \frac{C(d)}{C(0)} = \exp\left\{-\left(d/\theta_1\right)^{\theta_2}\right\} \qquad \theta_1 > 0,\ 0 \le \theta_2 \le 2 \tag{A.3}$$

where, $C(d)$ is the covariance for a distance lag *d*, θ_1 is the correlation distance defined as the distance for which the correlation coefficient is equal to 1/e and θ_2 is the shape parameter that controls the shape of the correlation function near the origin. We selected this correlation model because it is widely used to model such spatial processes as hydraulic conductivity and rainfall [e.g., *Ciach and Krajewski*, 1999; *Gebremichael and Krajewski*, 2004; *Stauffer*, 2005; *Ciach and Krajewski*, 2006]. Hereafter, we refer to the ratio of the correlation distance to the size of the estimation domain as the scale factor (SF). The simulation experiments are designed such that, unlike studies discussed in the previous paragraphs, we have a broad spectrum of SF and shape parameters.

We first focus our attention on the case of the spatially correlated Gaussian random fields (RF), and then we extend the analysis to the case of lognormally distributed RF, as it is a close approximation to such geophysical phenomena as hydraulic conductivity and rainfall [e.g., *Kedem and Chiu*, 1987; *Kedem et al.*, 1990; *Xian-Huan*, 1994; *Loáiciga et al.*, 2006]. Instead of employing parametric estimation techniques (e.g., ML, REML and GLS), which involve apriori selection of an underlying correlation

model, we present the correlations that are estimated using Pearson's product moment estimator and quantify the bias in a non-parametric manner.

This chapter is organized as follows. Following the Introduction, Section 2 describes the approach used to simulate Gaussian and lognormal RF. The procedure used to estimate their spatial correlation structure is described in Section 3. We present the results in Section 4, and Section 5 concludes the chapter.

A.2 Simulation of Random Fields

This section describes the procedure used to simulate stationary, isotropic, Gaussian and lognormal RF characterized by the correlation structure given in Equation (3). Note that in Equation (3), when $\theta_2 = 1$, the function is the simple exponential correlation function, and when $\theta_2 = 2$, the function reduces to the Gaussian correlation function. In this study, the correlation distance ranges from 5 to 250, and the shape parameter ranges from 0.1 to 2.0.

A.2.1 Gaussian fields

Many methods have been proposed to simulate stationary, isotropic and correlated Gaussian random fields. The choice among them depends on the desired correlation structure, the resolution and the computational resources available [*Gneiting et al.*, 2005]. We consider matrix factorization, spectral turning bands and circulant embedding techniques and investigate which one is most suitable for our study.

In the matrix factorization technique [e.g., *Davis*, 1987], the variance-covariance matrix is factorized using the Cholesky decomposition method and requires at least $O(n^5)$ floating point operations for simulations on an $n \times n$ grid [*Dietrich*, 1993]. Though the method is asymptotically exact, resulting in an ensemble of realizations with exactly the desired correlation structure, the technique is computationally intensive to simulate large two-dimensional fields. In the spectral turning bands method [e.g. *Mantoglou and Wilson*, 1982], the two-dimensional RF is obtained by summing a series of uniformly

distributed, equivalent one-dimensional processes generated on intersecting lines using a spectral approach. This method is computationally efficient, but the correlation structure of the simulated process will converge to that of the true process only for a number of lines tending to infinity. In Equation (A.3), when $\theta_2 \neq 1$ or $\theta_2 \neq 2$, the closed form expressions for the spectral density do not exist [*Yaglom*, 1987]. Hence, this method cannot be used in this study.

Dietrich and Newsam [1993] and *Wood and Chan* [1994] proposed a new approach to simulate multi-dimensional stationary Gaussian RF that is based on embedding the random field correlation matrix in a matrix with circulant structure. This technique, called circulant embedding, is fast and exact, but one of the main conditions for its validity is that the embedding matrix should be non-negative definite [*Dietrich and Newsam*, 1993; *Dietrich and Newsam*, 1997]. *Dietrich and Newsam* [1997] showed that for two-dimensional simulations, the embedding matrix is always non-negative definite if the correlation function is bounded (e.g., spherical or power models) and if it is completely sampled by the simulation grid. However, for correlation functions with infinite support, such as the one used in this study, they concluded that the condition of non-negativity is satisfied if the simulation domain is sufficiently larger than the correlation distance. For exponential and Gaussian correlation functions, they derived analytical expressions for the size of the domain required for the embedding matrix to be non-negative definite but also acknowledged that these bounds are pessimistic.

Gneiting et al. [2005] proposed two methods, termed cut-off and intrinsic embedding, which are variants of the circulant embedding method. These methods outperform the circulant embedding method in ensuring the non-negativity of the embedding matrix. This is achieved by approximating the desired correlation function with a class of functions that are compactly supported [*Gneiting,* 2002]. While cut-off embedding is for simulating a second order stationary process, the intrinsic embedding method can only be used to simulate fields that are intrinsically stationary. *Gneiting et*

al. [2005] also explored numerically the computational limits of standard circulant embedding, cutoff embedding and intrinsic embedding methods with three different correlation functions, including the two-parameter exponential function used in this study.

Based on the above discussion, we used the circulant embedding technique to simulate stationary, two-dimensional correlated Gaussian fields, and we used the cutoff embedding technique when this technique failed. However, as mentioned in the previous paragraph, the cut-off embedding method is used by approximating the exponential correlation function with a class of compactly supported functions. Furthermore, to ensure non-negativity of the embedding matrix, we simulated random fields on a 100×100 domain, though we used a 50×50 domain for estimation (SF therefore ranges from 0.1 to 5.0). We simulated 1000 realizations for each case with the ensemble mean and the variance equal to 0 and 1, respectively. The sample Gaussian simulations for varying correlation distances and shape factors are shown in Figures A.1 and A.2, respectively.

A.2.2 Lognormal fields

The lognormal fields are obtained by simulating the Gaussian fields and then applying the exponential transformation. However, in the transformation process, we had to ensure that the lognormal fields had the desired correlation structure. For a particular distance lag, the correlation in the Gaussian random field that would result in the desired correlation in the lognormal field after the transformation is obtained as follows: Let X_1 and X_2 be the correlated Gaussian random vectors with mean μ_G, variance σ_G^2 and a correlation coefficient ρ_G; the lognormal random vectors Y_1 and Y_2 were obtained by taking $Y_1 = \exp(X_1)$ and $Y_2 = \exp(X_2)$. Then the lognormal mean, μ_{LN}, variance, σ_{LN}^2, and the correlation coefficient, ρ_{LN}, were related to the corresponding Gaussian parameters as follows [*Mejía and Rodríguez-Iturbe*, 1974]

$$\mu_G = \ln\left[\frac{\mu_{LN}^2}{\sqrt{\mu_{LN}^2 + \sigma_{LN}^2}}\right] \tag{A.4}$$

$$\sigma_G^2 = \ln\left[1 + \frac{\sigma_{LN}^2}{\mu_{LN}^2}\right] \tag{A.5}$$

$$\rho_G = \frac{1}{\sigma_G^2}\ln\left[1 + \rho_{LN}\left(\exp(\sigma_G^2) - 1\right)\right] \tag{A.6}$$

Given the lognormal mean, μ_{LN}, variance, σ_{LN}^2, and correlations, ρ_{LN}, Equations 4-6 can be solved for the corresponding Gaussian parameters. However, for simulations using circulant embedding, the Gaussian correlations obtained from Equation (A.6) have to be expressed in a parametric form given by Equation (A.3). This is accomplished by performing a systematic two-dimensional search of the parameter space, minimizing the root mean squared error between the correlations from Equation (A.6) and the fitted function. For each lognormal correlation structure characterized by a certain combination of correlation distance and a shape parameter, we repeated the above procedure to obtain the corresponding Gaussian parameters, which are always larger than the lognormal ones (Table 1). For example, to obtain a lognormal field with a correlation distance of 40 and a shape parameter of 1.0, we simulated the Gaussian field with a correlation distance of 89.5 and a shape parameter of 1.24, and the field was exponentiated. For lognormal fields with correlation distances beyond 100, the corresponding Gaussian fields are too highly correlated to be simulated using either the circulant or cut-off embedding methods. For this reason, we limited the lognormal simulations to a correlation distance ranging from 5 to 100 with the simple exponential correlation structure (i.e., $\theta_2 = 1$), and the SF for the lognormal case consequently ranges from 0.1 to 2.0. We set the Gaussian mean and variance equal to 0.0 and 3.0, respectively. The corresponding lognormal mean and variance are 4.48 and 383.34. We simulated 2500 Gaussian fields for each scenario using the circulant/cutoff embedding technique and transformed them into lognormal fields.

A.3 Estimation of Spatial Correlation Structure

We used the Pearson's product moment estimator to obtain the correlation of the process $\{Z(\mathrm{u}) : \mathrm{u} \in \mathrm{R}^2\}$ for all u_i and u_j such that $\{\mathrm{u}_i - \mathrm{u}_j = d_t;\ i, j = 1,....,n^2\}$

$$\hat{r}(d_t) = \frac{\langle Z(\mathrm{u}_i)Z(\mathrm{u}_j)\rangle - \langle Z(\mathrm{u}_i)\rangle\langle Z(\mathrm{u}_j)\rangle}{\sqrt{\langle Z(\mathrm{u}_i)^2\rangle - \langle Z(\mathrm{u}_i)\rangle^2} \cdot \sqrt{\langle Z(\mathrm{u}_j)^2\rangle - \langle Z(\mathrm{u}_j)\rangle^2}} \tag{A.7}$$

where the angular brackets represent the sample average. In this study, d_t varied from 0 to approximately 50 units separated by 0.8 units (a total of 62 lags). Depending on the manner in which the sample averages in Equation (A.7) are estimated, we classified the resulting correlations as single-field and ensemble correlations.

Let $n(d_t)$ be the number of distinct pairs separated by distance lag d_t in a single field. The single-field correlation coefficient $\hat{r}_k(d_t)$ for k^{th} field at a certain distance lag d_t is obtained as

$$\hat{r}_k(d_t) = \left. \frac{n(d_t) \cdot \sum_{n(d_t)} Z(\mathrm{u}_i)Z(\mathrm{u}_j) - \sum_{n(d_t)} Z(\mathrm{u}_i) \cdot \sum_{n(d_t)} Z(\mathrm{u}_j)}{\sqrt{n(d_t) \cdot \sum_{n(d_t)} Z(\mathrm{u}_i)^2 - \left[\sum_{n(d_t)} Z(\mathrm{u}_i)\right]^2} \cdot \sqrt{n(d_t) \cdot \sum_{n(d_t)} Z(\mathrm{u}_j)^2 - \left[\sum_{n(d_t)} Z(\mathrm{u}_j)\right]^2}} \right|_k \tag{A.8}$$

where $1 \le k \le 1000$ for the Gaussian case and $1 \le k \le 2500$ for the lognormal case. The average single-field correlation (henceforth referred to as the average correlation) $\hat{r}_{avg}(d_t)$ at a distance lag d_t is then obtained as:

$$\hat{r}_{avg}(d_t) = \frac{1}{N}\sum_{k=1}^{N} \hat{r}_k(d_t) \tag{A.9}$$

Instead of a single realization, one can also estimate correlation at a distance lag d_t by using all the N available fields to compute the sample averages in Equation (A.7). We refer to such correlations as ensemble correlations. Since the simulations are on a regular two-dimensional grid, we obtained the ensemble correlations by replacing the term $n(d_t)$ in Equation (A.8) with $N \cdot n(d_t)$.

The single-field, average and ensemble correlations are estimated for varying distance lags to obtain the corresponding correlation functions. The mean error between the true and the estimated correlation functions is computed for each combination of SF and θ_2 as follows

$$MAE = \frac{1}{62}\sum_{t=1}^{62}\left|r_{true}(d_t) - \hat{r}_{est}(d_t)\right| \tag{A.10}$$

where $r_{true}(d_t)$ is the true correlation at lag d_t and $\hat{r}_{est}(d_t)$ is either an average or an ensemble correlation coefficient for the lag d_t.

In the next section, in addition to quantifying the bias resulting from the Pearson's estimator for various SF and θ_2, we compare correlation estimates with the semivariogram estimates obtained using the classical method-of-moments estimator [*Cressie*, 1993]

$$\hat{\gamma}_k(d_t) = \frac{1}{2\cdot n(d_t)}\cdot\sum_{n(d_t)}\left[Z(u_i) - Z(u_j)\right]^2\Big|_k \tag{A.11}$$

where $\hat{\gamma}_k(d_t)$ is the semivariogram estimate for the k^{th} field at a certain distance lag d_t and $1 \le k \le 1000$ for the Gaussian case and $1 \le k \le 2500$ for the lognormal case. The average and ensemble semivariogram estimates for lag d_t are then obtained in the same manner as their correlation counterparts. Upon repeating the process for varying d_t, we obtain the single-field, average and ensemble semivariograms.

A.4 Results and Discussion

A.4.1 Gaussian simulations

We begin this section with the estimation results for a simple exponential correlation function (θ_2 = 1.0) with varying SF. In Figure A.3, we show the single-field, the average and the true correlation structures for four different SF of 0.1, 0.5, 1.0 and 2.0. There is a clear systematic bias in the estimation of the single-field correlation

structure. The spread in the single-field correlation estimates, shown in Figure A.3 in terms of the 5th and 95th quantiles, increases with increasing distance. We also noticed that the bias in the estimation of the ensemble correlation function (not shown in the Figure A.3) is negligible compared to the bias in the single-field correlation function. The estimation bias and the quantiles for a particular lag cannot be obtained using Fisher's exact sampling distribution theory [*Johnson et al.,* 1995] of correlations as the pairs of pixels that correspond to d_t are not mutually independent due to the presence of spatial correlation.

