Multi-Sensor Rainfall Data Assimilation using Ensemble Approaches

by

Virat Chatdarong

B.S. Civil Engineering
Chulalongkorn University, 2000

M.Eng. Civil and Environmental Engineering
Massachusetts Institute of Technology, 2001

Submitted to the Department of Civil and Environmental Engineering
In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2006

Signature of Author ....
Department of Civil and Environmental Engineering
May 12th, 2006

Certified by ....
Dennis McLaughlin
H.M. King Bhumibol Professor of Civil and Environmental Engineering
Thesis Supervisor

Certified by ....
Dara Entekhabi
Professor of Civil and Environmental Engineering
Thesis Supervisor

Accepted by ....
Andrew Whittle
Chairman, Committee for Graduate Student

Multi-Sensor Rainfall Data Assimilation using Ensemble Approaches

by

Virat Chatdarong

Submitted to the Department of Civil and Environmental Engineering
on May 12th, 2006 in partial fulfillment of the
Requirements for the Degree of Doctor of Philosophy in
the field of Hydrology

Abstract

Rainfall is a major process transferring water mass and energy from the atmosphere to the surface. Rainfall data is needed over large scales for improved understanding of the Earth climate system. Although there are many instruments for measuring rainfall, none of them can provide continuous global coverage at fine spatial and temporal resolutions.

This thesis proposes an efficient methodology for obtaining a probabilistic characterization of rainfall over an extended time period and spatial domain. The characterization takes the form of an ensemble of rainfall replicates, each conditioned on multiple measurement sources. The conditional replicates are obtained from ensemble data assimilation algorithms (Kalman filters and smoothers) based on a recursive cluster rainfall model. Satellite measurements of cloud-top temperatures are used to identify areas where rainfall can possibly occur. A variational field alignment algorithm is used to estimate rainfall advective velocity field from successive cloud-top temperature images. A stable pseudo-inverse improves the stability of the algorithms when the ensemble size is small.

The ensemble data assimilation is implemented over the United States Great Plains during the summer of 2004. It combines surface rain-gauge data with three satellite-based instruments. The ensemble output is then validated with ground-based radar precipitation product. The recursive rainfall model is simple, fast and reliable. In addition, the ensemble Kalman filter and smoother are practical for a very large-scale data assimilation problem with a limited ensemble size.

Finally, this thesis describes a multi-scale recursive algorithm for estimating scaling parameters for popular multiplicative cascade rainfall models. In addition, this algorithm can be used to merge static rainfall data from multiple sources.

Thesis Supervisor: Dennis McLaughlin
Title: H.M. King Bhumibol Professor of Civil and Environmental Engineering,

Thesis Supervisor: Dara Entekhabi
Title: Professor of Civil and Environmental Engineering

Acknowledgments

There are a great number of people who have contributed significantly to my Ph.D. years at the M.I.T. First of all, I would like to thank my thesis supervisors, Prof. Dennis McLaughlin and Prof. Dara Entekhabi, for make this all possible. I am deeply indebted to them for having shared their knowledge and experience, from which I benefited tremendously over the course of my Ph.D. I would like to extend my sincere gratitude to my thesis committee, Dr. Ross Hoffman and Prof. Efi Foufoula-Georgiou for their invaluable suggestions and supports on this thesis. I also appreciate the assistance from Prof. David H. Staelin and C. Surusavadee for providing the AMSU-B data and from Dr. Christopher Grassotti at AER for providing the NOWRAD data.

I am most obliged to the Ananda Mahidol Foundation for supports over the past years. It is my most honorable opportunity to receive a scholarship from H.M. King Bhumibol Rama IX and HRH Princess Sirindhorn of Thailand. I would like to thank HE Dr. Chaovana Nasylvanta and A. Anukalya Israsena for giving me an opportunity of a life time, P'Pop Komsorn and Khun Manida for their helps and supports. I am also grateful to A. Sucharit Koontanakulvong, A. Chaiyuth Suksri, A. Tuantan Kitpaisalsakul, and other professors and lecturers at the Civil Engineering and at the Water Resources Department, Chulalongkorn University for their encouragements, supports, and advices.

I have also greatly appreciated the technical and social interaction with my fellow MIT students: Crystal Ng for editing my thesis, David Flagg for acquiring most of my data to a usable format, and Sai Ravela for the multi-resolution alignment algorithm. Thank to the previous and current students from the McLaughlin and Entekhabi research groups for many inspiring and useful inputs, discussions, and solutions. Thank to my fellow Thai friends for making my experience here so enjoyable and memorable.

Above all I am deeply grateful to my family: my dad, my mom, my sister Ann, and Bell for their love and understanding through these years. I hope we can finally find some pride and joy after these long hard years.

Contents

1. Introduction

2. The United States Great Plains Case Study

3. Dynamic Rainfall Model

4. Dynamic Rainfall Data Assimilation

5. Estimation of the Multiplicative Cascade Rainfall Model Parameters by the EM-SRE Algorithm

6. Thesis Conclusion and Future Researches

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Importance of Rainfall in the Earth Climate System

Water is essential to all life on our planet. It covers three quarters of the Earth's surface and is an active component of the atmosphere. The collection of stores of water that exist in the Earth system is called the hydrosphere. The hydrosphere extends approximately 15 kilometers up into the atmosphere and approximately 1 kilometer down into the lithosphere or the Earth's crust. Within the hydrosphere, water circulates between the atmosphere and the surface stores. The various pathways constituting the hydrologic cycle are illustrated in Figure 1-1.

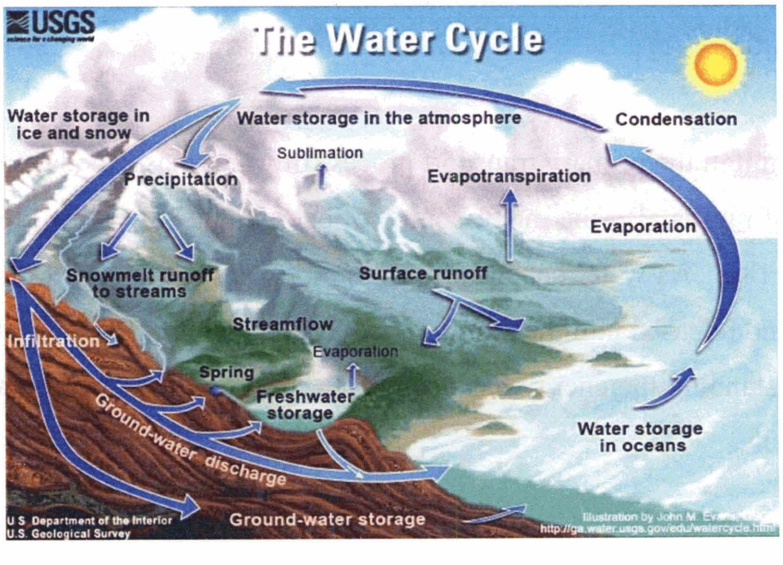


Figure 1-1: Schematic diagram of the hydrologic cycle (from U.S. Geological Survey)

In the hydrological cycle, water mass is transported from the atmosphere to the surface by the process called precipitation. The precipitation process can occur in liquid form or solid form. However, the majority of precipitation, especially over the tropics, is in a liquid form and is commonly referred to as rainfall. Rainfall has immediate and major impacts on the environment and human livelihood. It can infiltrate and run off to nearby streams shortly after it reaches the ground. Conversely, the frozen precipitation including snow, ice, and hail can remain inactive where it falls for a long time before it begins to melt and interact with the environment.

Rainfall amount and its variation are primary factors in many engineering and management decisions. Excess rainfall can cause flooding, while insufficient rainfall can cause drought and starvation. These extreme behaviors of rainfall result in enormous damage to properties and human lives. Rainfall couples with other environmental variables in complex manners and causes global scale weather anomalies as well. Perhaps the most well known example of a tropical climate anomaly is El Nino. It is a disruption of the ocean-atmosphere system in the tropical Pacific, which has important consequences for weather and climate around the globe. The El Nino-Southern Oscillation (ENSO) can cause anomalously wet weather in California, wetter and colder winters in the eastern United States and dryer summer monsoon seasons across the southern hemisphere [2].

Not only is rainfall the primary transport process of water mass, the latent heat absorption and release of rain is also the major source of energy that drives the global atmospheric circulation [118]. As water changes from liquid to vapor and back to liquid, latent heat is absorbed and released into the atmosphere. The variation of latent heat is an important part of understanding the energy balance, which in turn affects the climate on a regional and global scale.

1.2 Rainfall Data Acquisition

Rainfall data is an essential ingredient for better understanding the Earth's climate system. In many applications, we seek to obtain rainfall information at the spatial and temporal resolution of interest. This data can be collected from rainfall measuring instruments, or can be estimated from rainfall prediction models. However, obtaining accurate and comprehensive rainfall data from either source can be difficult to achieve because rainfall measurements are scattered in space and time. In addition, the intermittent dynamics of rainfall is too complex to simulate accurately over a large scale and an extended period.

1.2.1 Rainfall Measurements

There are many types of instruments employed worldwide to detect and collect rainfall data. They include rain gauges, ground-based radar stations, and remote sensing instruments onboard orbiting satellites. Although rainfall data is universally expressed as the depth of water falling on a level surface in inches or millimeters, the algorithms to obtain rainfall data for each instrument is usually unique. Even instruments from the same platform can vary greatly in their characteristics, scales, coverage and accuracy. A suitable choice instrument is based on many factors, such as, topography, accessibility, desired spatial and temporal resolution, etc.

The first and most basic instrument used by humans to measure rainfall is the rain-gauge. It directly collects water in an open container. Rainfall measurement from a rain-gauge station can be very useful if the continuous record of rainfall data over a particular location or over a small region is of interest. At the global scale, interpolating rainfall measurement from rain gauge station requires a dense network that can be very costly. In addition, it is impossible to set up and routinely operate rain gauge stations in remote areas or over the ocean.

A new era of rainfall measurement emerges from the development of radar during World War II. Microwave radiation at wavelengths between 1 – 20 centimeters can indicate the presence of rain [8, 133, 134]. Active microwave radiation or radar provides information on raindrop distribution, which can be directly converted to rainfall rate. A single radar station offers a means of obtaining rainfall distribution in a three-dimensional space that can only be crudely approximate with rain gauge data. By the end of the 20th century, dense networks of radar stations have been set up in many major populated areas including the Next Generation Weather Radar (NEXRAD) program in United States. Radar has quickly become the primary source of rainfall measurement in many regions. However, worldwide coverage of rainfall data from this source alone is not possible. Furthermore, accuracy and coverage of rainfall measurement from radar stations can be limited by many factors including topography. For example, the utility of the NEXRAD for estimating rainfall at the surface over the United States is highly compromised across the mountainous West [48, 78] as illustrated by data voids in Figure 1-2.