The estimation bias quantified in terms of MAE (Equation 9) is plotted in Figure A.4 for all of the SF. It increases rapidly for the SF between 0.1 and 1.0, and its rate of increase with the scale factor is very small beyond an SF of 4.0. Therefore, the size of the domain has a significant impact on the estimation of the correlation structure. The domain size should be at least 5 times or larger (SF less than 0.2 in Figure A.4) than the correlation distance of the underlying process for the sampling effects to be within reasonable limits. To further investigate the effect of SF on the estimation of the correlation structure, we computed the mean error (ME) separately for each of the terms in Equation 7 (Figure A.5). We selected the ME for the error budget analysis as the signs are preserved in the arithmetic operations. From Figure A.5, it is apparent that the non-centered covariance is unbiased. Both the numerator (which is the centered covariance) and the denominator in Equation 7 are biased and dependent on the SF. It is also apparent from Figure A.5 that bias is due to the behavior of the product of means, $\langle Z(\mathbf{u}_1)\rangle\langle Z(\mathbf{u}_2)\rangle$ in the covariance term and $\langle Z(\mathrm{u}_1)\rangle\langle Z(\mathrm{u}_1)\rangle$ and $\langle Z(\mathrm{u}_2)\rangle\langle Z(\mathrm{u}_2)\rangle$ in the denominator of Equation 7. This result is further justified by Figure A.6, which shows the mean and variance from each of 1000 realizations for an SF of 0.1 and 2.0 and θ_2 equal to 1.0. Even if the sample mean is unbiased, its scatter around the true value increases with the scale factor, resulting in the biased $\langle Z(\mathrm{u})\rangle^2$ and, subsequently, the

biased variance (Figure A.6). It should be noted that the bias discussed so far is for the simple exponential correlation function ($\theta_2 = 1.0$).

The bias resulting from varying SF (0.3, 0.6, 1.0 and 2.0) and θ_2 (0.1 to 2.0) is summarized in Figure A.7. However, for the scale factor of 2.0, we have the ME only for the shape parameter up to 1.0 because realizations cannot be simulated using the circulant/cut-off embedding method. For an SF equal to 0.3, its relative contribution to the overall bias is small, and the MAE is mainly due to the effect of smoothness of the process; it decreases with increasing θ_2. As the SF increases, it dominates the estimation bias, resulting in a bulge in the MAE function that can be noticed clearly for SF equal to 1.0 and 2.0. Again, the reason for the dependence of MAE on the shape parameter is the same as that of the aforementioned SF.

A.4.2 Lognormal simulations

Figure A.7 summarizes the effect of SF on the estimation of the correlation structure for lognormal RF. As mentioned earlier, the SF is limited to 2.0, and the number of realizations is 2500. A comparison of Figures A.4 and A.8 reveals that the bias for the lognormal case is similar to and even slightly smaller than the bias for the Gaussian case. Because of the skewness, one would expect that the bias in the case of lognormal simulations is larger than that for the Gaussian simulations. This contradiction can be explained from the results of *Habib and Krajewski* [2001]. Based on simulations from a bivariate mixed lognormal distribution, they reported that Pearson's correlation coefficient is overestimated, and that the bias depends on the skewness of the lognormal distribution. Figure A.9 shows the effect of skewness, which is in terms of the variance of the underlying Gaussian variable, on the estimation of Pearson's correlation coefficient. For the lognormal 2D simulations (with the Gaussian variance of 3.0), the bias in the Pearson's estimate is sometimes as large as 0.25 for true correlation equal to

0.4. Due to the limited domain size, this overestimation might have compensated for the bias.

A.4.3 Sensitivity to the number of realizations

We performed a sensitivity analysis to investigate the behavior of the bias as a function of the number of realizations available for the estimation of the ensemble correlation structure. From Figure A.10(a), it is clear that the bias for the Gaussian fields drops rapidly and is negligible for approximately N=50. That is, about 50 realizations are sufficient to estimate the product $\langle Z(\mathrm{u}_1)\rangle\langle Z(\mathrm{u}_2)\rangle$ correctly, resulting in unbiased correlation estimates. Even for the lognormal simulations, the bias drops rapidly after few realizations (Figure A.10(b)). However, the convergence for lognormal simulations is slower than in the Gaussian case.

A.4.4 Estimation of semivariogram

In . A11, we show the semivariogram estimates for the four cases considered in Figure A.3. Comparing the average semivariogram estimates with that of the true semivariogram, the size of the estimation domain has no effect on the single-field semivariogram estimates in terms of bias. The unbiasedness is expected because, unlike the correlation function, the semivariogram does not require the estimation of mean. Comparing Figure A.11 with Figure A.3, the distribution of single-field semivariogram estimates for a particular lag is skewed compared to that of correlation estimates.

Since the semivariogram estimates are unbiased, one might be tempted to use the following one-one relationship to obtain the covariance and, subsequently, the correlations.

$$\gamma(h) = \{C(0) - C(h)\} \tag{A12}$$

However, this is strictly a theoretical relationship and is not valid for the estimators. *Cressie* [1993] suggests that the relationship can be extended to the

estimators if the ratio of the number of pairs for a distance lag to the total number tends to 1. We nevertheless obtained correlations from the semivariogram using the above relation, and they would still be biased as the variance used in the normalization is biased. Figure A.12 shows the correlations obtained from semivariogram estimates, and it is apparent that not only are the correlations biased, but they are also not bounded between -1 and 1. On the other hand, the correlations estimated by Pearson's product moment estimator are by definition always bounded between -1 and 1.

A.5. Summary and Conclusions

In this study, we focused on the bias in the estimation of the spatial correlation structure of Gaussian and lognormal processes for a fixed domain size. Selection of a smaller domain size results in severe underestimation, and this bias increases as the correlation distance increases relative to the domain size. Our results confirm earlier findings of *Krajewski and Duffy* [1988] and *Dubin* [1994]. We assert that one should exercise caution when estimating the correlation structure from a single realization, and as a rule of thumb, the domain size should be at least 5 times larger than the correlation distance for the sampling problems to be within the reasonable limits. The study also showed that the higher the variability of the process at the small scales, the higher the bias in the estimation of correlation structure.

If multiple realizations are used in the estimation, a sample size of around 50 is sufficient to have a negligible bias for both normal and lognormal fields. This requirement can be easily met in such fields as radar hydrology and meteorology, given the availability of remotely sensed observation of frequently reoccurring processes of interest, e.g. rainfall. However, in such areas as ground water hydrology, obtaining 50 independent realizations may not be feasible. In that case, our results provide information about the bias that would affect the estimation of correlation structure. The study showed that the bias we see in the estimation of correlation structure comes from

the bias in the estimation of the product of means. Though it is well documented that the biasedness in correlation comes from the estimation of the mean, this study gives a quantitative idea of the size of the two-dimensional domain, the role of the small scale variability of the process and the number of realizations required to obtain an unbiased correlation function. The study also shows that the size of the estimation domain has no effect on the estimation of the semivariogram. However, the correlations obtained by transforming the semivariogram estimates are still biased and tend to go beyond -1, which is unrealistic.

Table A.1* Parameters of the correlation function (Equation 3.1) of Gaussian random fields required to obtain lognormal random fields with correlation distance indicated in column 1 for a fixed shape parameter of 1.0.

Correlation Distance Lognormal	Correlation Distance Gaussian	Shape Parameter Gaussian	RMSE
5	11.0	1.26	0.0105
10	22.5	1.23	0.0087
15	33.5	1.24	0.0087
20	45.0	1.23	0.0087
25	56.0	1.24	0.0085
30	67.0	1.24	0.0087
35	78.5	1.24	0.0086
40	89.5	1.24	0.0086
45	101.0	1.24	0.0085
50	112.0	1.24	0.0085
60	134.5	1.24	0.0086
70	157.0	1.24	0.0086
75	168.0	1.24	0.0085
80	179.0	1.24	0.0086
90	201.5	1.24	0.0085
100	224.0	1.24	0.0085

*The parameters are obtained by a systematic search of the parametric space minimizing the RMSE.

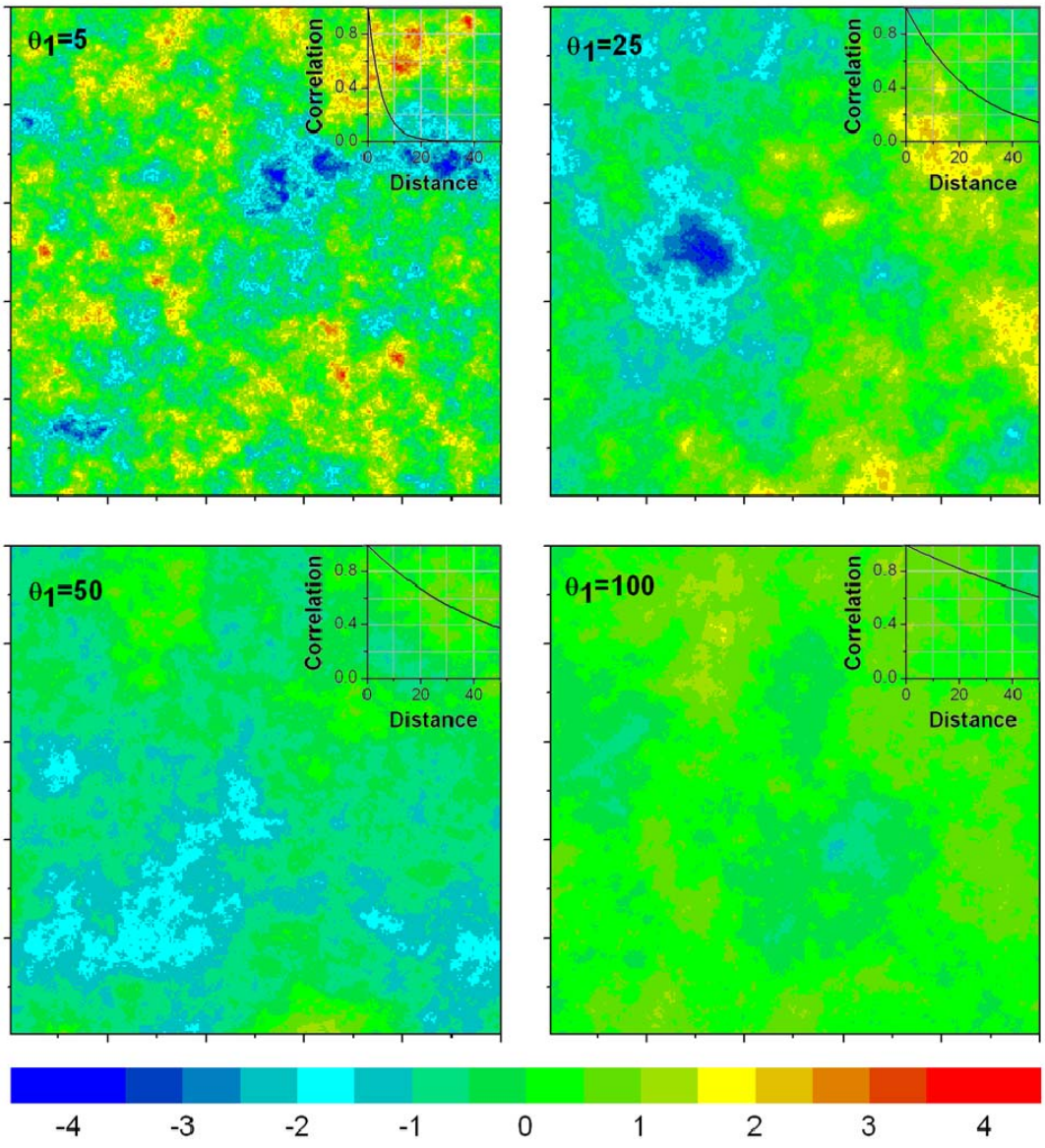

Figure A.1 Sample realizations simulated using the circulant embedding technique with the exponential correlation function (Equation A.3). The shape parameter (θ_2) is fixed at 1.0, and the correlation distance (θi) is indicated on each field. The size of the simulation domain is 100 (arbitrary length units), and the resolution is 0.4. The corresponding correlation functions are shown in the insets.