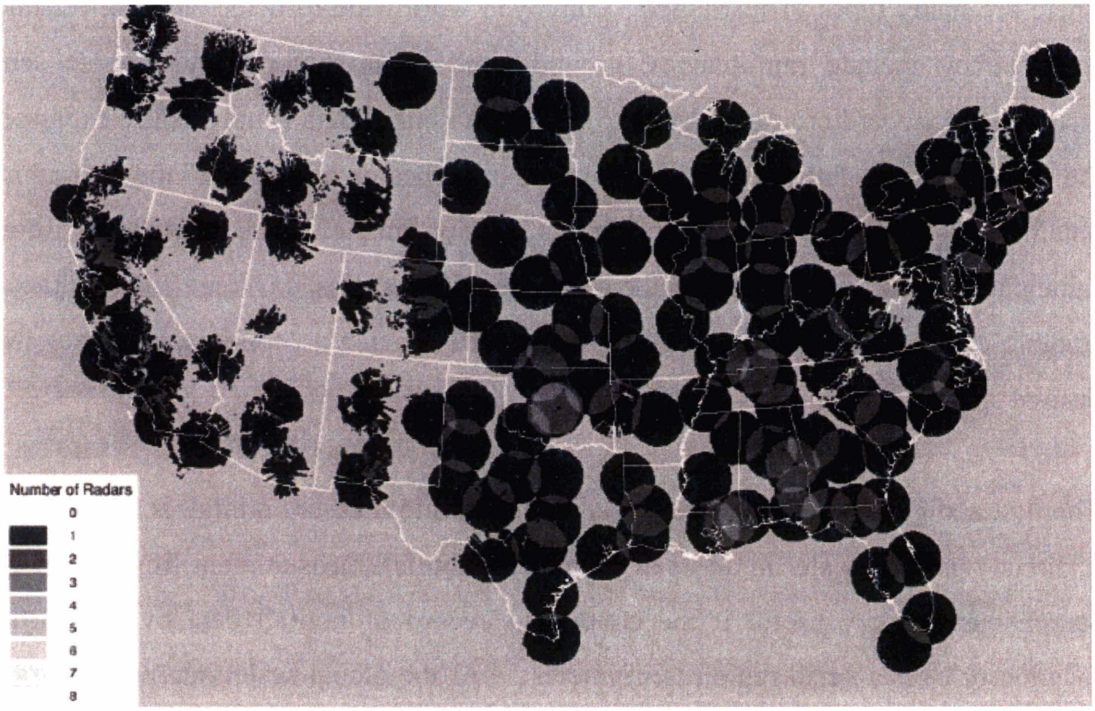


Figure 1-2: Coverage of the NEXRAD network at height of 2 km above the ground level [78]

Another breakthrough in monitoring global rainfall is the development of passive and active microwave satellites. Similar to radar, these air-borne instruments detect emitted and scattered radiation by raindrops and ice particles, which can also be accurately converted to rainfall intensity. Microwave radiation can penetrate through cirrus clouds (i.e., high and thin clouds with no rain) and provide information at various levels of the atmosphere. Currently, there are substantial numbers of satellites with active and passive microwave instruments orbiting around the Earth. Some examples of the current rainfall-measuring satellite projects operated by the United States are the Tropical Rainfall Measuring Mission (TRMM), the Special Sensor Microwave Imager (SSM/I) onboard the Defense Meteorological Satellite Program (DMSP), the Advanced Microwave Sounding Unit (AMSU) onboard National Oceanic and Atmospheric Administration (NOAA) satellites.

Rainfall measurements from microwave satellites are useful for studies of the Earth climate system at a global scale. These satellites can provide information over remote regions and over the ocean whose measurement from rain gauges and radar stations are unavailable. However, since microwave radiations have relatively low energy, these satellites must be in low orbit around the Earth in order to obtain data at a usable resolution. Data from a low satellite is given as the satellite progresses along its track, as in the example shown in Figure 1-3. Measurements at different locations along the track are observed at different times, making it difficult to integrate with other data sources. In addition, the revisit time of the satellite usually takes several hours to several days. At this temporal frequency, rainfall measurement from low orbit satellite is not practical to track rainfall dynamically over any particular region.

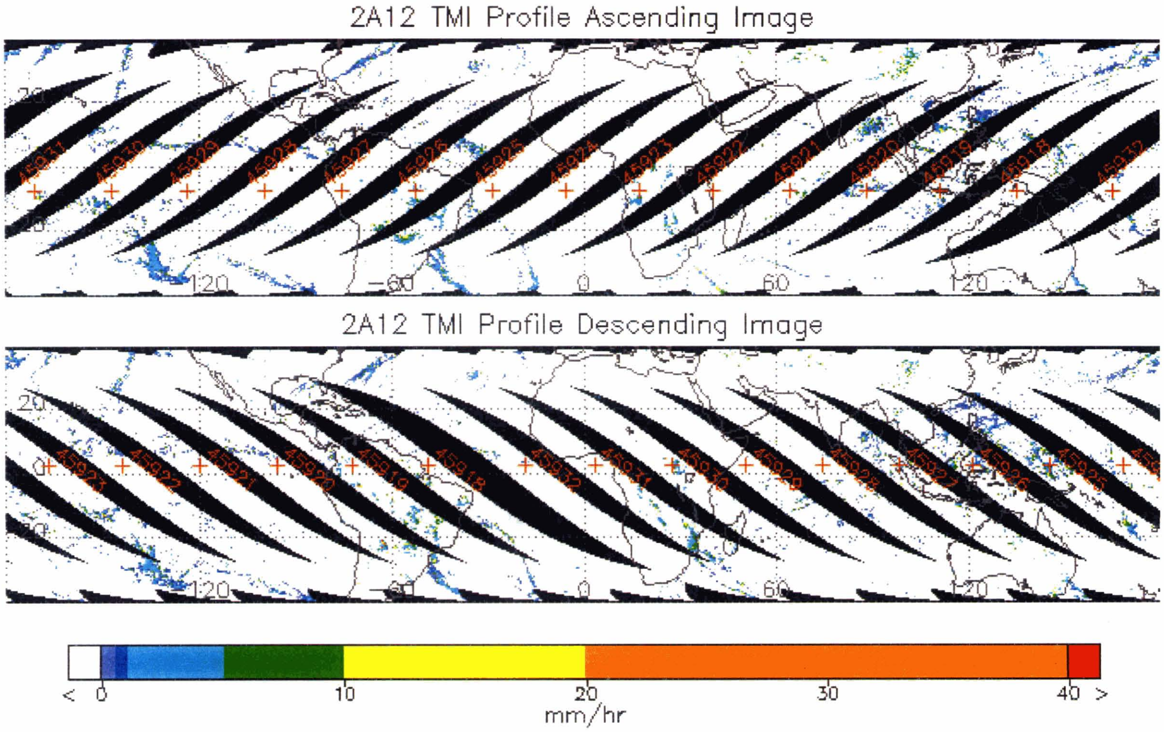
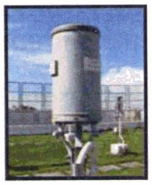
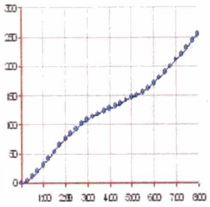
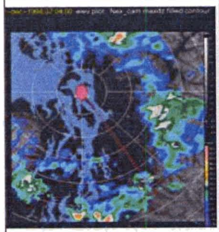
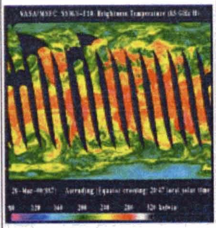

2A12 TMI Profile Ascending Image
20
0
-120
-60
0
60
120
2A12 TMI Profile Descending Image
20
0
-120
-60
0
60
120
< 0
10
20
30
40 >
mm/hr

Figure 1-3: Tracks of TRMM Microwave Imager (TMI) measurement on December 6, 2005 from TRMM Orbital Data Products <http://disc.sci.gsfc.nasa.gov/data/datapool/TRMM/>

As opposed to a low orbit satellite, a geostationary satellite orbits the Earth at a speed matching the Earth's rotation. At 22,300 miles (or 35,800 kilometers) above the Earth, a geostationary satellite is high enough to continuously monitor the Earth with a full-disk view. This seems to be a perfect source of comprehensive rainfall measurement. Unfortunately, microwave radiation, which is commonly used to measure rainfall, has insufficient energy to be detected at this altitude. In addition, any higher energy radiation cannot penetrate through cirrus cloud and accurately estimate rainfall at the surface. The variable detectible at this altitude that is most related to rainfall is the cloud-top temperature from infrared and visible radiation. Although these continuous cloud-top images cannot be directly converted to rainfall intensity, they provide information on cloud-drift winds, cloud-thickness, and moisture contents, which can help with estimating rainfall.

Table 1.1: Characteristics, advantages, and limitations of rainfall-measuring instruments

Instrument Type	 Rain-gauge Station	 Ground-based Radar	 Low-Orbit Satellites	 High-Orbit Satellites
Measurement Data	Cumulative Rain	Rain Reflectivity	Rain Reflectivity	Cloud-top Temp.
Data Example	 			
Estimation Accuracy	Locally Accurate	Accurate	Accurate	Poor
Coverage Area	Point Location	Regional Scale	Global Scale	Continental Scale
Spatial Availability	Poor	Good (US) Poor (Global)	Good	Very Good
Temporal Availability	Very Frequent	Very Frequent	Infrequent	Very Frequent

In conclusion, each rainfall-measuring instrument has its advantages and limitations over one another, as summarized in Table 1.1. At present, no single instrument can provide accurate and comprehensive rainfall measurement at a usable spatial and temporal resolution over the global scale. The optimal rainfall data must be obtained by combining many sources of rainfall measurement together. However, merging these measurements with different characteristics is not a straightforward task. It requires well-established knowledge of the rainfall process and understanding of measurement error characteristics from each instrument.

1.2.2 Rainfall Models and Simulations

Rainfall is a complex environmental variable that is difficult to describe either deterministically or statistically. It is affected by turbulent and chaotic physical processes, varies over a wide range of spatial and temporal scales, and it is intermittent. The rainfall process can couple with other environmental variables in complex manners and may never be fully understood. However, needs for rainfall data in many applications motivate us to search for a model that can accurately and effectively simulate and estimate rainfall at spatial and temporal resolutions of interest. A rainfall model is required to propagate rainfall information temporally and/or spatially.

There are countless rainfall models in the literature. Some aim to capture long-term variations, some aim to provide rainfall information where no or insufficient rainfall data are available, and some aim to simulate and forecast rainfall into the future. Rainfall models may be classified into two categories: 1) meteorological rainfall models, and 2) stochastic rainfall models.

Meteorological models seek for a complete physical description of the rainfall process by accounting for dynamical and thermo-dynamical relationships of the atmosphere [81]. Meteorological models are the most sophisticated and computationally demanding among the two types of rainfall models. However, they provide relatively accurate rainfall estimation and are commonly used for short term rainfall forecasting. Meteorological models that are used in short term forecasting are referred to as Quantitative Precipitation Forecasting (QPF) systems. The popular QPF models used in the United State are the Pennsylvania State University- National Center for Atmospheric Research Mesoscale Model (MM5) [22, 51], the National Center for Environmental Prediction's Eta model [11, 12, 115], and the Regional Atmospheric Modeling System (RAMS) [95, 102].

Stochastic models attempt to capture rainfall characteristics in space and/or time using only a few parameters. Statistical rainfall models that use past statistics of rainfall to estimate trends into the future may be classified in this group as well. Stochastic models are useful for representing general spatial and temporal trends or correlations among many climate variables. They can be used to provide short term forecasts but generally are relatively less accurate than the forecasts obtained from meteorological models.

Stochastic models can also be further categorized into three types: 1) spatial, 2) temporal, and 3) spatiotemporal stochastic models. Spatial stochastic models are static models that describe the spatial pattern of rainfall mostly using scaling methods. These include multi-scale, multi-fractal and cascade models [47, 54, 75, 97, 117]. Spatial stochastic models are useful for merging multi-resolution rainfall data obtained at the same measurement time. However, they cannot deal with temporal dynamics of rainfall. Conversely, temporal stochastic models focus on the dynamics of rainfall process at one or multiple discrete locations. Examples of temporal stochastic models include the single-site Bartlett-Lewis rectangular pulse model [67, 129], the Neyman-Scott rectangular pulse model [25, 41], and single or multi-site temporal models [17a, 18b, 119]. They provide reasonable characteristics of temporal dynamics of rainfall but fail to characterize spatial features of rainfall. Finally, spatiotemporal stochastic models account for both spatial and temporal characteristics of rainfall process [94]. Spatiotemporal stochastic models are normally adapted from the spatial and temporal stochastic models.