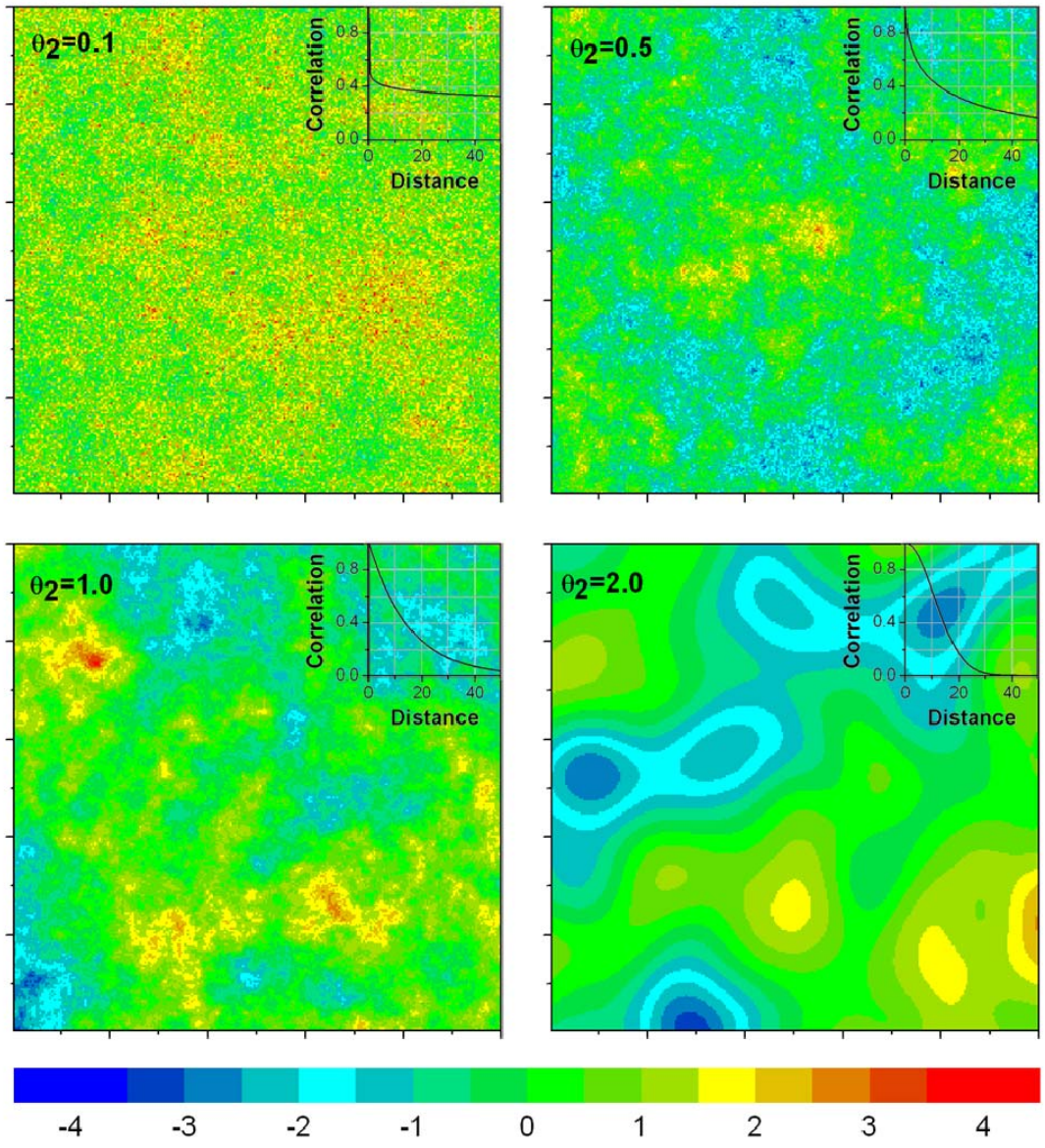

Figure A.2 Sample realizations simulated using the circulant embedding technique with the exponential correlation function (Equation A.3). The shape parameter (θ_2) is fixed at 1.0, and the correlation distance (θi) is indicated on each field. The size of the simulation domain is 100 (arbitrary length units), and the resolution is 0.4. The corresponding correlation functions are shown in the insets.

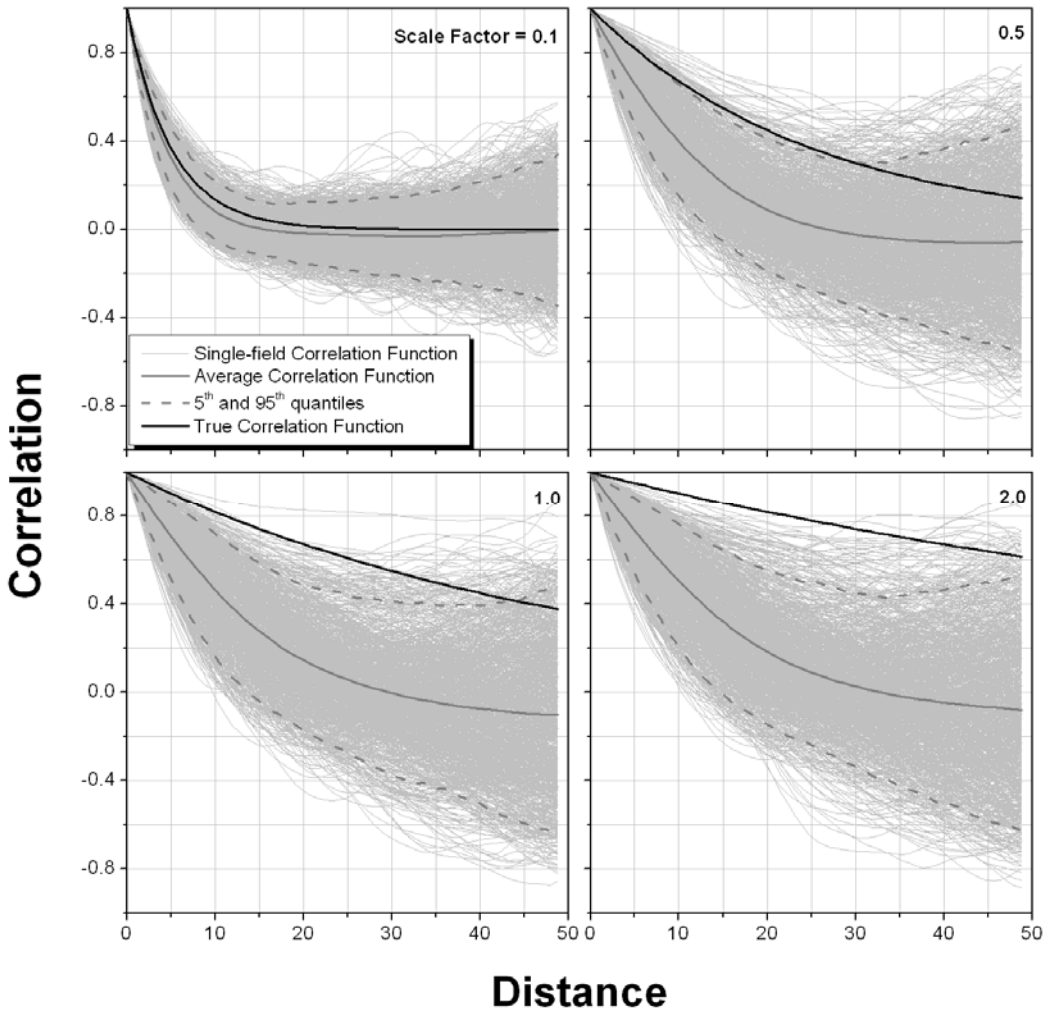

Figure A.3 Comparison of correlation estimates of Gaussian random fields for various scale factors. The shape parameter (θ_2 in Equation A.3) is equal to 1.0, and the number of realizations is 1000. The shaded area represents two-standard deviation of the correlation function.

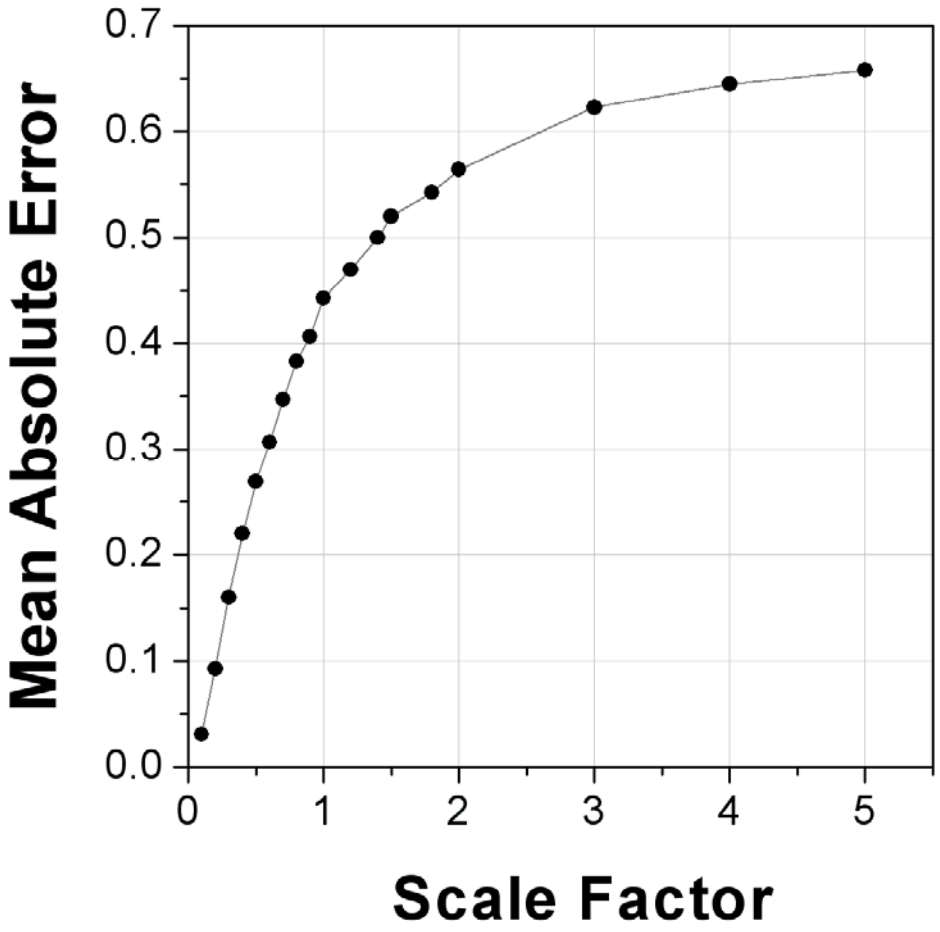

Figure A.4 Effect of scale factor on the estimation of the correlation structure using Pearson's estimator (Equation A.7).

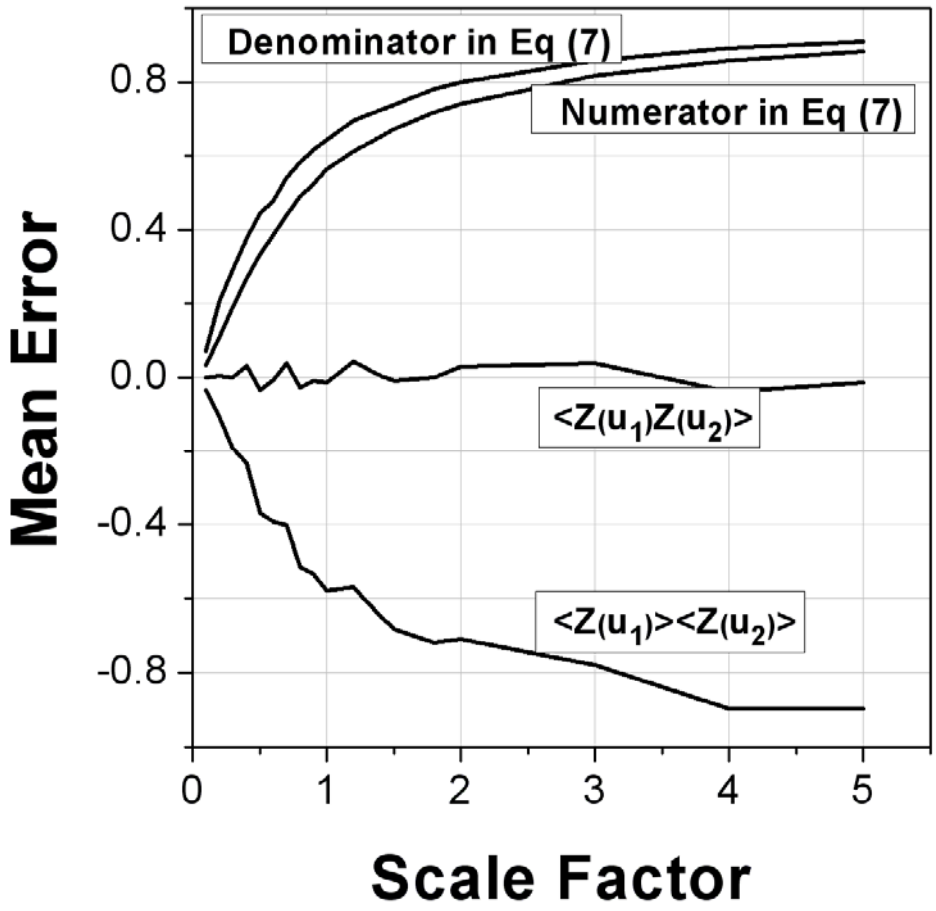

Figure A.5 Effect of the scale factor on the estimation of each component of Equation A.7 in terms of mean error. Angular brackets indicate the sample average for each realization.

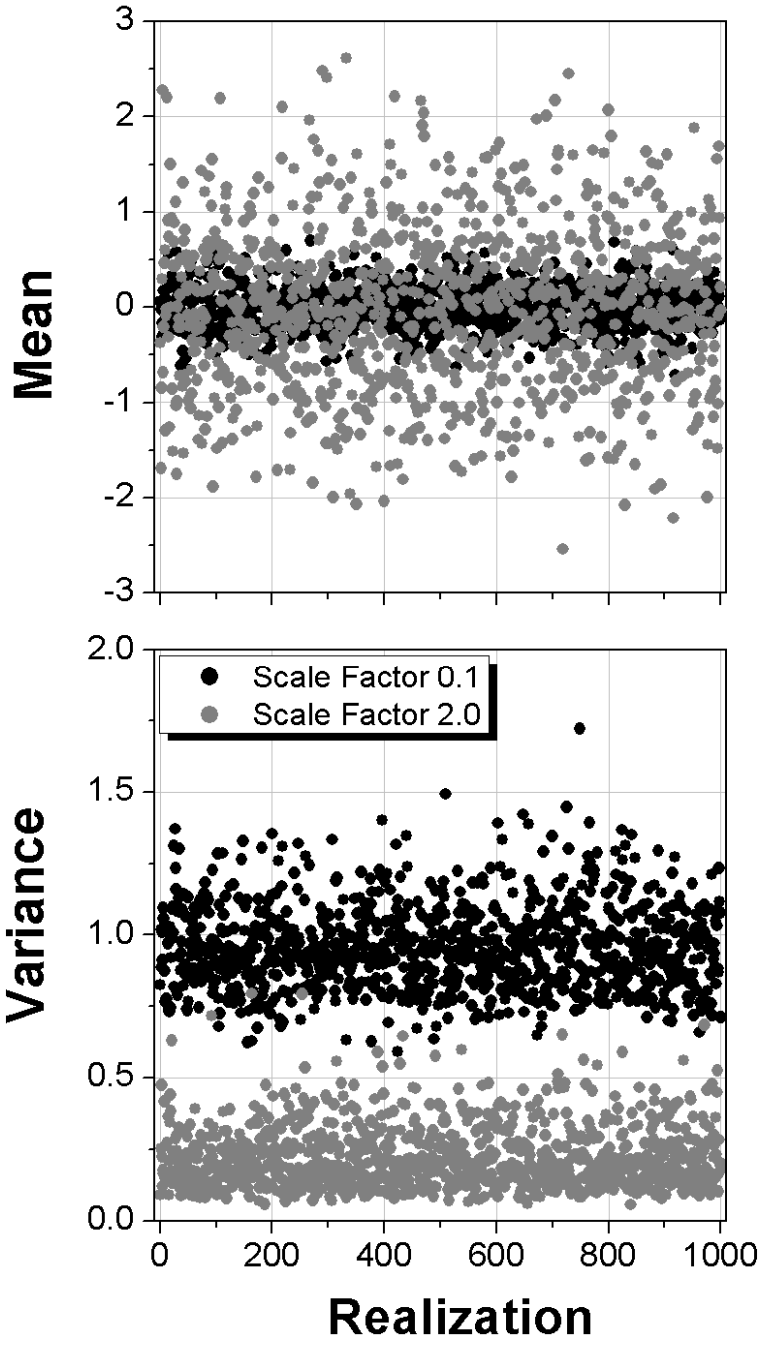

Figure A.6 Estimates of mean and variance for each realization of the Gaussian random fields with the correlation distances of 5 and 100 and the shape parameter of 1.0. Corresponding scale factors are indicated on the Figure.

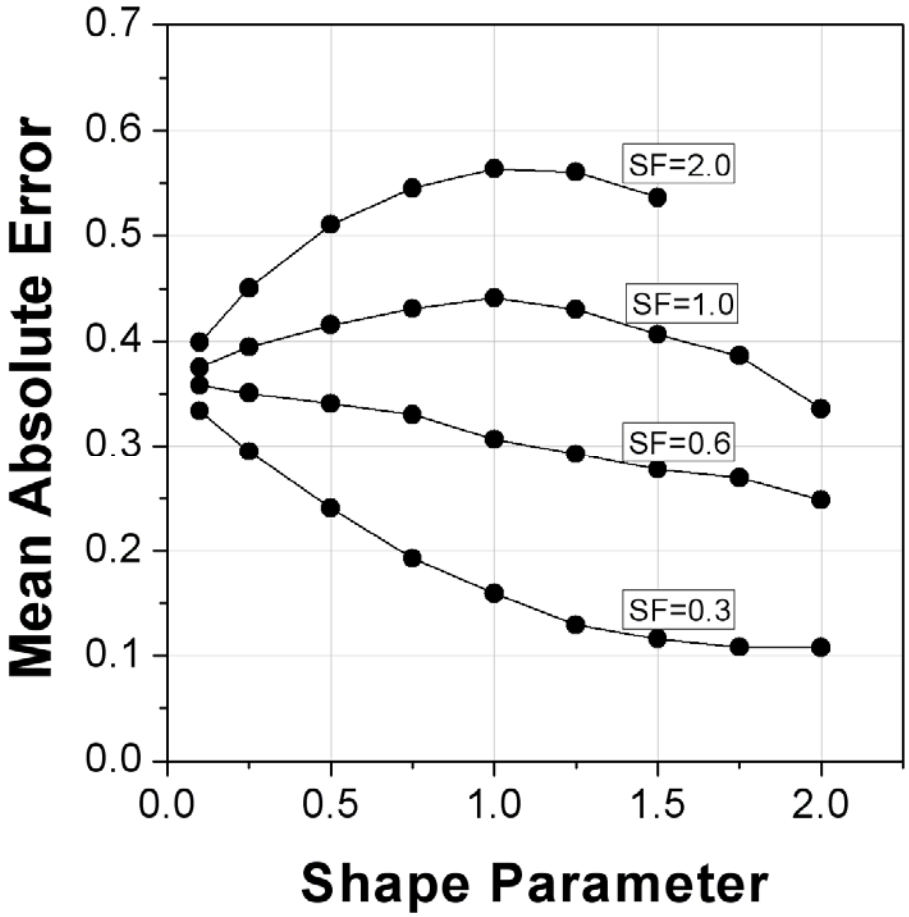

Figure A.7 Effect of shape parameter on the estimation of the correlation structure of Gaussian random fields with a fixed correlation distance (Corresponding scale factor is shown on each curve). The missing points for the scale factor of 2.0 are because the realizations cannot be simulated using the circulant/cut-off embedding method. Number of realizations for each case is equal to 1000.

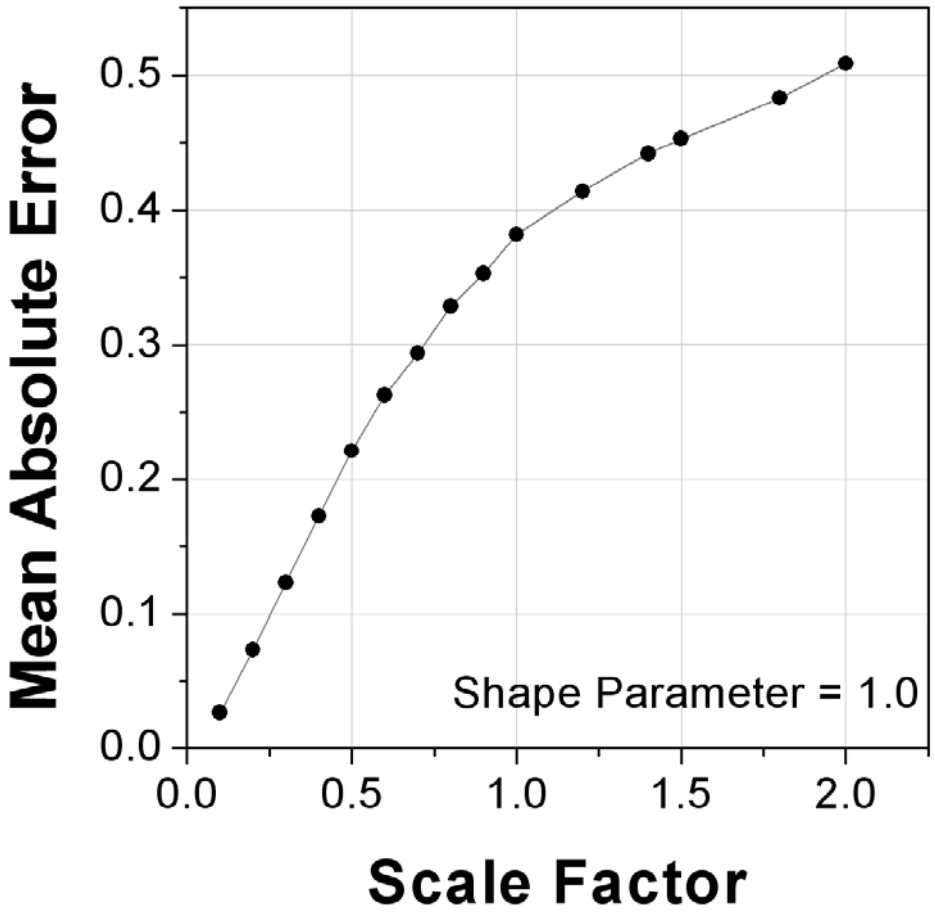

Figure A.8 Effect of the scale factor on the estimation of the correlation structure for the lognormal random fields. The shape parameter is fixed at 1.0.

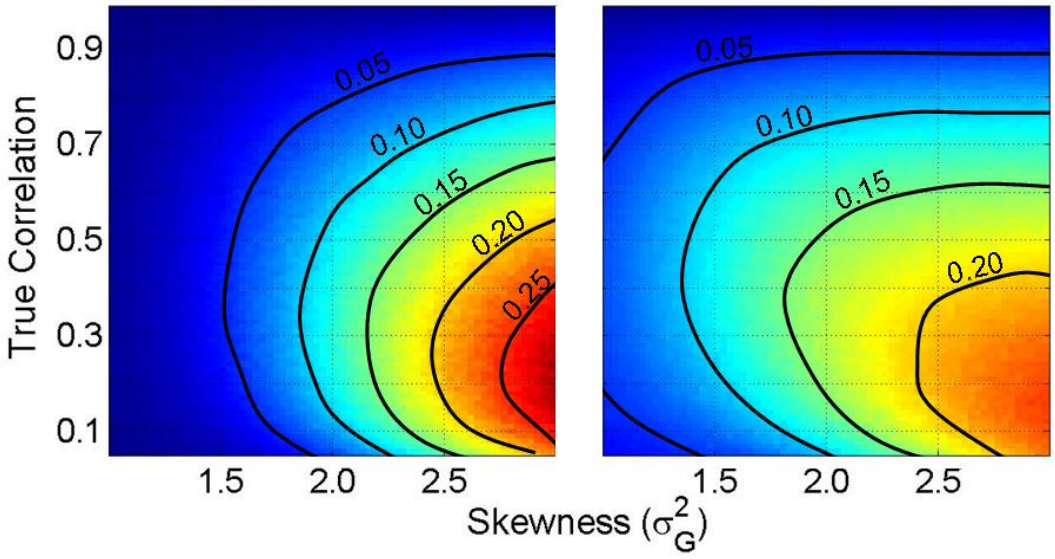

Figure A.9 Bias (left panel) and standard deviation (right panel) in the estimation of Pearson's correlation coefficient for the lognormal data. The size of the sample is equal to 1000, and the number of simulations is equal to 5000.

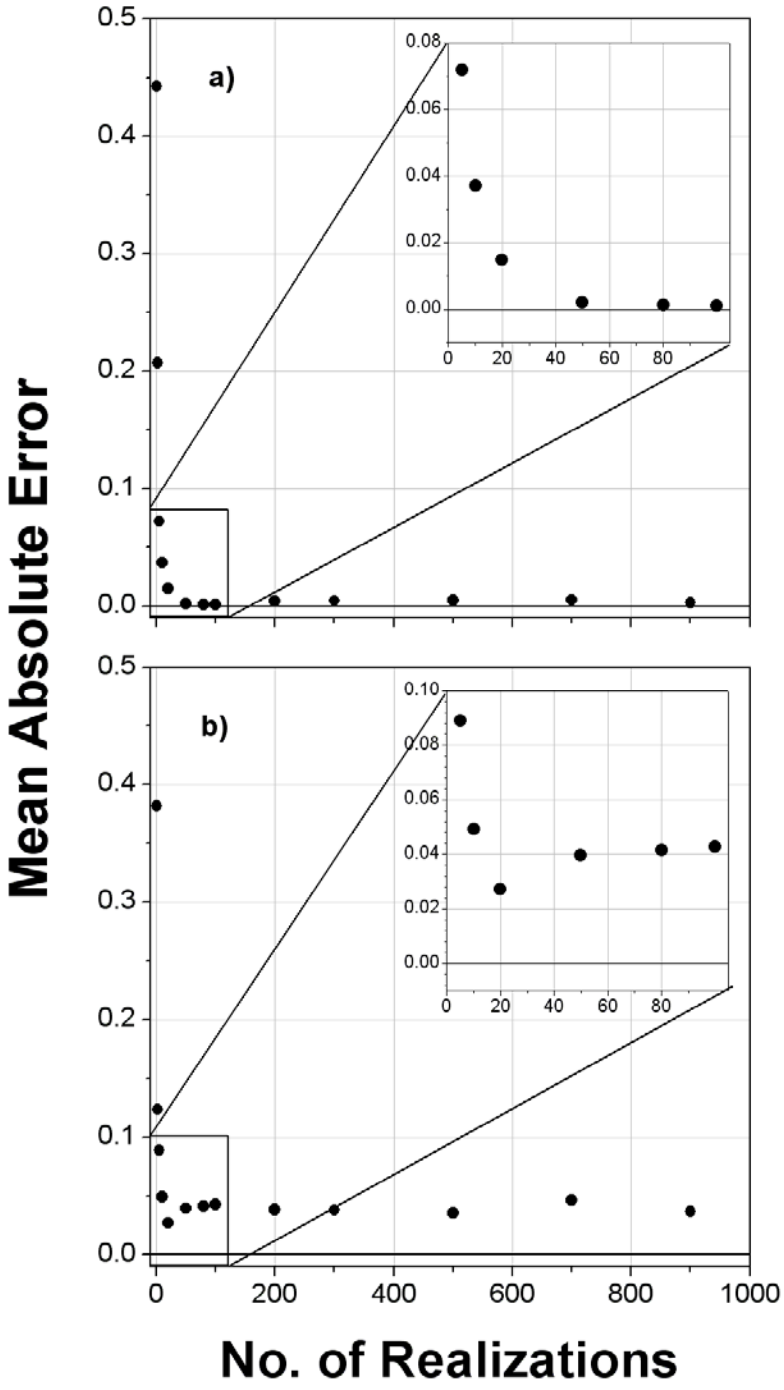

Figure A.10 Mean error in the estimation of the ensemble correlation structure as a function of number of realizations for the a) Gaussian and b) lognormal realizations. The scale and shape parameters are fixed at 1.0.