1.2.3 Rainfall Data Assimilation

Data assimilation is a data merging technique used to combine measurement information with prior knowledge from a dynamic model to produce an analysis state. This technique is useful if measurements are scattered or indirectly related to the variables of interest. In this case, the dynamic model will provide a flow of information from local measurements to all variables in the space and time of interest. Data assimilation can be used to merge multiple sources of rainfall measurements together. By selecting a suitable rainfall model to propagate information, we can ultimately obtain comprehensive rainfall at spatial and temporal resolution of interest.

The concept of data assimilation can be illustrated by a simple calculus problem. Suppose there are two pieces of information, x_1 representing forecast from a model, and x_2 representing measurement of the state x. What combination of x_1 and x_2 gives the best estimate of true x? The answer depends on uncertainty of x_1 and x_2, statistically described by variances σ_1^2 and σ_2^2, respectively. The analysis state or the best estimate of x that give the least uncertainty, denoted by $\hat{x}$, is given by the weighted sum of x_1 and x_2 in the following form,

$$\hat{x} = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} x_1 + \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} x_2 \tag{1.1}$$

The estimate $\hat{x}$ will be optimal if the probability density of x_1 and x_2 are Gaussian, (e.g., full probabilistic characterization can be described only by the mean and (co)variance.).

Over the past decades, many data assimilation algorithms proposed. These algorithms can drastically differ in their formulations, efficiency, concepts and appropriateness for specific applications. Well-known data assimilation algorithms include the 3DVAR and 4DVAR methods [23, 24, 122, 123], the representer methods [9, 10], the Kalman Filter and Extended Kalman Filter methods [35, 43, 86], the Ensemble Kalman Filter methods

[16, 36, 38, 39], and the Particle Filter [6, 100, 132]. Among the many algorithms, those based on the ensemble Kalman Filter (EnKF) may be the most commonly used in hydrologic community. It is simple and efficient, yet accurate enough for many applications. The particle filter is more accurate for dealing with very complex and non-linear systems; however, its computation expense makes the algorithm impractical to apply for a large data assimilation problem.

1.3 Comprehensive Rainfall Ensemble for Land-Surface Models

Many recent climate forecasting and land-surface models are based on ensemble approaches where many possible outcomes are simulated [33, 38, 74, 77, 84, 108, 136]. These models usually require a probabilistic characterization of rainfall over an extended time period and spatial domain. This can be obtained from the ensemble of rainfall conditioned on the available measurements. Therefore, a single realization of comprehensive rainfall data is no longer sufficient. We need a way to generate rainfall ensembles and feed them to ensemble-based climate and hydrological models. A large number of land-surface model studies create the input rainfall ensemble by simply perturbing each rainfall pixel randomly or using a simple statistical model that does not provide realistic correlation of rainfall in space and time [33, 84, 108]. An unrealistic rainfall ensemble can degrade the outputs from these models or may cause the model to become unstable. On the other hand, other applications use rainfall ensembles generated from a QPF models [19, 20, 95]. Although, a rainfall ensemble produced by a QPF model is realistic, it is time consuming and impractical for large scale problems. Therefore, having a fast, efficient and reliable algorithm to merge and provide comprehensive rainfall ensembles at desired spatial and temporal resolution would greatly benefit climate system studies.

1.4 Thesis Outline

The outline of this thesis is given as following. In chapter 2, we introduce the atmospheric forcing and rainfall measurements across the United States Great Plains (USGP) region. These measurement sources include the Geostationary Operational Environmental Satellite (GOES), the NOWRAD precipitation product, the Automated Surface Observing Station (ASOS), the Tropical Rainfall Measuring Mission (TRMM), the Special Sensor Microwave Imager (SSM/I), and the Advance Microwave Sounding Unit-B (AMSU-B). This study area and these measurements will be used to illustrate the rainfall data assimilation technique. We also propose a multi-resolution alignment method to estimate position error statistics and a regression method to estimate the intensity error statistics, which are required in the data assimilation framework.

In Chapter 3, we provide details of the Recursive Clustered Rainfall (RCR) for propagating rainfall ensemble through space and time. The chapter begins by introducing the cluster-point process rainfall model and forming the RCR model by assuming the Markov properties. The GOES cloud-top temperature measurements are then used to handle rainfall intermittency. In addition, we employ multi-resolution alignment with two consecutive cloud images to estimate the velocity field. The velocity field is then used to advect rainfall. Finally, we use the RCR model to propagate rainfall over the USGP region.

Chapter 5 presents a dynamic rainfall data assimilation, which is the core topic of the thesis. The approach is the widely applied Ensemble Kalman Filter (EnKF). We also improve the stability of the EnKF for small ensembles by utilizing the stable pseudo-invert technique. Next, the Ensemble Kalman Smoother (EnKS) algorithm is used in order to incorporate measurements taken later than the estimate time. The EnKS may be more practical in the reanalysis applications. Then we utilize the state-augmentation technique to estimate for the unknown parameters in the RCR model. Lastly, we perform the dynamic rainfall data assimilation using the RCR model, the EnKF, and the EnKS algorithm and provide comprehensive rainfall ensembles over the USGP region.

Chapter 6 is a standalone section. It presents the Expectation-Maximization technique on the Scale-Recursive Estimation framework (EM-SRE) to estimate the multiplicative cascade rainfall model parameters. This rainfall model is well known for its ability to provide spatial characteristic of rainfall field and can be used to statically merge multiple sources of rainfall measurement given at the same time. We present the general form of the Scale-Recursive Estimation (SRE) algorithm for estimating static rainfall and employ the multiplicative cascade rainfall model into the SRE framework. In addition, we propose a tree pruning technique to deal with rainfall intermittency, as well as the Expectation-Maximization (EM) algorithm to estimate the scaling parameters on the tree efficiently. The identifiability and uniqueness of the scaling parameters are emphasized. Finally, we apply the EM algorithm to estimate the scaling parameters from real rainfall estimate inside the USGP region.

Finally, this thesis conclude with Chapter 6 by summarizing the major contributions of rainfall data assimilation and suggesting possible future research directions associated with the work presented in the preceding chapters.

Chapter 2

The United States Great Plains Case Study

2.1 The United States Great Plains (USGP) Region

The United States Great Plains (USGP) case study is set up to study large-scale hydrological process. It is used here to demonstrate the utility of the rainfall data assimilation technique in providing the forcing for a land-surface soil moisture estimation study [136]. The spatial domain is delineated by the United States Geological Survey (USGS) hydrological unit boundaries as shown in Figure 2-1. The region is located between 25.85°N to 49.01°N latitude and 114.07°W to 90.12°W longitude. The time interval of interest is from June 1st, 2004 at 0:00 GMT to August 31st, 2004 at 23:00 GMT. All rainfall and weather related variables used in the USGP project are interpolated to latitude/longitude grid. The spatial resolution depends on the native resolution of measurements but is not finer than 0.05°. The temporal resolution will be rounded to a 15-minute interval. The temporal resolution is chosen to facilitate the data assimilation procedure, especially for continuous measurements from satellite-based instruments.

The large size of the study region is chosen to allow examination of the scaling properties and intermittency effects of rainfall, while the time window is chosen to avoid snow and ice. In addition, the USGP region is relatively flat. We exclude the mountainous areas across the West in order to minimize topography effects on the accuracy of ground-based measurements. The measurements over the USGP region will be provided over a rectangular area containing the USGP boundary; however, the evaluation of the measurements and rainfall assimilation results will only be done over the shaded area in Figure 2-1.

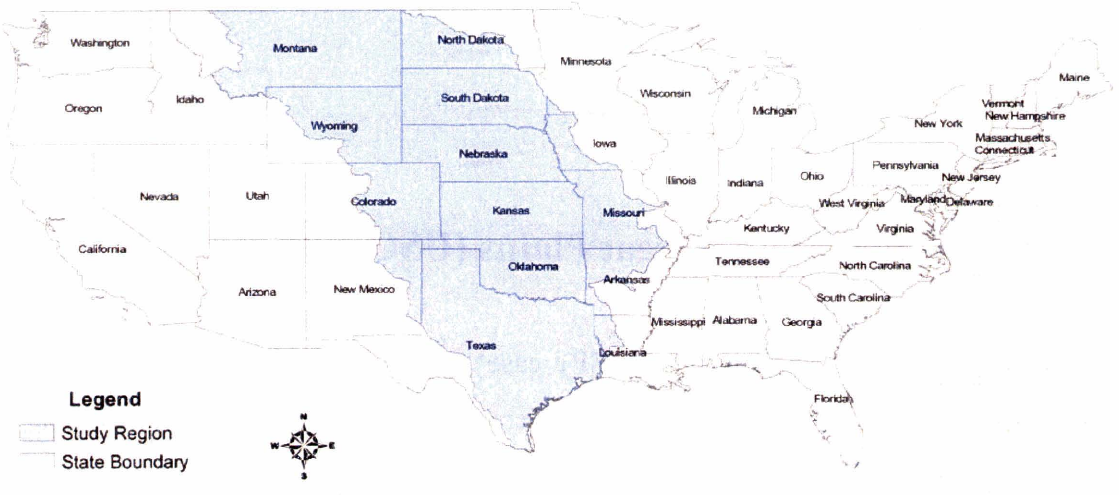


Figure 2-1: The United States Great Plains (USGP) case study region

2.2 Atmospheric Forcing and Rainfall Measurement in the USGP Region

The United States Great Plains (USGP) case study incorporates one atmospheric forcing and five rainfall measurement sources. The atmospheric forcing is the cloud-top temperature images from the Geostationary Operational Environmental Satellite (GOES). The remaining five rainfall data sources are the NOWRAD precipitation product, the Automated Surface Observing Station (ASOS), the Special Sensor Microwave Imager (SSM/I), the Tropical Rainfall Measuring Mission (TRMM), and the Advance Microwave Sounding Unit (AMSU). The utilization of these measurements for the USGP case study, their units and resolutions are illustrated in Table 2.1. The following subsections will provide detailed descriptions, the spatial and temporal characteristics, and the process used to convert measurements to the format used in the USGP project.

Table 2.1: Measurements summary and roles in the USGP case study

Sources	**Utilization in the USGP case study**	**Unit**	**Resolution**	
			spatial	**temporal**
1. GOES	Implement the rainfall model	Kelvin	0.05 °	1 hr
2. NOWRAD	Ground-truth for validation	mm/15min	0.05 °	15 min.
3. ASOS	Gauge Rainfall Measurement	mm/hr	0.05 °	1 hr
4. TRMM	Satellite Rainfall Measurement	mm/hr	0.05 °	2 /day*
5. SSM/I	Satellite Rainfall Measurement	mm/hr	0.25 °	6 /day*
6. AMSU	Satellite Rainfall Measurement	mm/hr	0.15 °	6 /day*

Note: * Revisit time of the satellite-based measurement

2.2.1 The Geostationary Operational Environmental Satellite (GOES)

The Geostationary Operational Environmental Satellite (GOES) is a weather satellite system operated by the National Oceanic and Atmospheric Administration (NOAA). It provides key information on a short-range weather warning and forecasting for the United States. The GOES system consists of two main meteorological satellites: the GOES-West positioned at 135 °W to monitor the Pacific ocean and western United States, and the GOES-East positioned at 75 °W to monitor most of North and South America and Atlantic ocean. The coverage area of these two geostationary satellites is illustrated in Figure 2-2. These satellites encircle the Earth in a geosynchronous orbit at the same speed of the Earth's rotation, so they are capable of continuously monitoring the atmospheric variables. Each satellite carries two main instruments to observe atmospheric and weather condition. First, the imager instrument measures radiant and reflected solar energy from the Earth's surface and atmosphere in the visible and infrared spectrum. Second, the sounder unit provides the vertical profiles of important variables such as temperature, moisture, and ozone distribution.

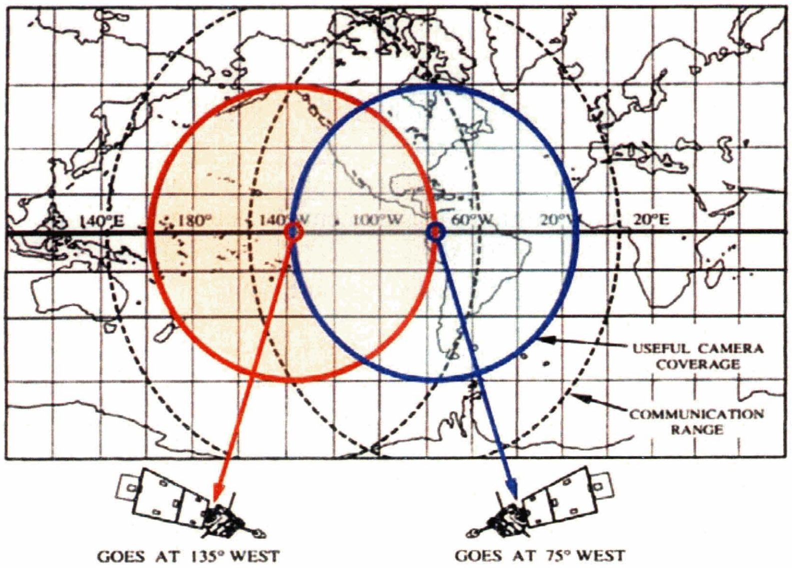


Figure 2-2: Data coverage of GOES East and GOES West spacecraft [58]

The GOES satellites provide cloud-top temperature images for the USGP project. The raw dataset is from the infrared imager, which provides the radio-brightness temperature at a wavelength range between 10.23μm to 11.24μm and centered at 10.7μm. This dataset is known as the long-wave infrared channel or the "window channel". The radiation at this wavelength range is not significantly absorbed by atmospheric gases and can represent actual temperature with minimal interference. This dataset is widely used to determine cloud-top heights and to track synoptic or mesoscale features.

The GOES raw data is obtained from the National Oceanographic and Atmospheric Administration (NOAA) Comprehensive Large Array-data Stewardship System (CLASS). The data is in a netCDF format at a spatial resolution of approximately 2.3 x 4 km^2 (E/W x N/S), and at a temporal resolution of roughly 30-minute intervals. We obtain the primary source of GOES data from GOES-12 satellite, and use data from GOES-10 to fill in missing measurements. Because significant amounts of the 30-minute interval dataset are corrupt or missing, we use the GOES dataset at a temporal resolution of 1 hour instead. GOES cloud-top temperature data is interpolated to fit the latitude-longitude grid at 0.05° resolution, as illustrated in Figure 2-3.

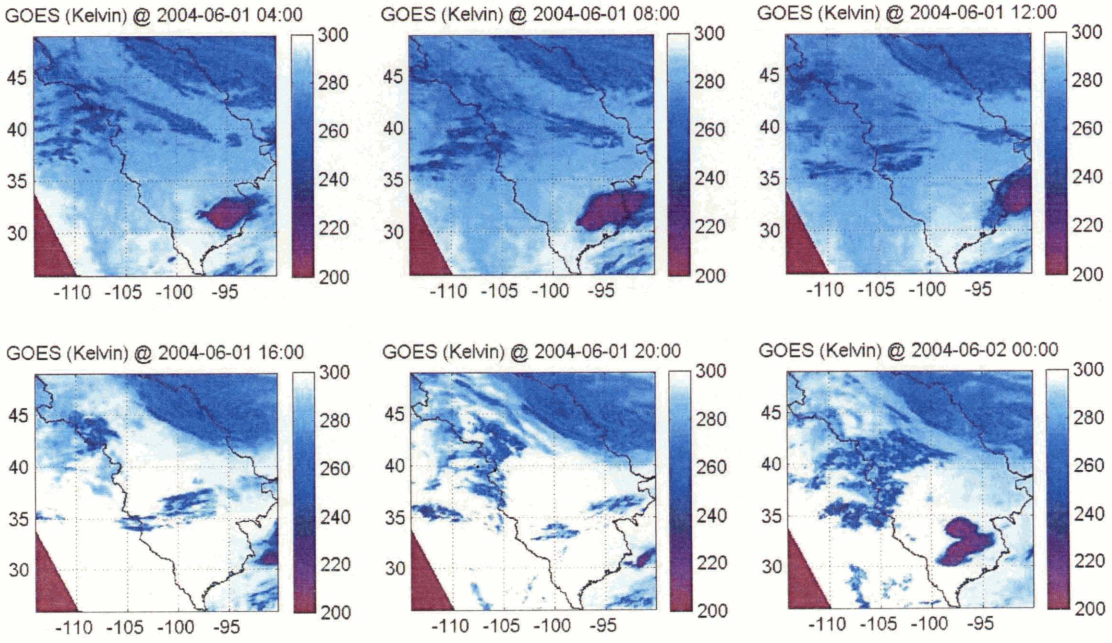

GOES (Kelvin) @ 2004-06-01 04:00
GOES (Kelvin) @ 2004-06-01 08:00
GOES (Kelvin) @ 2004-06-01 12:00
GOES (Kelvin) @ 2004-06-01 16:00
GOES (Kelvin) @ 2004-06-01 20:00
GOES (Kelvin) @ 2004-06-02 00:00
45
40
35
30
-110 -105 -100 -95
300
280
260
240
220
200

Figure 2-3: Examples of GOES infrared cloud-top temperature from the USGP case study

The USGP case study uses the GOES dataset to provide information about cloud location and thickness for a rainfall model. Although the data is comprehensive in space and time, it will not be used directly to estimate rainfall intensity because various attempts to convert the cloud-top temperature to rainfall measurements [1, 5, 130, 135] fail to provide accurate results. However, when it is used in conjunction with other sources of rainfall measurements, they provide much more reliable estimates of the potential presence of rainfall [62, 86, 90, 125]. Usages of the cloud-top temperature in the dynamic rainfall model will be discussed in Chapter 4.

2.2.2 The NOWRAD Precipitation Product

The NOWRAD products are value-added commercial data created by the Weather Services International (WSI) Corporation. The dataset is based on the Weather Surveillance Radat-1988 Doppler (WSR-88D) system from the Next-Generation Weather Radar (NEXRAD) program. However, the commercial NOWRAD products are subject to a great degree of quality control in comparison to the raw NEXRAD measurement. The details of the algorithm used to enhance raw NEXRAD imagery to produce NOWRAD data are proprietary, but general concepts are given in [50]. NOWRAD rainrate significantly depends on the reflectivity-rainfall rate (Z-R) relationship; hence, it may be less accurate than estimates based on sophisticated algorithms that include adjustment for reflectivity, vertical profile, visibility, attenuation, and reflectivity rainfall rate variations, etc. [44, 45]. Nevertheless, the NOWRAD rainfall product is reasonably accurate for a large scale rainfall reanalysis application [50]. In addition, the NOWRAD data set is much more comprehensive in time and space over the continental United States than other measurements of rainfall. Therefore, it is widely used in many news and media channels, consulting companies, and research projects.

The NOWRAD rainfall dataset for the USGP case study is obtained from the Atmospheric and Environmental Research (AER), Inc[1]. The raw dataset represents 15-minute cumulative rainfall over the continental United States on a 2 km regular grid. The precision of the product is at 0.254 mm, and the maximum 15-minute cumulative rainfall estimate allowed is approximately 20 mm. In the USGP project, we spatially interpolate NOWRAD rainfall data to a 0.05° latitude-longitude grid at a temporal resolution of 15 minutes. Examples of NOWRAD rainfall measurement over the USGP region are illustrated in Figure 2-4.

[1] NOWRAD dataset is provided by Dr. Ross Hoffman and Dr. Christopher Grassotti for educational or research purposes only.

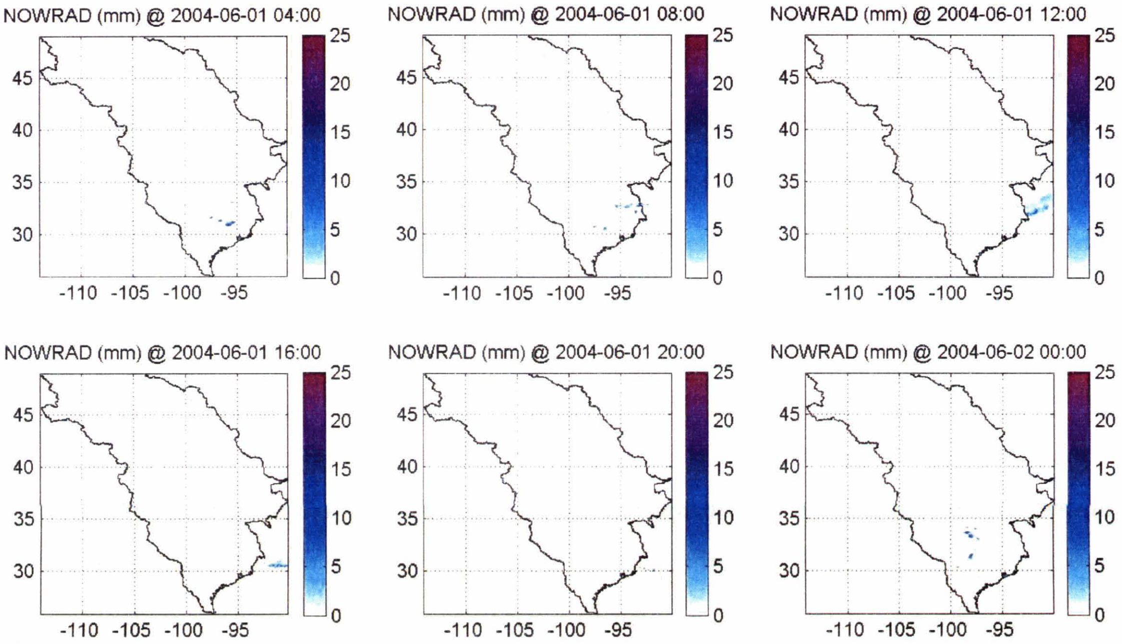

NOWRAD (mm) @ 2004-06-01 04:00
NOWRAD (mm) @ 2004-06-01 08:00
NOWRAD (mm) @ 2004-06-01 12:00
NOWRAD (mm) @ 2004-06-01 16:00
NOWRAD (mm) @ 2004-06-01 20:00
NOWRAD (mm) @ 2004-06-02 00:00
45
40
35
30
-110 -105 -100 -95
25
20
15
10
5
0

Figure 2-4: Examples of NOWRAD 15-minute cumulative rainrate from the USGP case study

The NOWRAD rainfall product is relatively accurate and comprehensive in space and time. When the NOWRAD data is available, it is unnecessary to employ the data assimilation technique to merge multiple measurements in order to provide comprehensive rainfall data because the NOWRAD data will dominate the other measurement sources. Hence, it is more appropriate to exclude the NOWRAD when utilizing the data assimilation scheme and instead employ it to validate the results. The exclusion of the NOWRAD data in the data assimilation problem makes sense because rainfall measurement with the quality and availability of NOWRAD is very rare. It is important to note that the accuracy and coverage of NOWRAD data drops drastically in mountainous regions, e.g. over the Rockies on the west of USGP region, as described by [78]. Consequently, we will use only the NOWRAD data strictly inside the USGP boundary.

2.2.3 The Automated Surface Observing Station (ASOS)

The Automated Surface Observing Station (ASOS) is a surface observing network used primarily for weather forecast activities and aviation operation in the United States. The ASOS program is a joint effort of the National Weather Service (NWS), the Federal Aviation Administration (FAA), and the Department of Defense (DOD). It provides basic micrometeorological measurements including precipitation accumulation. There are other surface observing networks in the United States as well. Examples are NWS Cooperative Station Network (COOP) and the World Meteorological Organization stations (WMO). However, these networks provide less consistent measurements compared to the ASOS network.