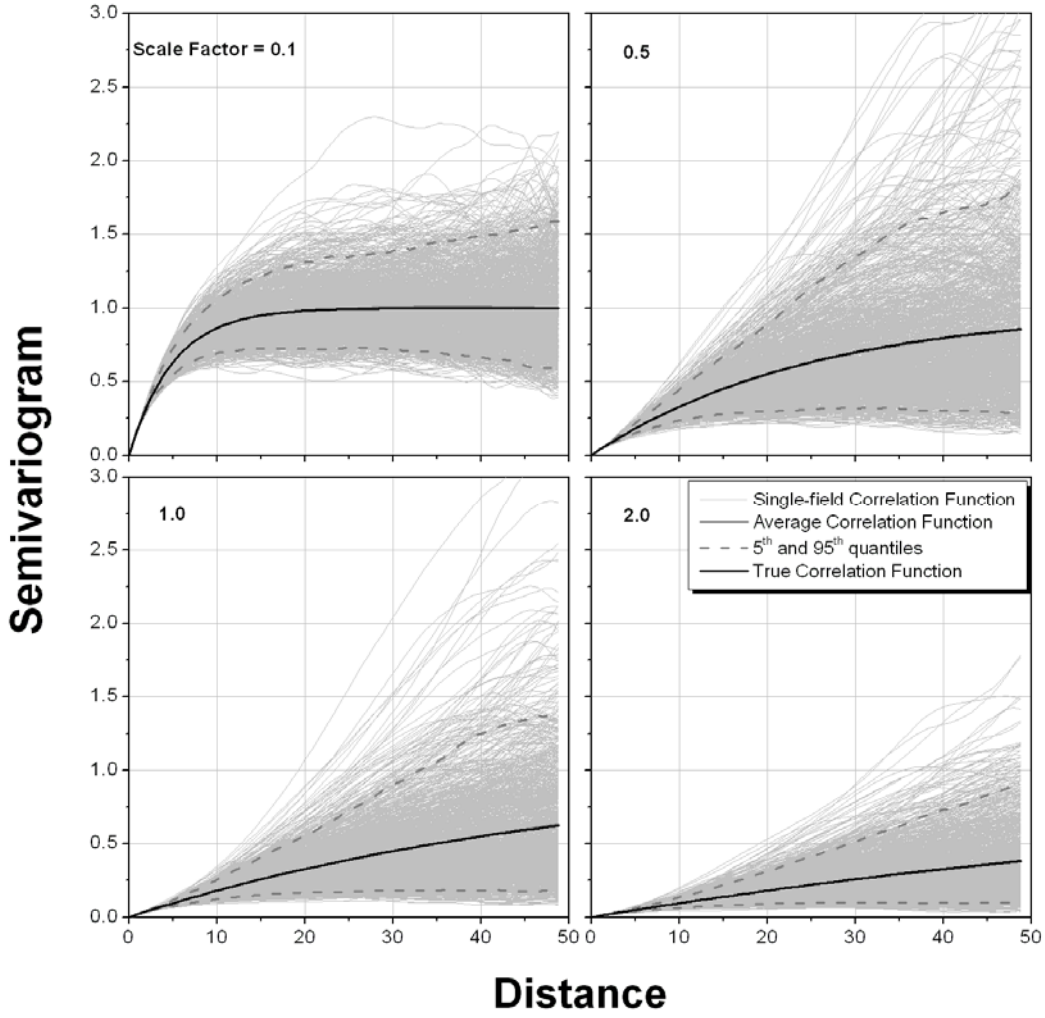

Figure A.11 Comparison of semivariogram estimates of Gaussian random fields for various scale factors. The correlation structure is the same as the one used in Figure A.3, and the number of realizations is 1000.

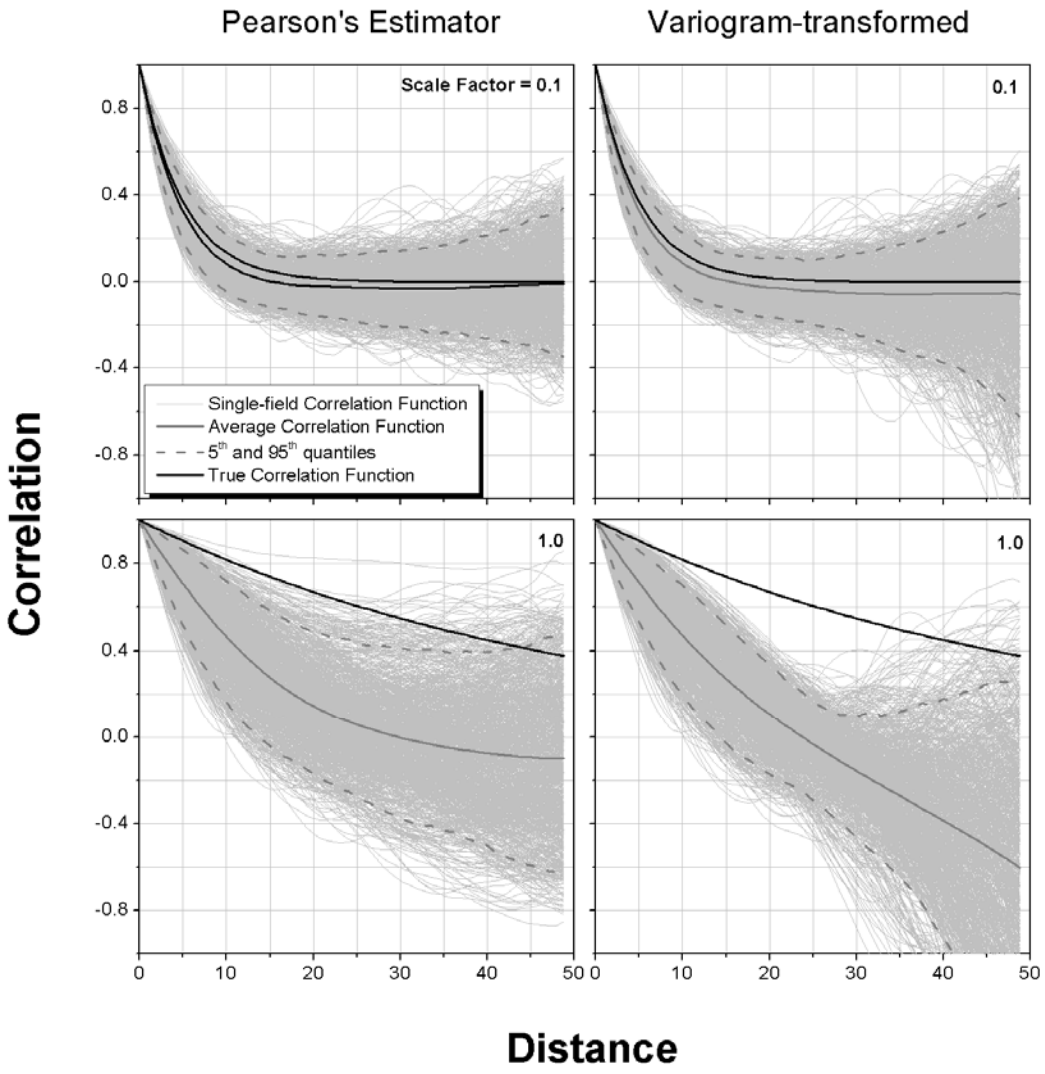

Figure A.12 Comparison of correlations obtained from Pearson's estimator and from a one-to-one relation between variogram and covariance for Gaussian random fields with an exponential correlation structure. The top panels are for the scale factor of 0.1, and the bottom panels are for the scale factor of 1.0.

REFERENCES

Abbott, M. B., J. C. Bathurst, J. A. Cunge, P. E. O'Connell, and J. Rasmussen (1986a), An introduction to the European Hydrological System–Syst`eme Hydrologique Europ´een, 'SHE' 1. History and philosophy of a physically-based, distributed modeling system, *Journal of Hydrology*, 87, 45–59.

Abbott, M. B., J. C. Bathurst, J. A. Cunge, P. E. O'Connell, and J. Rasmussen (1986b), An introduction to the European Hydrological System–Syst´eme Hydrologique European, 'SHE' 2. Structure of physically-based, distributed modeling system, *Journal of Hydrology*, 87, 61–77.

Alonso, C. V. and R. L. Binger (2000), Goodwin Creek experimental watershed: a unique field laboratory. *Journal of Hydraulic Engineering*, 126, 174-177.

Aitchison, J., and J. A. C. Brown (1957), The lognormal distribution with special reference to its uses in economics. University of Cambridge, *Department of Applied Economics, Monograph: 5*, Cambridge University Press, pp. 176.

Allen, P. B., and J. W, Naney (1991), Hydrology of the Little Washita River Watershed, Oklahoma: Data, and Analyses, *United States Department of Agriculture, Agricultural Research Service, ARS-90*, pp. 73.

Anagnostou, E. N., W. F. Krajewski, and J. A. Smith (1999), Uncertainty quantification of mean-areal radar-rainfall estimates, *Journal of Atmospheric and Oceanic Technology*, 1999, 16, 206-215.

Anagnostou, E. N., W. F. Krajewski, D.-J. Seo and E.R. Johnson (1998), Mean-field rainfall bias studies for WSR-88D, *Journal of Hydrologic Engineering*, 3, 149-159.

Anderson, M. G. and T. P Burt (1990), *Process Studies in Hillslope Hydrology*, Wiley, Chichester, 539 pp.

Arnaud, P., C. Bouvier, L. Cisneros, and R. Dominguez (2002), Influence of rainfall spatial variability on flood prediction, *Journal of Hydrology*, 260, 216 – 230.

Ashley, S. T., and W. S. Ashley (2008), Flood fatalities in the United States, *Journal of Applied Meteorology and Climatology*, 47, 805-818.

Austin, P. M (1987), Relation between measured radar reflectivity and surface rainfall, *Monthly Weather Review*, 115, 1053-1070.

Bachhi, B., and N. T. Kottegoda (1995), Identification And Calibration Of Spatial Correlation Patterns Of Rainfall, *Journal of Hydrology*, 165, 311-348.

Bell, T. L. (1987), A space-time stochastic model of rainfall for satellite remote sensing studies. *Journal of Geophysical Research*, 92, 9631-9643.

Bell, T. L., A. Abdullah, R. L. Martin and G. R. North (1990), Sampling errors for satellite-derived tropical rainfall—Monte Carlo study using a space–time stochastic model, *Journal of Geophysical Research*, 95, 2195–2205.

Benson, M. A. (1962), Evolution of methods for evaluating the occurrence of floods, *U.S. Geological Survey Water-Supply Paper 1580-A*, Washington, D. C., 30 pp.

Berndtsson, R. (1988), Temporal variability in the spatial correlation of daily rainfall, *Water Resources Research*, 24, 1511-1517.

Beven, K. J (1985), Distributed models, in M.G. Anderson and T.P. Burt (Eds) *Hydrological forecasting*, John Wiley and Sons, New York, pp 405-435

Beven, K.J (2001), *Rainfall-Runoff modeling - The Primer*, Wiley, Chichester, 356 pp.

Beven, K. J., and M. J. Kirby (1979), A physically-based variable contributing area model of basin hydrology, *Hydrology Science Bulletin* 24, 43–69.

Beven, K. J., and G. M. Hornberger (1982), Assessing the effect of spatial pattern of precipitation in modeling streamflow hydrographs, *Water Resources Bulletin*, 18, 823-829.

Bhaskar, R. (1989), *Reclaiming reality: A critical introduction to contemporary philosophy,* Verso, London.

Bloschl, G., and M. Sivapalan (1997), Process controls on regional flood frequency: Coefficient of variation and basin scale, *Water Resources Research*, 1997, 33, 2967-2980.

Borga, M (2002), Accuracy of radar rainfall estimates for streamflow simulation. *Journal of Hydrology*, 267, 26-39.

Borga, M., and F. Tonelli (2000),. Adjustment of range dependent bias in radar rainfall estimates. *Journal of Physics and Chemistry of the Earth*, 10, 909-914.

Bormann, H., L. Breuer, S. Giertz, J. A. Huisman, N. R. Viney (2009), Spatially explicit versus lumped models in catchment hydrology – experiences from two case studies, *Uncertainties in Environmental Modelling and Consequences for Policy Making,* P.C. Baveye et al. (eds.), Springer Science + Business Media B.V., pp 3-26.

Bras, R. L., and I. Rodriguez-Iturbe (1993), *Random Functions and Hydrology*, 559 pp, Dover Publications INC.