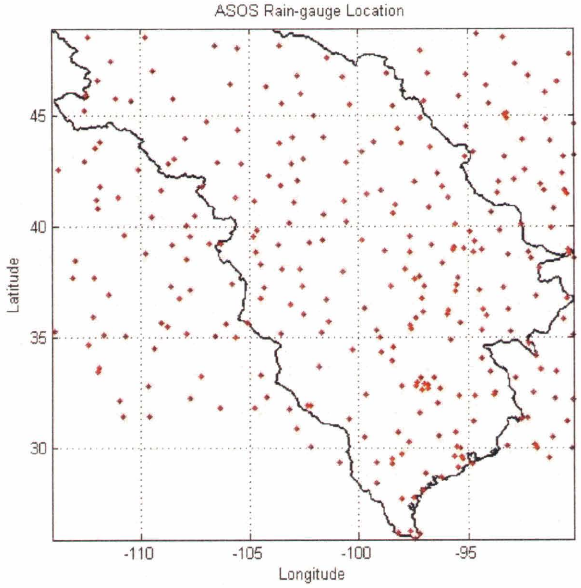


Figure 2-5: The ASOS locations over the USGP study domain

There are 308 ASOS stations between the latitudes of 25.85°N – 49.01°N and the longitudes of 114.07°W – 910.12°W. The rectangular area contains the USGP domain. The locations of these ASOS stations are indicated by red symbols in Figure 2-5. For the USGP project, we are interested in the cumulative rainfall dataset. This dataset is available at the end of every hour at accuracy of 0.25 mm. We use point measurement from ASOS station to represent average rainfall over the 0.05° grid call where the station is located.

The samples of cumulative rainrate from ASOS stations are shown in Figure 2-6. Each plot shows spatial interpolation of rainfall from 308 ASOS stations over the rectangular are surrounding the USGP region. The interpolation is done using the simple plate metaphor by solving a direct linear system of equations for missing cells. The red crosses represent ASOS stations with no rainfall detected, while the blue crosses represent ASOS stations with positive rainfall measurements. Note that the spatial interpolation of ASOS data in Figure 2-6 is just for the illustration purposes. In the rainfall data *assimilation*, we only use the point measurement at the ASOS stations.

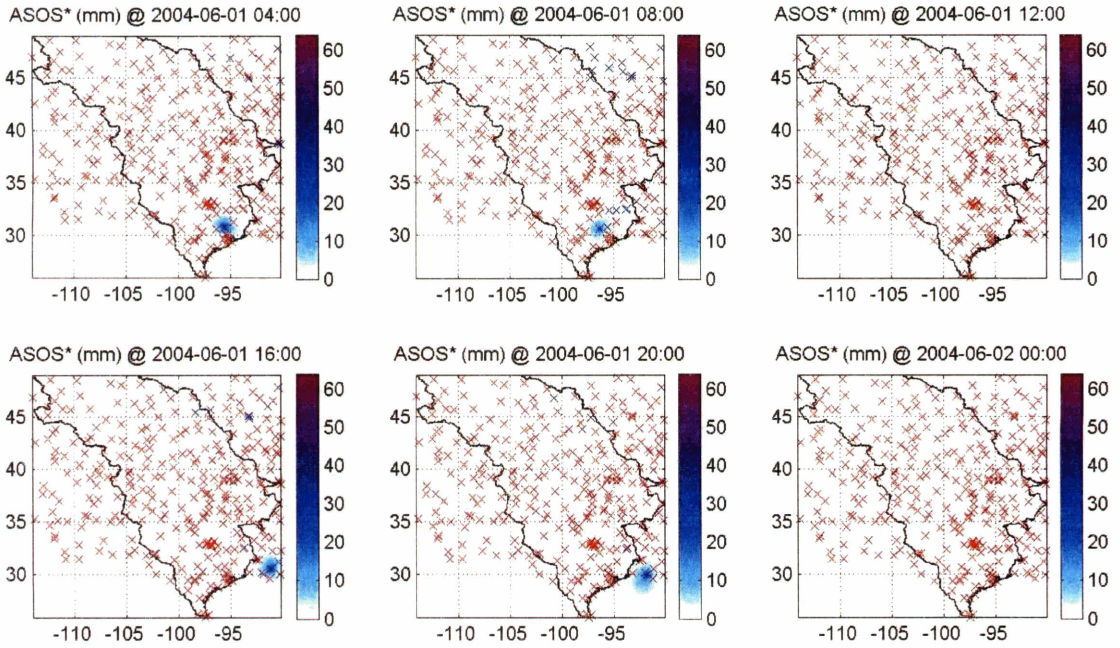

ASOS* (mm) @ 2004-06-01 04:00
ASOS* (mm) @ 2004-06-01 08:00
ASOS* (mm) @ 2004-06-01 12:00
ASOS* (mm) @ 2004-06-01 16:00
ASOS* (mm) @ 2004-06-01 20:00
ASOS* (mm) @ 2004-06-02 00:00
45
40
35
30
-110 -105 -100 -95
60
50
40
30
20
10
0

Figure 2-6: Examples of ASOS cumulative rainrate interpolated over the USGP study domain

2.2.4 The Tropical Rainfall Measuring Mission (TRMM)

The Tropical Rainfall Measuring Mission (TRMM) is a cooperative satellite mission between United States and Japan to monitor rainfall and other related atmospheric variables in the tropical and subtropical regions. The TRMM satellite is in circular, non-synchronous orbit with an inclination of 35 degrees relative to the Equator. The satellite provides meteorological data between the latitudes of 138°S and 38°N with the revisit time of about 12 hours. There are several instruments onboard the TRMM satellite, but the main instruments for measuring rainfall are the TRMM Microwave Imager (TMI), the Precipitation Radar (PR), and the Visible Infrared Scanner (VIRS.) The TMI is a multi-channel passive microwave radiometer and the PR is an active microwave radiometer. Their measurements of radiation are used to estimate integrated column precipitation contents, liquid water, ice, rain intensity, rainfall types, and precipitation layer depth. Finally, the VIRS supplies information on cloud coverage, cloud type and cloud-top temperature.

The raw data from TRMM satellite is processed by the TRMM Science Data and Information System (TSDIS), and the post-process data is distributed by Distributed Active Archive Center (DAAC). TRMM standard products consist of three levels. The level-1 products are mainly raw observations such as VIRS-calibrated radiances, the TMI brightness temperatures and the PR reflectivity measurements. The level-2 products contain derived geophysical variables including rainrate, rain type, and drop size distribution at the same resolution and location as the level-1 observations. Lastly, the level-3 products are temporal and spatial climatology of geophysical variables projected onto a uniform space-time grid.

For the USGP project, we are interested in the surface rainrate from TRMM satellite. This rainfall dataset is retrieved using the TRMM Microwave Imager (TMI) profiling algorithm, referred to as 2A-12. The algorithm provides vertical profiles of hydrometeors and instantaneous surface rainrate. The rainrate measurement has a resolution of 5.1 km and a precision of 0.1 mm/hr. The raw data is presented by the satellite track spanning 878 km, across and there are roughly 5 tracks that pass over the USGP region per day. In this format, the measurement time will vary as the satellite progresses along the track. To simplify the dataset, we aggregate all measurement inside a 15-minute interval and express it as a snapshot of instantaneous rainrate at the end of time interval. For example, measurement labeled "*0:15*" represents rainfall measured between the time 0:00 and 0:15. We interpolate the TRMM measurement to fit a latitude-longitude grid at the resolution of 0.05°. Examples of TRMM surface rainfall measurement in the USGP project are given in Figure 2-7.

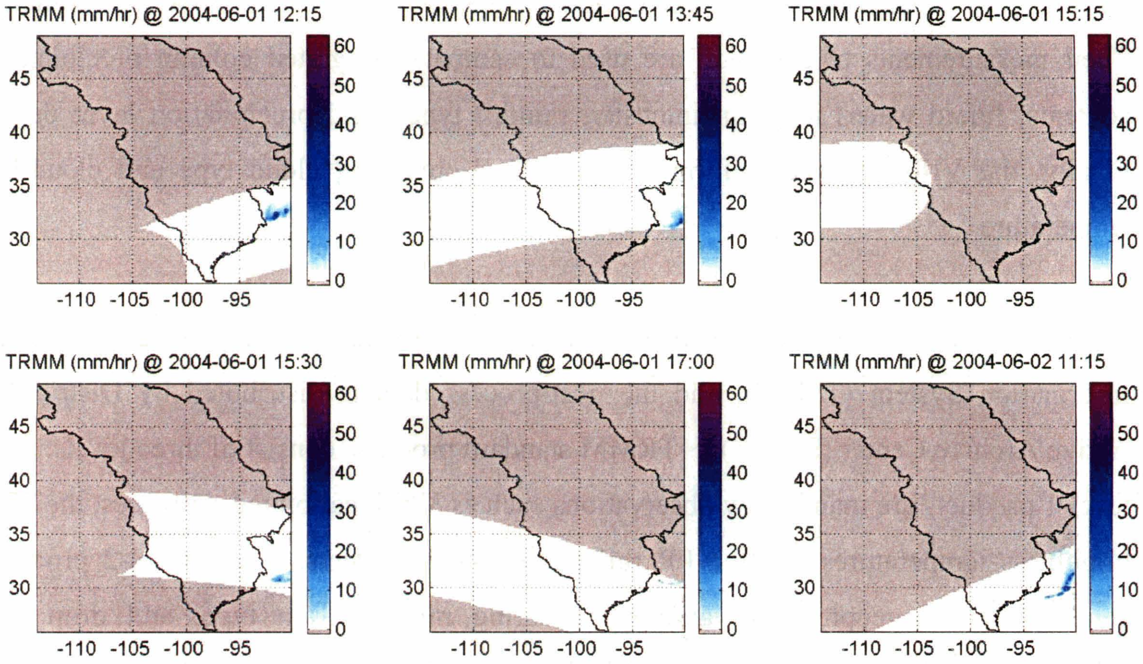

TRMM (mm/hr) @ 2004-06-01 12:15
TRMM (mm/hr) @ 2004-06-01 13:45
TRMM (mm/hr) @ 2004-06-01 15:15
TRMM (mm/hr) @ 2004-06-01 15:30
TRMM (mm/hr) @ 2004-06-01 17:00
TRMM (mm/hr) @ 2004-06-02 11:15
45
40
35
30
-110 -105 -100 -95
60
50
40
30
20
10
0

Figure 2-7: Examples of TRMM instantaneous surface rainrate from the USGP study domain

2.2.5 The Special Sensor Microwave Imager (SSM/I)

The Special Sensor Microwave Imager (SSM/I) is a multi-channel passive radiometer installed onboard the F13, the F14, and the F15 Defense Meteorological Satellite Program (DMSP) platforms. These satellites are in a sun-synchronous polar orbit with a period of 102 minutes monitoring almost the entire globe. The revisit time for each satellite is approximately 12 hours. The SSM/I instrument can penetrate through cirrus clouds and sense radiation emitted and scattered by raindrops and precipitation-sized ice particles. The satellite radiation observations are processed by the Hydrology Data Support Team (HDST) at the NASA Goddard Space Flight Center. The raw data is converted to the surface precipitation using the Goddard Profiling Algorithm (GPROF), which is similar to the TMI profiling algorithm used in TRMM [118].