Brath, A., and A. Montaniri (2003), Sensitivity of the peak flows to the spatial variability of the soil infiltration capacity for different climatic scenarios, *Physics and Chemistry of the Earth*, 28, 247–254.

Brath, A. A. Montanari and E. Toth (2004), Analysis of the effects of different scenarios of historical data availability on the calibration of a spatially-distributed hydrological model, *Journal of Hydrology*. 291, 232–253

Burnash, R. J. C., R. L. Ferral, R. A. McGuire (1973). A General Streamflow Simulation System - Conceptual Modeling for Digital Computers, *Report by the Joint Federal State River Forecasts Center*, Sacramento, California.

Carpenter, T. M., and K. P. Georgakakos (2004), Impacts of parametric and radar rainfall uncertainty on the ensemble streamflow simulations of a distributed hydrologic model. *Journal of Hydrology*, 298, 202-221.

Carpenter, T. M., and K.P. Georgakakos (2006), Discretization scale dependencies of the ensemble flow range versus catchment area relationship in distributed hydrologic modeling. *Journal of Hydrology*, 328, 242-257.

Chumchean, S., A. Sharma, and A. Seed (2003), Radar rainfall error variance and its impact on the radar rainfall calibration, *Journal of Physics and Chemistry of the Earth*, 28, 27-39.

Ciach, G. J., and W. F. Krajewski (1999a), On the estimation of radar-rainfall error variance, *Advances in Water Resources*, 22, 585-595.

Ciach, G. J. and W. F. Krajewski (1999b), Radar-rain gauge comparisons under observational uncertainties. *Journal of Applied Meteorology*, 38, 1519-1525.

Ciach, G. J., M. L. Morrissey, and W. F. Krajewski (2000), Conditional bias in radar rainfall estimation, *Journal of Applied Meteorology*, 39, 1941-1946.

Ciach, G. J., E. Habib, and W. F. Krajewski (2003), Zero-covariance hypothesis in the error variance separation method of radar rainfall verification, *Advances in Water Resources*, 26, 573-580.

Ciach, G.J., and W.F. Krajewski (2006), Analysis and modeling of spatial correlation structure in small-scale rainfall in central Oklahoma, *Advances in Water Resources*, 29, 1450-1463.

Ciach, G.J., W. F. Krajewski, and G. Villarini (2007), Product-Error driven uncertainty model for probabilistic quantitative precipitation estimation with NEXRAD data, *Journal of Hydrometeorology*, 8, 1325-1347.

Crawford, N. H., and R. K. Linsley (1966), Digital simulation in hydrology: Stanford Watershed Model IV, *Stanford University, Palo Alto, California., Tech. Rep. No. 39*.

Cressie, N.A.C (1993), *Statistics for Spatial Data*, 462 pp, Wiley and Sons.

Cressie, N. A. C. and D. M. Hawkins (1980), Robust estimation of the variogram, I. *Journal of the International Association of Mathematical Geology*, 12, 115-125.

Cressie, N., and M. O. Grondona (1992), A comparison of variogram estimation with covariogram estimation, *The Art of Statistical Science,* Ed. K.V. Mardia, 191-208.

Crum, T. D., and R. L. Alberty (1993), The WSR-88D and the WSR-88D operational support facility, *Bulletin of the American Meteorological Society*, 74, 1669-1687.

Davis, M.W. (1987), Production of conditional simulations via the LU triangular decomposition of the covariance matrix, *Mathematical Geology*, 19, 91-98.

Dawdy, D. R. (1961), Variation of flood ratios with size of drainage area, *US Geological Survey professional paper*, 424-C, C36.

Dietrich, C.R. (1993), Computationally efficient Cholesky factorization of a covariance matrix with block Toeplitz structure, *Journal of Statistical Computation and Simulation*, 45, 203-218.

Dietrich, C.R., and G.N. Newsam (1993), A fast and exact method for multidimensional Gaussian stochastic simulations, *Water Resources Research*, 29, 2861-2869.

Dietrich, C.R., and G.N. Newsam (1997), Fast and exact simulation of stationary Gaussian processes through circulant embedding of covariance matrix, *SIAM Journal of Scientific Computing*, 18, 1088-1107.

Di Lazarro (2009), Regional analysis of storm hydrographs in the rescaled width function framework, *Journal of Hydrology*, 373, 352-365.

D'Odorico, P., and R. Rigon (2003), Hillslope and channel contributions to the hydrologic response, *Water Resources Research*, 39, 1113, doi:10.1029/2002WR001708.

Dubin, R.A. (1994), Estimating correlograms: A Monte Carlo study*, Geographical Systems*, 1, 189-202.

Eaton, B., M. Church, and D. Ham (2002), Scaling and regionalization of flood flows in British Columbia, Canada*, Hydrologic Processes,* 16, 3245-3263.

Fabry, F., G. L. Austin, and D. Tees (1992), The accuracy of rainfall estimates by radar as a function of range*, Quarterly Journal of Royal Meteorological Society*, 118, 435-453.

Finnerty, B. D., M. B. Smith, D.-J. Seo, V. Koren, and G. E. Moglen (1997), Space-time scale sensitivity of the Sacramento model to radar-gage precipitation inputs, *Journal of Hydrology*, 203, 21-38.

Freeze, R. A., Harlan, R. L. (1969), Blueprint for a physically-based, digitally-simulated hydrologic response model, *Journal of Hydrology*, 9, 237-258.

Fulton, R. A., J. P. Breidenbach, D.-J. Seo, and D. A. Miller (1998), The WSR-88D Rainfall Algorithm, *Weather and Forecasting*, 13, 377-395.

Furey, P. R., and V. K. Gupta (2005), Effects of excess rainfall on the temporal variability of observed peak-discharge power laws. *Advances in Water Resources*, 28, 1240-1253.

Furey, P. R., and V. K. Gupta (2007), Diagnosing peak-discharge power laws observed in rainfall–runoff events in Goodwin Creek experimental watershed, *Advances in Water Resources*, 30, 2387-2399.

Furey, P. R., and B. M. Troutman (2008), A consistent framework for Horton regression statistics that leads to a modified Hack's law, *Geomorphology*, 102, 603-614.

Gagnon, J.-S., S. Lovejoy, and D. Schertzer (2006), Multifractal earth topography, *Nonlinear Processes in Geophysics*, 13, 541–570.

Gebremichael, M., and W.F. Krajewski (2004), Assessment of statistical characterization of small scale variability from radar: analysis of TRMM ground validation datasets, *Journal of Applied Meteorology*, 43, 1180-1199.

Gebremichael, M., W. F. Krajewski, M. L. Morrissey, D, Langerud, G. J. Huffman, and R. Adler (2003), Error uncertainty analysis of GPCP monthly rainfall products: A data based simulation study, *Journal of Applied Meteorology*, 42, 1837-1848.

Gebremichael, M., W. F. Krajewski, M. Morrissey, G. Huffman, and R. Adler (2005), A detailed evaluation of GPCP one-degree daily rainfall estimates over the Mississippi River Basin, *Journal of Applied Meteorology*, 44, 665–681.

Gebremichael, M., W. F. Krajewski, T. M. Over, Y. N. Takayabu, P. Arkin, M. Takayabu (2008), Scaling of tropical rainfall as observed by TRMM precipitation radar, *Atmospheric Research*, 88, 337-354.

Georgakakos, K. P., and T. M. Carpenter (2003), A methodology for assessing the utility of distributed model forecast applications in an operational environment, In: Tachikawa, Y., Vieux, B.E., Georgakakos, K.P., Nakakita, E. (Eds.). *Weather Radar Information and Distributed Hydrological Modeling*. IAHS publications No. 282. IAHS Press, 85-92.

Germann, U., and J. Joss (2001), Variograms of radar reflectivity to describe the spatial continuity of alpine precipitation, *Journal of Applied Meteorology*, 40, 1042-1059.

Germann, U., M. Berenguer, D. Sempere-Torres, and M. Zappa, M. (2009), REAL – ensemble radar precipitation for hydrology in a mountainous region, *Quarterly Journal of the Royal Meteorological Society*, 135, 445–456.

Gneiting, T. (2002), Compactly supported correlation functions, *Journal of Multivariate Analysis*, 83, 493-508.

Gneiting, T., H. Sevcikova, D. B. Percival, M. Schlather, and J. Yindeng (2005), Fast and exact simulation of large Gaussian lattice systems in R2: Exploring the limits, *Journal of Computational and Graphical Statistics*, 15, 483-501.

Goodrich, D.C., L. J. Lane, R. M. Shillito, and S. N. Miller (1997), Linearity of basin response as a function of scale in a semiarid watershed. *Water Resources Research*, 33, 2951-2965.

Goodrich, D. C., T. O. Keefer, C. L. Unkrich, M. H. Nichols, H. B. Osborn, J. J. Stone, and J. R. Smith (2008), Long-term precipitation database, Walnut Gulch Experimental Watershed, Arizona, United States. *Water Resources. Research*, 44, W05S04, doi:10.1029/2006WR005782.

Grayson, R. B., I. D. Moore, and T. A. McMahon (1992), Physically based hydrologic modeling. 2. Is the concept realistic? *Water Resources Research*, 28, 2659–2666.

Guillot, G., and T. Lebel (1999), Disaggregation of Sahelian mesoscale convective system rain fields: Further developments and validation, *Journal of Geophysical Research*, 1999, 104: 31533-31551.

Gupta, V. K (2004), Emergence of statistical scaling in floods on channel networks from complex runoff dynamics, *Chaos, Solitons and Fractals*, 19, 357-365.

Gupta, V. K., and D. R. Dawdy (1995), Physical interpretation of regional variations in the scaling exponents of flood peaks, *Hydrologic Processes*, 9, 347-361.

Gupta, V. K., and E. C. Waymire (1993), A statistical analysis of mesoscale rainfall as a random cascade, *Journal of Applied Meteorology*, 32, 251-267.

Gupta, V. K., and E. Waymire (1998), Spatial variability and scale invariance in hydrologic regionalization. *Scale Dependence and Scale Invariance in Hydrology*, G. Sposito (Ed.), 88-135.

Gupta, V. K., Mesa, O. J., and D. R. Dawdy (1994), Multiscaling theory of flood peaks: Regional quantile analysis, *Water Resources Research*, 30, 3405-3421.

Gupta, V. K., Castro, S. L., and T. M. Over (1996), On Scaling exponents of spatial peak flows from rainfall and river network geometry, *Journal of Hydrology*, 187, 81-104.

Gupta, V. K., Troutman, T. M., and D. R. Dawdy (2007), Towards a nonlinear geophysical theory of floods in river networks: An overview of 20 years of progress. *Nonlinear Dynamics in Geosciences*, Edited by Tsonis, A.A., and J. Elsner, pp 1-31.

Habib, E., and W. F. Krajewski (2001), Estimation of interstation correlation, *Journal of Hydrometeorology*, 2, 621-629.

Habib, E., and W. F. Krajewski (2002), Uncertainty analysis of TRMM ground validation radar-rainfall products: Application to the TEFLUN-B field campaign. *Journal of Applied Meteorology*, 41, 558-572.

Habib. E., G. J. Ciach, and W. F. Krajewski (2004), A method for filtering out raingauge representativeness errors from verification distributions of radar and raingauge rainfall, *Advances in Water Resources*, 27, 967-980.

Habib, E., W. F. Krajewski, and G. J. Ciach (2001), Estimation of rainfall interstation correlation. *Journal of Hydrometeorology*, 2, 621-629.

Harris, D., A. Seed, M. Menabde, and G. Austin (1997), Factors affecting multiscaling analysis of rainfall time series, *Nonlinear Processes in Geophysics*, 4, 137-155.

Harris, D., M. Menabde, A. Seed, and G. Austin (1998), Breakdown coefficients and scaling properties of rain fields, *Nonlinear Processes in Geophysics*, 5, 93-104.

Harris, D., E. Foufoula-Georgiou, and C. Kummerow (2003), Effects of underrepresented hydrometeor variability and partial beam filling on microwave brightness temperatures for rainfall retrieval, *Journal of Geophysical Research,* 108, 8380, doi:10.1029/2001JD001144.

Hooper, R. P., A. Stone, N. Christophersen, E. de Grosbois, and H. M. Seip (1988), Assessing the Birkenes model of stream acidification using a multisignal calibration methodology, *Water Resources Research*, 24, 1308-1316.

Hornberger, G., Germann, P., and Beven, K. (1991), Through flow and solute transport in an isolated sloping soil block in a forested catchment, *Journal of Hydrology*, 124, 81-99.

Hossain, F., E. N. Anagnostou, and T. Dinku (2004), Sensitivity analysis of satellite retrieval and sampling error on flood prediction uncertainty, *IEEE Transactions on Geoscience and Remote Sensing*, 42, 130-139.

Hubert P., Y. Tessier, S. Lovejoy, D. Schertzer, P. Ladoy, J. P. Carbonnel, S. Violette (1993), Multifractals and extreme rainfall events, *Geophysical Research Letters*, 20, 931-934.

Hunter, S. M (1996), WSR-88D radar rainfall estimation: capabilities, limitations and potential improvements, *National Weather Digest*, 20, 26-38.

Hutchinson, T. P (1997), A comment on correlation in skewed distributions, *Journal of General Psychology*, 124, 211-215.

IFRCRCS (1997), *World Disasters Report 1997*, Oxford University Press, 173 pp.

Jakeman, A. J., and G. M. Hornberger (1993), How much complexity is warranted in a rainfall-runoff model? *Water Resources Research*, 29, 2637-2649.