The SSM/I rainfall dataset in USGP project is acquired from the Distributed Active Archive Center (DAAC) archive. The dataset represents real-time orbit-by-orbit instantaneous rainfall rate at 0.25-degree resolution on an equal latitude-longitude projection at a precision of 0.1 mm/hr. The SSM/I orbital data is presented by the satellite track. There are approximately 15 tracks per day over the USGP region from 3 SSM/I satellites. As with the TRMM dataset, we aggregate SSM/I data measured within 15-minute time intervals to represent a snap-short of instantaneous rainrate at the end of the time interval. Because the original SSM/I data is on an latitude-longitude grid, there is no need to further interpolate the data. Examples of SSM/I instantaneous rainfall images in the USGP project are given in Figure 2-8.

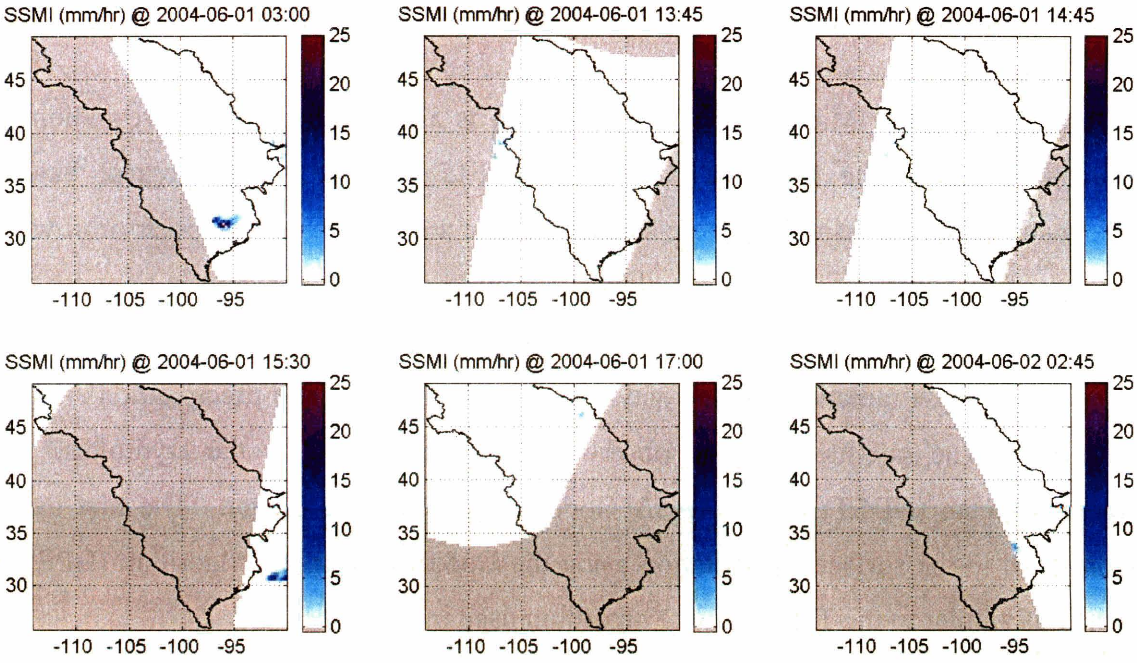

SSMI (mm/hr) @ 2004-06-01 03:00
SSMI (mm/hr) @ 2004-06-01 13:45
SSMI (mm/hr) @ 2004-06-01 14:45
SSMI (mm/hr) @ 2004-06-01 15:30
SSMI (mm/hr) @ 2004-06-01 17:00
SSMI (mm/hr) @ 2004-06-02 02:45
45
40
35
30
-110
-105
-100
-95
25
20
15
10
5
0

Figure 2-8: Examples of SSM/I instantaneous surface rainrate from the USGP study domain

2.2.6 The Advance Microwave Sounding Unit-B (AMSU-B)

The Advance Microwave Sounding Unit-B (AMSU-B) is a new five-channel microwave-sounding instrument developed by the UK Meteorological Office. The instrument is placed onboard the National Ocean and Atmospheric Administration (NOAA) polar orbiting satellites, e.g. NOAA-K, NOAA-L, and NOAA-M. The AMSU-B is deployed to measure radiation from various layers of the atmosphere and estimate global data on humidity profiles. The microwave frequency used by the AMSU-B can penetrate through clouds and provide the signature of rainfall and snow; thus, allowing the instrument to be used to map precipitation.

The AMSU-B instantaneous surface rainrate in the USGP project is the experimental product from the Research Laboratory of Electronics at Massachusetts Institute of Technology [121]. The rainrate estimated from the AMSU-B satellite is validated using the mesoscale numerical weather prediction model (MM5) and the two-stream radiative transfer model (TBSCAT). The final rainfall product is the orbital surface rainrate at 15 km resolution. As with TRMM and SSM/I, we aggregate AMSU-B data measured within 15-minute time intervals and represent it as a snapshot of instantaneous rainrate at the end of the time interval. We spatially interpolate AMSU-B from [121] to the latitude-longitude grid at a 0.15° resolution. Examples of AMSU instantaneous rainfall images in the USGP project are given in Figure 2-9.

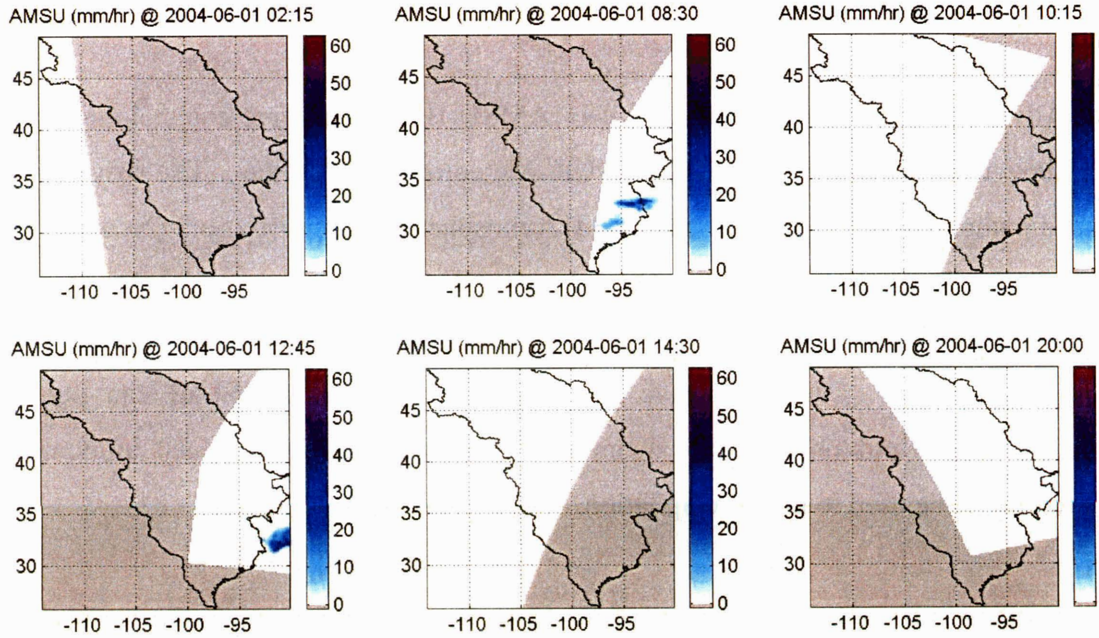

AMSU (mm/hr) @ 2004-06-01 02:15
AMSU (mm/hr) @ 2004-06-01 08:30
AMSU (mm/hr) @ 2004-06-01 10:15
AMSU (mm/hr) @ 2004-06-01 12:45
AMSU (mm/hr) @ 2004-06-01 14:30
AMSU (mm/hr) @ 2004-06-01 20:00
45
40
35
30
-110 -105 -100 -95
60
50
40
30
20
10
0

Figure 2-9: Examples of AMSU-B instantaneous surface rainrate from the USGP study domain

2.3 Rainfall Measurement Error Statistics

In order to utilize the rainfall measurements in the data assimilation framework, it is essential to know its uncertainty described by the error statistics. As seen in (1.1), we need these uncertainties to obtain the best estimate of the state. Measurement uncertainty is generally obtained by field experiment and ground-truth validation processes. In most data assimilation studies, the measurement error statistics are assumed to be known or taken directly from the raw data source. Unfortunately, rainfall data in the USGP project does not include the measurement uncertainty. Therefore, we must estimate these statistics.

The following sections will provide procedures and detailed analysis for estimating error statistics of the USGP rainfall measurements with respect to NOWRAD dataset. We choose the NOWRAD precipitation product to represent the truth because it is relatively accurate over the USGP region. In addition, NOWRAD measurements are comprehensive in space and available at every 15-minute interval.

2.3.1 Measurement Error Classification

Measurement errors can be classified into two groups based on their characteristics: (1) position error, and (2) scale error. First, the position or displacement error is defined by a position difference or an offset between the measurement and the truth. The position error is highly correlated in space. It depends on the shape of the underlying measurement and the truth. The position error may be a result of using an unaligned reference point, using different map projections, or shifting measurement times. Scale intensity error (i.e., amplitude or magnitude error) is defined as a difference in the magnitude between the measurement and the truth when there is no position error. This error is normally assumed to be spatially independent and uncorrelated to the truth.

Detection and correction of position and amplitude errors have been the focus of many recent studies [49, 56, 57, 71, 92]. Because the special characteristics of the position and amplitude errors are very different, these two types of errors should be separated whenever possible. To illustrate their differences, we set up a simple experiment in one-dimension as shown in Figure 2-10. We generate a synthetic truth (red solid line) and create two measurement datasets from it. The first dataset is generated by independently altering the position of the truth (Figure 2-10a), and the second one is generated by altering the magnitude of the truth (Figure 2-10b). The blue lines in Figure 2-10a and 2-10b are the ensemble mean and the cyan lines are the individual replicates. Even though these two datasets have very different characteristics and means, they possess the same statistics of the residual, as shown in Figure 2-10c (e.g., the red line represent a Gaussian distribution with mean and standard derivation given in the title.)

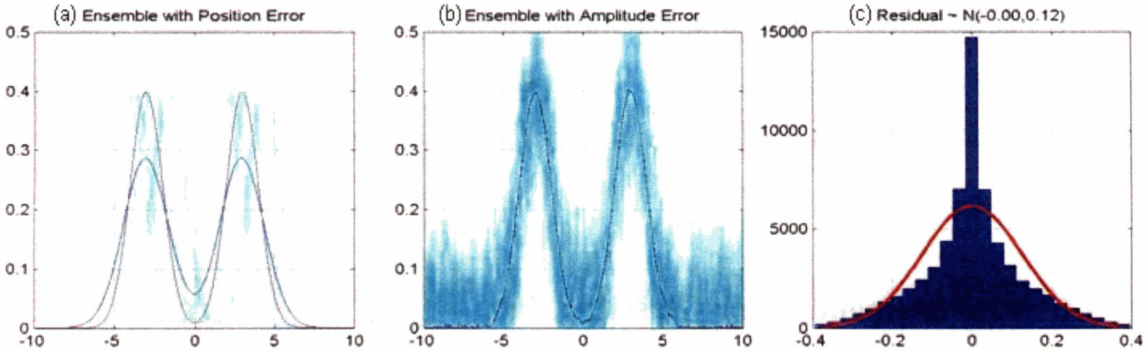


Figure 2-10: A synthetic example of (a) measurement position errors, (b) amplitude errors, and (c) histogram of the residuals (x-axis representing the position and y-axis representing the rainrate)

In the USGP case study, there are four types of rainfall measurements: ASOS, TRMM, SSM/I, and AMSU-B. The ASOS rain-gauge measurements are very scattered in space, and there are only a few measurements with positive rainfall rate at each measurement time. Because it is a point measurement, estimating the position error is more complex than with the other satellite-based measurements. Therefore, in the following section, we will first estimate the error statistics for the satellite-based measurements (e.g., TRMM, SSM/I, and AMSU-B). We will estimate the error statistics of the ASOS rain-gauge data separately.