Jennings, M. E., W. O. Thomas, Jr., and H. C. Riggs (1994), Nationwide Summary of U.S. Geological Survey Regional Regression Equations for Estimating Magnitude and Frequency of Floods for Ungaged Sites, 1993: *U.S. Geological Survey Water-Resources Investigations Report* 94-4002, 196 pp.

Johnson, N. L., S. Kotz, and N. Balakrishnan (1995), *Continuous univariate distributions*, Vol 2, Wiley, New York, 719 pp.

Journel, A. G., and C. I. Huijbregts (1978), *Mining geostatistics*, Academic Press, New York, 597 pp.

Kedem, B., and L. S. Chiu (1987), On the lognormality of rain rate, *Proceedings of the National Academy of Sciences*, 84, 901-905.

Kedem, B., L. S. Chiu, and G. R. North (1990), Estimation of mean rain rate: Application to satellite observations, *Journal of Geophysical Research*, 95, 1965-1972.

Kitchen, M., and R. M. Blackall (1992), Representativeness errors in comparisons between radar and gauge measurements of rainfall. *Journal of Hydrology*, 132, 13-33.

Kitchen, M., and P. M. Jackson (1993), Weather radar performance at long range simulated and observed, *Journal of Applied Meteorology*, 1993, 32, 975-985.

Kirkby, M. J. (1978), *Hillslope Hydrology, Landscape Systems*, Wiley, Chichester, 389 pp.

Kirkby, M. (1988), Hillslope runoff processes and models, *Journal of Hydrology*, 100, 315–339.

Klazura, G. E., and D. A. Imy (1993), A description of the initial set of analysis products available from the NEXRAD WSR-88D system, *Bulletin of the American Meteorological Society*, 74, 1293-1311.

Koren, V. I., B. D. Finnerty, J. C. Schaake, M. B. Smith, D.-J. Seo, and Q.-Y. Duan (1999), Scale dependencies of hydrologic models to spatial variability of precipitation, *Journal of Hydrology*, 217, 285-302.

Krajewski, W. F., and C. J. Duffy (1988), Estimation of correlation structure for a homogeneous isotropic random field: A simulation study, *Computers and Geosciences*, 14, 113-122.

Krajewski, W. F., and J. A. Smith (2002), Radar hydrology: Rainfall estimation, *Advances in Water Resources*, 25, 1387-1394.

Krajewski, W. F., E. N. Anagnostou, and G. J. Ciach (1996), Effect of radar observation process on the inferred rainfall statistics, *Journal of Geophysical Research*, 101, 26,493-26,502.

Krajewski, W. F., G. J. Ciach, and E. Habib (2003), An analysis of small-scale rainfall variability in different climatological regimes, *Hydrologic Sciences Journal*, 48, 151-162.

Krajewski, W. F., V. Lakshmi, K. P. Georgakakos, and S.C. Jain (1991), A Monte Carlo study of rainfall sampling effect on a distributed catchment model, *Water Resources Research*, 27, 119-128.

Krajewski, W. F., G. J. Ciach, J. R. McCollum, and C. Bacotiu (2000), Initial validation of the global precipitation climatology project monthly rainfall over the United States. *Journal of Applied Meteorology*, 39, 1071-1086.

Krajewski, W. F., A. Kruger, R. Lawrence, J. A. Smith, P. Domaszczynski, B. -C. Seo, R. Goska, C. Gunyon, M. L. Baeck, M. Steiner, M. K. Ramamurthy, J. Weber, and S. A. DelGreco (2009), Hydro-NEXRAD: Informational Technology Framework for Enhancing Hydrologic Use of WSR-88D Data, submitted to *Journal of Hydroinformatics.*

Lark, R. M., A Comparison of some robust estimators of variogram for use in soil survey, *European Journal of Soil Science*, 2000a; 51: 137-157.

Lark, R.M. (2000b), Estimating variograms of soil properties by the method-of-moments and maximum likelihood, *European Journal of Soil Science*, 51, 717-728.

Lashermes, B. and E. Foufoula-Georgiou (2007), Area and width functions of river networks: New results on multifractal properties, *Water Resources Research*, 43, W09405, doi:10.1029/2006WR005329.

Lavallee, D., S. Lovejoy, D. Schertzer, P. Ladoy (1993), Nonlinear variability and landscape topography: analysis and simulation, In: *Fractals in Geography*, De Cola, L., and N. Lam, Editors, Prentice Hall, pp 158-192.

Lazarev A, D. Schertzer, S. Lovejoy, Y. Chigirinskaya (1994), Multifractal analysis of tropical turbulence, part II: vertical scaling and Generalized Scale Invariance, *Nonlinear Processes in Geophysics*. 1, 115-123.

Lebel, T., G. Bastin, C. Obled, and J.D. Creutin (1987), On the accuracy of areal rainfall estimation: A case study, *Water Resources Research*, 23, 2123-2134.

Liu H. H, and F. J. Molz (1997), Multifractal analysis of hydraulic conductivity distributions, *Water Resources Research*, 33, 2483-2488.

Lilley, M., S. Lovejoy, N. Desaulniers-Soucy, D. Schertzer (2006), Multifractal large number of drops limit in rain, *Journal of Hydrology*, 328, 20-37.

Loáiciga, H.A., W.-G. Yeh, and M.A. Ortega-Guerrero (2006), Probability density functions in the analysis of hydraulic conductivity data, *ASCE Journal of Hydrologic Engineering*, 11, 442-450.

Lovejoy, S., and D. Schertzer (1990), Multifractals, universality classes and satellite and radar measurements of cloud and rain fields, *Journal of Geophysical Research*, 95, 2021-2034.

Lovejoy, S., and D. Schertzer (2006), Multifractals, cloud radiances and rain, *Journal of Hydrology*, 322, 59-88.

Lovejoy, S., D. Schertzer, and V. C. Allaire (2008), The remarkable wide range spatial scaling of TRMM precipitation, *Atmospheric Research*, 90, 10-32.

Mantilla, R. Physical basis of statistical self-similarity in peak flows on random self-similar networks, *PhD dissertation 2007*, University of Colorado, Boulder.

Mantilla, R., and V. K. Gupta (2005), A GIS numerical framework to study the process basis of scaling statistics in river networks, *IEEE Geoscience and Remote sensing letters*; 2, 404-408.

Mantilla, R., V. K. Gupta, and O.J. Mesa (2006), Role of coupled flow dynamics and real network structures on Hortonian scaling of peak flows, *Journal of Hydrology,* 322, 155-167.

Mantoglou, A., and J.L. Wilson (1982), The turning bands method for the simulation of random fields using line generation by spectral method, *Water Resources Research*, 18, 1379-1394.

Marsan D., D. Schertzer D, and S. Lovejoy (1996), Causal space-time multifractal modelling of rain, *Journal of Geophysical Research*, 31, 333-346.

McCollum, J. R., W. F. Krajewski, R. R. Ferraro, and M. B. Ba (2002), Evaluation of biases of satellite rainfall estimation algorithms over the continental United States, *Journal of Applied Meteorology*, 41, 1065-1080.

Mejia, J.M., and I. Rodriguez-Iturbe (1974a), Correlation links between normal and lognormal processes, *Water Resources Research*, 10, 689-690.

Mejia, J., and I. Rodriguez-Iturbe (1974b), On the synthesis of random fields sampling from the spectrum: An application to the generation of hydrologic spatial processes. *Water Resources Research,* 10, 705-711.

Menabde, M., and M. Sivapalan (2001), Linking space-time variability of river runoff and rainfall fields: a dynamic approach, *Advances in Water resources,* 24, 1001-1014.

Menabde, M., A. Seed, D. Harris, and G. Austin (1997a), Self-similar random fields and rainfall simulation, *Journal of Geophysical Research*, 102, 13,509-13515.

Menabde, M., D. Harris, A. Seed, G. Austin, and D. Stow (1997b), Multiscaling properties of rainfall and bounded random cascades, *Water Resources Research*, 33, 2823-2830.

Menabde, M., S. Veitzer, V. K. Gupta, and M. Sivapalan (2001), Tests on peak flow scaling in simulated self-similar river networks, *Advances in Water Resources,*24, 991-999.

Michalak, A.M., L. Bruhwiler, and P.P. Tans (2004), A geostatistical approach to surface flux estimation of atmospheric trace gases, Journal *of Geophysical Research*, 109, D14109.

Milly, P. C. D., and P. S. Eagleson (1988), Effect of storm scale on surface runoff volume, *Water Resources Research*, 24, 620– 624.

Mohanty, B.P., M.D. Ankeny, R. Horton, and R.S. Kanwar (1994), Spatial analysis of hydraulic conductivity measured using disc infiltrometers, *Water Resources Research*, 30, 2489-2498.

Morales, J. E., and G. Poveda (2009), Diurnally driven scaling properties of Amazonian rainfall fields: Fourier spectra and order-q statistical moments, *Journal of Geophysical Research*, 114, D11104, doi:10.1029/2008JD011281.

Morin, E., D. C. Goodrich, R. A. Maddox, X. Gao, H. V. Gupta, and S. Sorooshian (2006), Spatial patterns in thunderstorm rainfall events and their coupling with watershed hydrological response, *Advances in Water Resources*, 29, 843– 860, doi:10.1016/j.advwatres.2005.07.014.

Morris, M.D. (1991), On counting number of data pairs for semivariogram estimation, *Mathematical Geology*, 23, 929-943.

Morrissey, M.L., J.A. Maliekal, J.S. Green, and J. Wang (1995), The uncertainty of simple spatial averages using rain gauge networks, *Water Resources Research*, 31, 2011-2017.

Morrissey, M.L., J. A. Maliekal, J. S. Greene, and J. Wang (1995), The uncertainty of simple spatial averages using raingauge networks, *Water Resources Research* 31, 2011-2017.

Morrison, J. E., and J. A. Smith (2001), Scaling properties of flood peaks, *Extremes*, 4, 5-22.

Naden, P. S. (1992), Spatial variability in flood estimation for large catchments: The exploitation of channel network structure, *Hydrological Sciences Journal*, 37, 53–71.

Naud C., D. Schertzer, and S. Lovejoy (1996), Fractional integration and radiative transfer in multifractal atmospheres, In: Woyczynski W, Molchansov S (Eds.) *Stochastic Models in Geosystems*, pp. 239-267, New York, Berlin, Heidelberg: Springer-Verlag.

Nicotina, L., E. Alessi Celegon, A. Rinaldo, and M. Marani (2008), On the impact of rainfall patterns on the hydrologic response. *Water Resources Research*, 44, W12401, doi:10.1029/2007WR006654.

Nicholson, S. (1986), Spatial coherence of African rainfall anomalies: Interhemispheric teleconnections, *Journal of Climate and Applied Meteorology*, 25, 1365-1381.

Nijssen, B., and D. P. Lettenmaier (2004), Effect of precipitation sampling error on simulated hydrological fluxes and states: Anticipating the Global Precipitation Measurement satellites, *Journal of Geophysical Research,* 109: D02103, doi:10.1029/2003JD003497.

Novikov, E. A. (1990), Effects of intermittency on statistical characteristics of turbulence and the scale similarity of breakdown coefficients, *Physics of Fluids A*, 2, 814-820.

Ntelekos, A. A., J. A. Smith, M.-L. Baeck, W.F. Krajewski, A. J. Miller, and R. Goska (2008), Extreme hydrometeorological events and the urban environment: Dissecting the 7 July 2004 thunderstorm over the Baltimore, MD, metropolitan region, *Water Resources Research*, 44, W08446, doi:10.1029/2007WR006346.

Ntelekos, A. A., J. A. Smith, L. J. Donner, J. D. Fast, W. I. Gustafson Jr., E.G. Chapman, and W. F. Krajewski (2009), The effects of aerosols on intense convective precipitation in the Northeastern US, *Quarterly Journal of the Royal Meteorological Society*, 135, 1367-1391.

Nykanen, D. K., and D. Harris (2003), Orographic influences on the multiscale statistical properties of precipitation, *Journal of Geophysical Research*, 108(D8), 8381, doi:10.1029/2001JD001518.

Obled, C., J. Wedling, and K. Beven (1994), The sensitivity of hydrological models to spatial rainfall patterns: An evaluation using observed data, *Journal of Hydrology*, 159, 305– 333.

O'Callaghan, J. F., and D. Mark (1984), The extraction of drainage networks from digital elevation data. *Computer Vision, Graphics, Image Processing,* 28, 328–344.

Ogden, F. L., and D. R. Dawdy (2003), Peak discharge scaling in small Hortonian watershed, *Journal of Hydrologic Engineering*, 8, 64-73.

Ogden, F. L., and B. A. Watts (2000), Saturated area formation on nonconvergent hillslope topography with shallow soils: a numerical investigation. *Water Resources Research*, 36, 1795–1804.

Ogden, F. L., and P. Y., Julien (1993), Runoff sensitivity to the temporal and spatial rainfall variability at runoff plane and small basin scales, *Water Resources Research,* 29, 2589-2597.

Peckham, S. D., and V. K. Gupta (1999), A reformulation of Horton's laws for large river networks in terms of statistical self-similarity, *Water Resources Research*, 35, 2763-2777.

Pessoa, M. L., R. L. Bras, and E. R. Williams (1993), Use of weather radar for flood forecasting in the Sieve River basin: A sensitivity analysis, *Journal of Applied Meteorology*, 32, 462-475.