2.3.2 Position Error Statistics of the USGP Satellite Measurements

The position error of the satellite-based rainfall measurements in the USGP project can be estimated by the position difference between the measurement data and the NOWRAD data. There are several solutions to the position error problem and image alignment [31, 49, 63, 92, 105, 106]. In this study, we employ the multi-resolution alignment (MRA) algorithm [106] to estimate the offset between the NOWRAD and the satellite rainfall measurement. The MRA algorithm aligns satellite measurement field to the NOWRAD data at the same observation time by minimizing the misfit. It calculates the displacement in x- and y-direction, and establishes aligned field that best fit the NOWRAD data. This algorithm is similar to the feature calibration and alignment (FCA) technique [49, 55, 92]. Details and derivation of the MRA algorithm is given in Appendix A.

To illustrate the position errors and the correction after applying the MRA algorithm, the TRMM measurements are plotted against the NOWRAD data on June 3rd, 2004 at 12:00 GMT. Figure 2-11 provides an example of position errors before applying the MRA algorithm. Figure 2-11a shows NOWRAD data (e.g., it is multiplied by 4 to give results in mm/hr, Figure 2-11b shows TRMM measurement data, and Figure 2-11c show the difference between these two measurements. In this example, there is significant position error between the TRMM and NOWRAD datasets. The TRMM dataset is offset to the North-East direction of the NOWRAD data. We then align the TRMM measurements inside the USGP region with the NOWRAD data and plot the difference. Figure 2-12 illustrates the results from aligning TRMM with NOWRAD data on June 3rd, 2004 at 12:00 GMT. Figure 2-11a represents the displacement field (i.e., offset distance) used to align TRMM to NOWRAD, Figure 2-11b shows TRMM measurement inside the USGP region after alignment with NOWRAD data, and Figure 2-11c shows the difference between aligned TRMM and NOWRAD rainfall measurement. By comparing Figure 2-11c with Figure 2-10c, it is evident that position error is minimized after applying the MRA algorithm to align the satellite measurement with its associated NOWRAD measurement. In addition, the displacement field obtained from the MRA algorithm can be used to quantify the amount of position error for each measurement type relative to NOWRAD measurement.

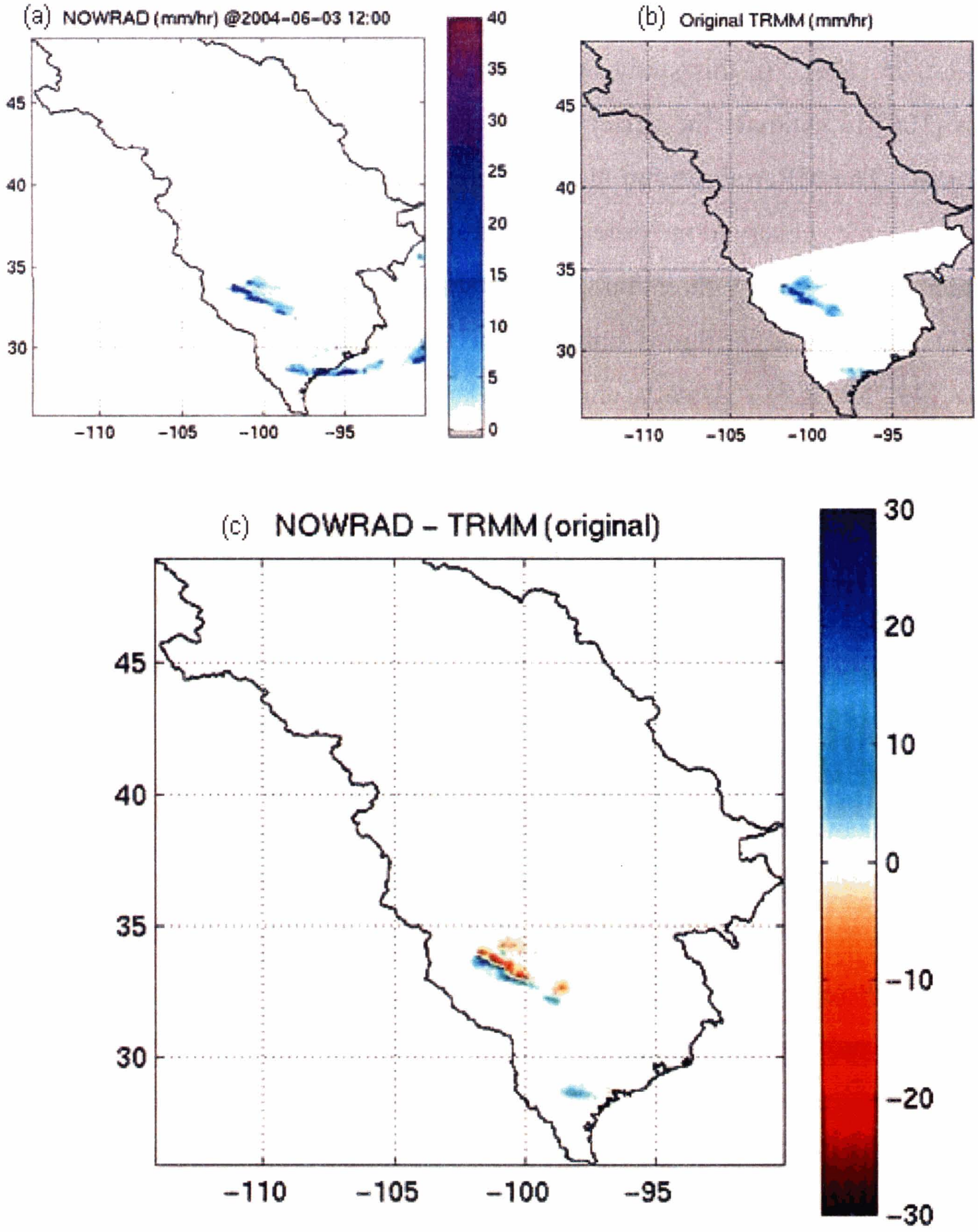


Figure 2-11: Example of TRMM position error relative to NOWRAD measurements over the USGP region on June 3rd, 2004 at 12:00GMT; (a) NOWRAD in mm/hr, (b) TRMM in mm/hr, and (c) difference between TRMM and NOWRAD in mm/hr inside the USGP region

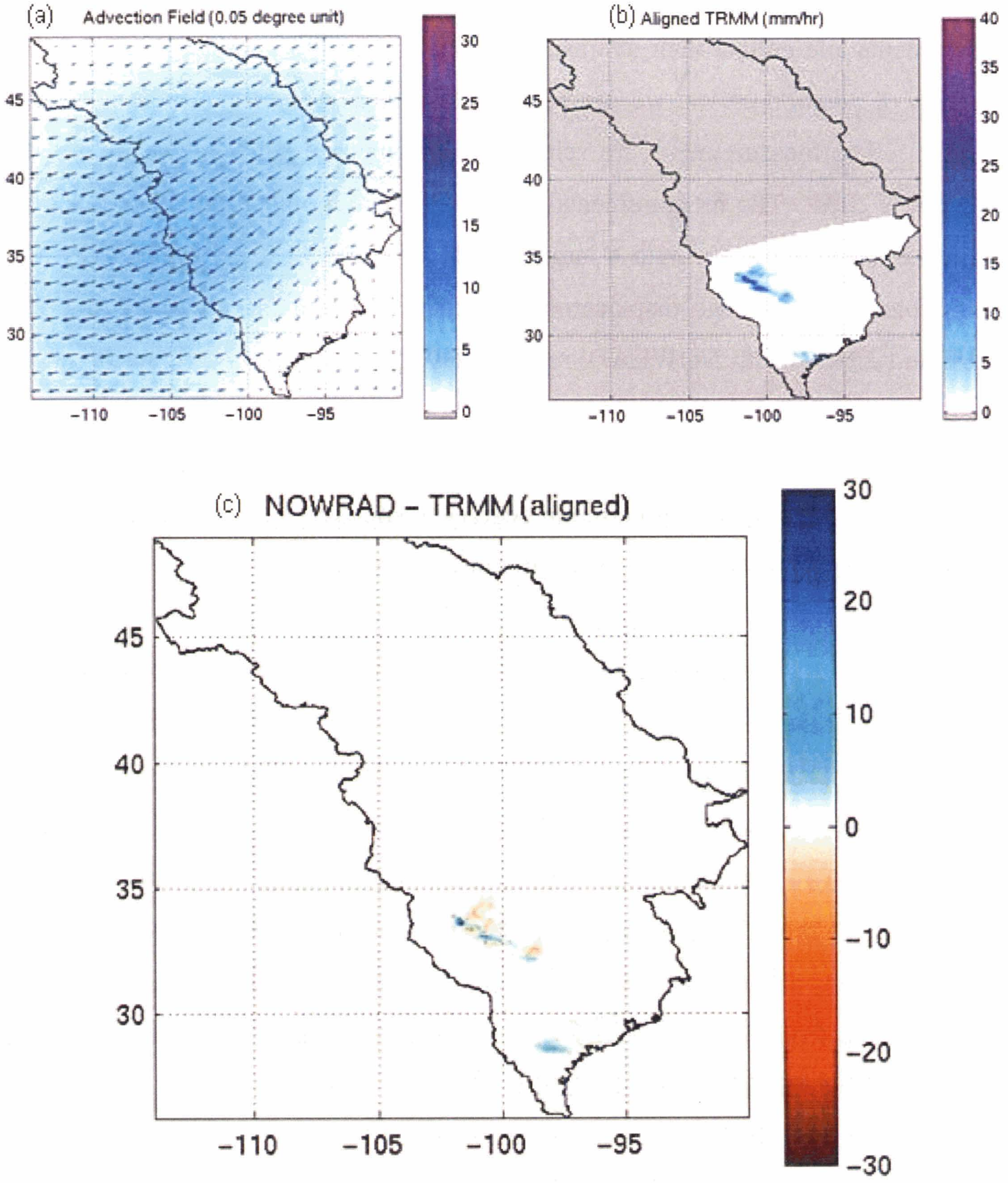


Figure 2-12: Example of TRMM position error relative to NOWRAD measurements *after* applying the multi-resolution alignment algorithm on June 3rd, 2004 at 12:00GMT; (a) displacement field in 0.05°, (b) aligned TRMM measurement in mm/hr, and (c) difference between aligned TRMM and NOWRAD inside the USGP region

To obtain the position error statistics, we repeat the alignment with many satellite measurements and collect their average position error in the x- and y-direction. With a collection of position errors, we can plot the histogram and estimate the position error statistics. The measurements are selected from our case study between June 1st and August 31st, 2004. The measurements used to estimate the position errors must contain significant amounts of rain over a large region. Figure 2-13, 2-14 and 2-15 show the distribution of the average displacement over the rainy pixels after aligning TRMM, SSM/I, and AMSU with NOWRAD, respectively. In each plot, figure (a) shows the joint distribution of the displacement in the x- and y-direction (e.g., Q_x and Q_y) in latitude/longitude degrees. Figure (b) and (c) show the marginal distribution of the displacement in the x- and y-direction in degrees, respectively. The mean and variance of the marginal distribution is given in the title of figure (b) and (c).

From the results, the position errors of the satellite-based instruments relative to NOWRAD are relatively low. The mean position errors in the x- and y-directions are close to zero. In addition, the standard deviation of the position errors depends on the resolution of the measurements (e.g., standard deviation of the TRMM position error < AMSU < SSM/I). The standard deviations are, however, relatively small in comparison to the size of the rainfall events, which normally extended over several degrees. With the joint distributions or the marginal distributions, we can sample the position error in the x- and y-direction from this distribution and perturb the position of the true measurement when performing rainfall data assimilation.