Purdy, J. C., D. Harris, G. L. Austin, A.W. Seed, and W. Gray (2001), A case study of orographic rainfall processes incorporating multiscaling characterization techniques, *Journal of Geophysical Research,* 106, 7837-7845.

Rea, J., and R. Knight (1998), Geostatistical analysis of ground-penetrating radar data: A means of describing spatial variation in the subsurface, *Water Resources Research*, 34, 329-339.

Reed, S. M., and D. R. Maidment (1999), Coordinate transformations for using NEXRAD data in GIS-based hydrologic modeling, *Journal of Hydrologic Engineering*, 4, 174-182.

Refsgaard, J. C., and J. Knudsen (1996). Operational validation and intercomparison of different types of hydrological models, *Water Resources Research*, 32, 2189-2202.

Ricciardulli, L., and P. D. Sardeshmukh (2002), Local time- and space scales of organized tropical deep convection, *Journal of Climate*, 15, 2775-2790.

Ridolfi, L., P. D'Odorico, A. Porporato, and I. Rodriguez-Iturbe (2003), Stochastic soil moisture dynamics along a hillslope, *Journal of Hydrology*, 272, 264-275.

Riggs, H. C. (1973), Regional Analyses of Streamflow Characteristics, *U.S. Geological Survey Techniques of Water-Resources Investigations*, Book 4, Chap. B3, 15 pp.

Ritsema, C. J., J. Stolte, K. Oostindie, and E. Van der Elsen (1996), Measuring and modelling of soil water dynamics and runoff generation in an agriculture hillslope, *Hydrologic Processes*, 10, 1081–1089.

Robinson, J. S., and M. Sivapalan (1997), An investigation into the physical causes of scaling and heterogeneity of regional flood frequency, *Water Resources Research*, 33 1045-1059.

Robinson, J. S., M. Sivapalan, and J. D. Snell (1995), On the relative roles of hillslope processes, channel routing, and network geomorphology in the hydrologic response of natural catchments, *Water Resources Research*, 31, 3089-3101.

Russo, D., and W.A. Jury (1987), A theoretical study of the estimation of the correlation scale in spatially variable fields, 1, Stationary fields, *Water Resources Research*, 23, 1257-1268.

Saulnier, G.-M., and M. Le Lay (2009), Sensitivity of flash-flood simulations on the volume, the intensity, and the localization of rainfall in the Ce´vennes-Vivarais region (France), *Water Resources Research*, 45, W10425, doi:10.1029/2008WR006906.

Schertzer, D., and S. Lovejoy (1987), Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes, *Journal of Geophysical Research*, 92(D8), 9693-9714.

Schmitt F., D. Schertzer,, S. Lovejoy, and G. Brunet (1994), Estimation of universal multifractal indices for atmospheric turbulent velocity fields,. *Fractals* 1, 569-575.

Schmitt F, D. Schertzer,, S. Lovejoy, and G. Brunet (1996), Universal multifractal structure of atmospheric temperature and velocity fields, *Europhysics Letters*, 34, 195-200.

Schuurmans, J. M., and M. F. P. Bierkens (2007), Effect of spatial distribution of daily rainfall on interior catchment response of a distributed hydrological model, *Hydrology and Earth System Science*, 11, 677– 693.

Seo, B.-C., W. F. Krajewski, A. Kruger, P. Domaszczynski, M. Steiner, and J. A. Smith (2009), Hydro-NEXRAD Radar-Rainfall Estimates Algorithm, *to be submitted to Journal of Hydroinformatics*.

Seo, D.-J., and J. P. Breidenbach (2002), Real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements, *Journal of Hydrometeorology*, 3, 93-111.

Seo, D.-J., J. P. Breidenbach, and E. R. Johnson (1999), Real-time estimation of mean field bias in radar rainfall data, *Journal of Hydrology*, 223, 131-147.

Shah, S. M. S., P. E. O'Connell, J. R. M. Hosking (1996), Modelling the effects of spatial variability in rainfall on catchment response. 2. Experiments with distributed and lumped models, *Journal of Hydrology*, 175, 89-111.

Sharif, H.O., F. L. Ogden, W. F. Krajewski, and M. Xue (2002), Numerical simulations of radar rainfall error propagation, *Water Resources Research*, 38, 10.1029/2001WR000525.

Singh V.P., and D.K. Frevert (eds) (2002a), *Mathematical models of small watershed hydrology*. Water Resources Press, Highlands Ranch, Colorado, 972 pp.

Singh V.P., and D.K. Frevert (eds) (2002b), *Mathematical models of large watershed hydrology*. Water Resources Press, Highlands Ranch Colorado, 972 pp.

Sivapalan, M. Takeuchi, K. Franks, S.W., Gupta, V.K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J.J., Mendiondino, E.M., O'Connell, P.E., Oki, T., Pomeroy, J.W., Schertzer, D., Uhlenbrook, S., and E. Zehe (2003), IAHS decade on prediction in ungauged basins (PUB), 2003-2012: Shaping an exciting future for the hydrologic sciences. *Hydrologic Sciences Journal*, 48, 857-880.

Smith, J.A. Representation of basin scale in flood peak distribution (1992), *Water Resources Research*, 28, 2993-2999.

Smith, J. A., and W. F. Krajewski (1991), Estimation of the mean field bias of radar rainfall estimates, *Journal of Applied Meteorology*, 30, 397-412.

Smith, J.A., D.-J Seo, M. Baeck, and M. Hudlow (1996), An intercomparison study of NEXRAD precipitation estimates, *Water Resources Research*, 32, 2035-2045.

Smith, M. B., D.-J. Seo, V. I. Koren, S. M. Reed, Z. Zhang, Q. Duan, F. Moreda, and S. Cong (2004a), The distributed model intercomparison project (DMIP): motivation and experiment design, *Journal of Hydrology*, 298, 4-26.

Smith, M. B., V. I. Koren, S. M. Reed, Z. Zhang, J.-J. Pan, and F. Moreda (2004b), Runoff response to spatial variability in precipitation: an analysis of observed data, *Journal of Hydrology*, 298, 267-286.

Smith, T. M., P. A. Arkin, J. J. Bates, and G. J. Huffman (2006), Estimating bias of satellite-based precipitation products, *Journal of Hydrometeorology*, 7, 841-856.

Stauffer, F. (2005), Uncertainty estimation of path lines in groundwater models, *Ground Water*, 43, 843-849.

Steiner M, J. A. Smith, S. J. Burges, C. V. Alonso, and R.W. Darden (1999), Effect of bias adjustment and rain gauge data quality control on radar rainfall estimation, *Water Resources Research,* 35, 2487-2503.

Strahler, A. N (1957), Quantitative analysis of watershed geomorphology, *Transactions of American Geophysical Union*, 38, 913-920.

Sumner, G. N. (1982), The spatial organization of daily rainfall in eastern New South Wales, *International Journal of Climatology*, 3, 361-374.

Tchiguirinskaia, I., S. Lu, F. J. Molz, T. M. Williams, D. Lavallee (2000), Multifractal versus monofractal analysis of wetland topography, *Stochastic Environmental Research and Risk Assessment*, 14, 8-32.

Tessier Y., S. Lovejoy, D. Schertzer (1993a), Universal multifractals in rain and clouds: theory and observations, *Journal of Applied Meteorology*, 32, 223-250

Tessier Y., S. Lovejoy, D. Schertzer, D. Lavallee, and B. Kerman (1993b) Universal multifractal indices for the ocean surface at far red wavelengths, *Geophysical Research Letters*, 20, 1167-1170.

Tetzlaff, D., and S. Uhlenbrook (2005), 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 Science*, 9, 29-41.

Troutman, B. M., and T. M. Over (2001), River flow mass exponents with fractal channel networks and rainfall, *Advances in Water Resources*, 24, 967-989.

Vanmarcke E (1983), *Random Fields: Analysis and Synthesis,* MIT Press 1983; p. 382.

Veitzer, S.A., and V. K. Gupta (2001), Statistical self-similarity of width function maxima with implications to floods, *Advances in Water Resources*, 24, 955-965.

Villarini, G. and W.F. Krajewski (2007), Evaluation of the research-version TMPA three-hourly 0.25°×0.25° rainfall estimates over Oklahoma, *Geophysical Research Letters*, 34, L05402, doi:10.1029/2006GL029147.

Villarini, G., and W. F. Krajewski (2008), Empirically-based modeling of spatial sampling uncertainties associated with rainfall measurements by rain gauges. *Advances in Water Resources,* 31, 1015-1023.

Villarini, G., and Krajewski, W. F. (2009a), Review of different sources of uncertainty in single polarization radar-based estimates of rainfall, *Reviews in Geophysics*, doi: 10.1007/s10712-009-9079-x, In Press.

Villarini, G., and W. F. Krajewski (2009b), Inference of spatial scaling properties of rainfall: Impact of radar-rainfall estimation uncertainties, *IEEE Geoscience and Remote Sensing Letters*, 6(4), 812-815, 2009.

Villarini, G., and W. F. Krajewski (2009c), Empirically based modelling of uncertainties in radar rainfall estimates for a C-band weather radar at different time scales, *Quarterly Journal of the Royal Meteorological Society*, 135, 1424-1438.

Villarini, G., and W. F. Krajewski (2009d), Sensitivity studies of the models of radar-rainfall uncertainties, *Journal of Applied Meteorology and Climatology*, doi:10.1175/2009JAMC2188.1, In Press.

Villarini, G., G. J. Ciach, W. F. Krajewski, K. M. Nordstrom, and V. K. Gupta (2007a), Effects of systematic and random errors on the spatial scaling properties in radar-estimated rainfall. In: *Nonlinear Dynamics in Geosciences*, Tsonis, A.A. & J. Elsner, Editors, Springer, pp 37-51.

Villarini, G., J. B. Lang, F. Lombardo, F. Napolitano, F. Russo, and W. F. Krajewski (2007b), Impact of different regression frameworks on the estimation of the scaling properties of radar rainfall, *Atmospheric Research*, 86(3-4), 340-349.

Villarini, G., P. V. Mandapaka, W. F. Krajewski, and R. J. Moore (2008), Rainfall and Sampling Uncertainties: A Rain Gauge Perspective, *Journal of Geophysical Research*, doi:10.1029/2007JD009214.

Villarini, G., W. F. Krajewski, G. J. Ciach and D. L. Zimmerman (2009a), Product-error-driven generator of probable rainfall conditioned on WSR-88D precipitation estimates, *Water Resources Research*, 45, W01404, doi:10.1029/2008WR006946.

Villarini, G., W. F. Krajewski, and J. A. Smith (2009b), New paradigm for statistical validation of satellite precipitation estimates: Application to a large sample of the TMPA 0.25-degree three hourly estimates over Oklahoma, *Journal of Geophysical Research*, 114, D12106, doi:10.1029/2008JD011475.

Vivoni, E. R., D. Entekhabi, and R. N. Hoffman (2007), Error propagation of radar rainfall nowcasting fields through a fully distributed flood forecasting model, *Journal of Applied Meteorology and Climatology*, 46, 932-940.

Vogel,R.M. and A.Sankarasubramanian (2000), Scaling properties of annual streamflow in the continental United States, *Hydrological Sciences Journal*,45,465-476.

Warrick, A.W., and D. E. Myers (1987), Optimization of sampling locations for variogram calculations, *Water Resources Research*, 23, 496-500.

Weiler, M., and J. McDonnell (2004), Virtual experiments: a new approach for improving process conceptualization in hillslope hydrology, *Journal of Hydrology*, 285, 3-18.

Western, A.W., R. B. Grayson, G. Bloschl, G.R. Willgoose, and T.A. McMahon (1999), Observed spatial organization of soil moisture and its relation to terrain indices, *Water Resources Research*, 35, 797-810.

Wood, A.T.A., and G. Chan (1994), Simulation of stationary Gaussian processes in [0; 1]d, *Journal of Computational and Graphical Statistics*, 3, 409-432.

Xian-Huan, W. (1994), Estimation of statistical parameters for censored lognormal hydraulic conductivity measurements, *Mathematical Geology*, 26, 717-731.

Yaglom, M. A (1987), *Correlation theory of stationary and related random functions*, V. I and II, Springer-Verlag.

Young, C. B., B. R. Nelson, A. A. Bradley, J. A. Smith, C. D. Peters-Lidard, A. Kruger, and M. L. Baeck (1999), An evaluation of NEXRAD precipitation estimates in complex terrain, *Journal of Geophysical Research*, 104, 19691-19703.

Young, C. B., A. A. Bradley, W. F. Krajewski, A. Kruger, and M. L. Morrissey (2000), Evaluating NEXRAD multisensor precipitation estimates for operational hydrologic forecasting, *Journal of Hydrometeorology*, 1, 241-254.

Zawadzki, I (1975), On the radar-raingauge comparison, *Journal of Applied Meteorology* 14, 1430-1436.

Zawadzki, I (1984), Factors affecting the precision of radar measurements of rain, *Preprints of 22nd Conference on Radar Meteorology*, American Meteorological. Society, Zurich, pp. 251-256.

Zheng, L., and S.E. Silliman (2000), Estimating the theoretical semivarogram from finite numbers of measurements, *Water Resources Research*, 36, 361-366.

Zhang Y., T. Adams, and J. V. Bonta (2007), Sub-pixel scale rainfall variability and the effects on separation of radar and gauge rainfall errors, *Journal of Hydrometeorology*, 8, 1348-1363.

Zimmerman D. L., and M. B. Zimmerman (1991), A comparison of spatial semivariogram estimators and corresponding ordinary kriging predictors, *Technometrics*, 33, 77-91.