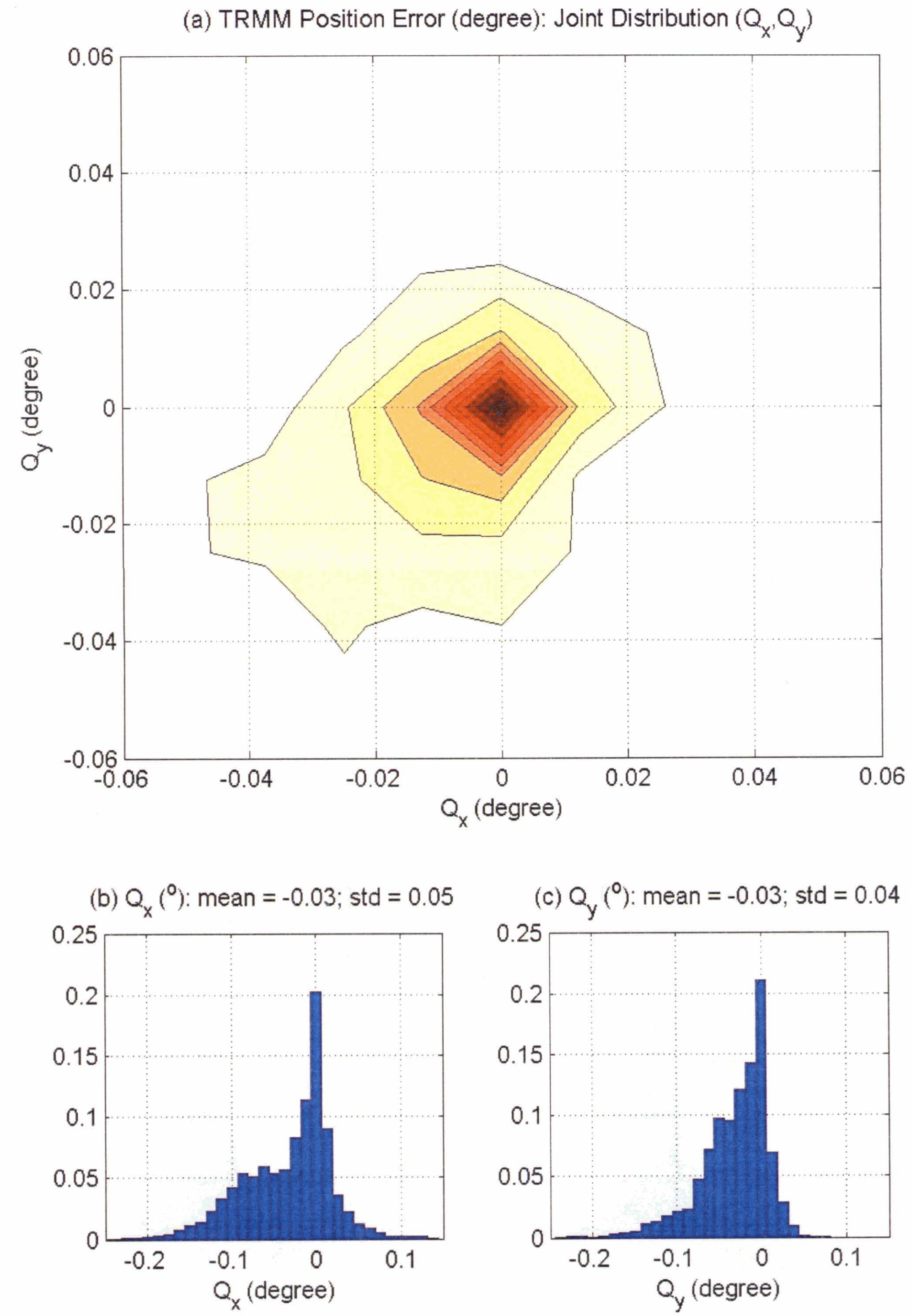


Figure 2-13: Distributions of average TRMM position error over rainy pixels inside the USGP region – (a) joint distribution of the position error, (b) marginal distribution in the x-direction, and (c) marginal distribution in the y-direction

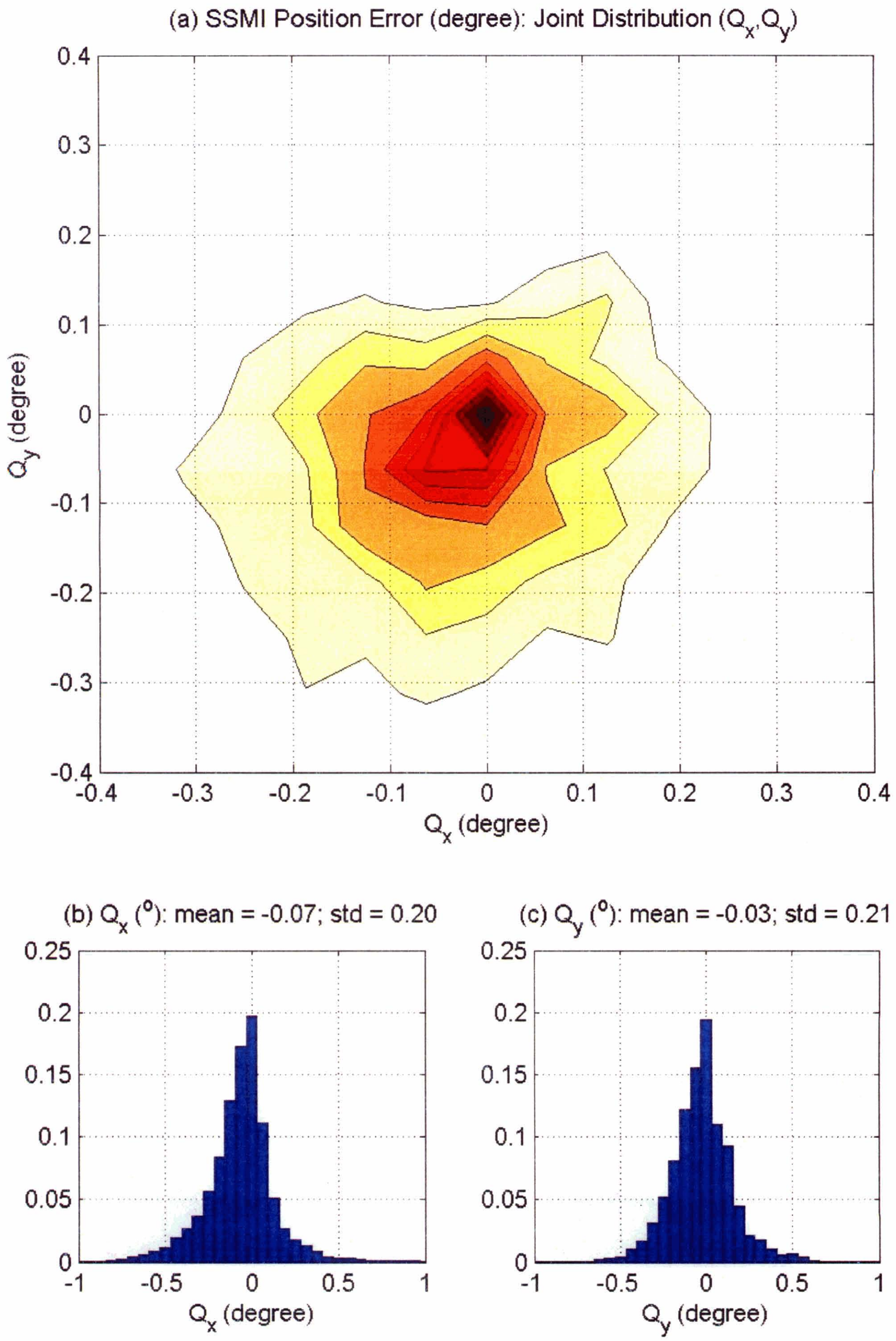


Figure 2-14: Distributions of average SSM/I position error over rainy pixels inside the USGP region – (a) joint distribution of the position error, (b) marginal distribution in the x-direction, and (c) marginal distribution in the y-direction

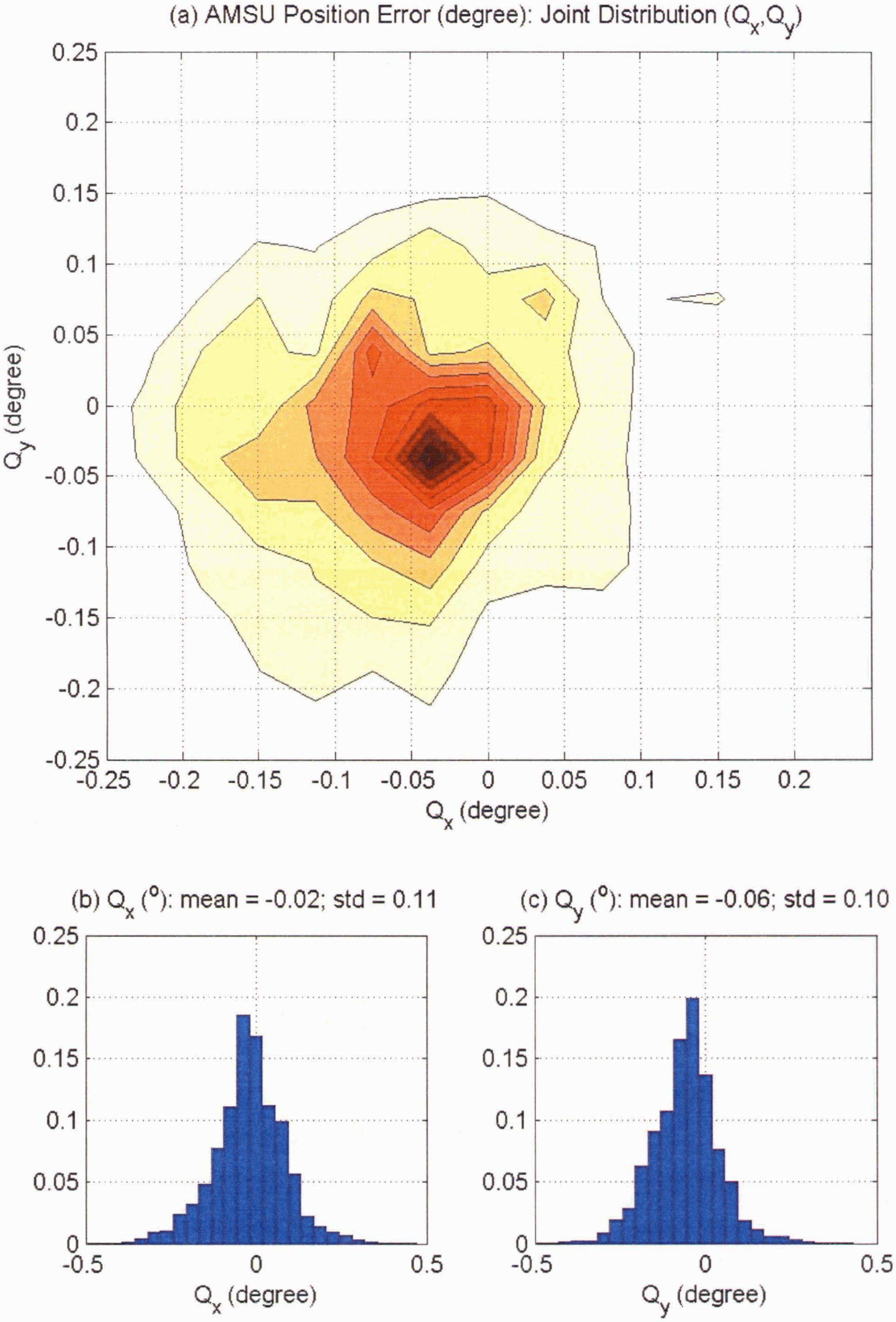


Figure 2-15: Distributions of average AMSU-B position error over rainy pixels inside the USGP region – (a) joint distribution of the position error, (b) marginal distribution in the x-direction, and (c) marginal distribution in the y-direction

2.3.3 Intensity Error Statistics of the USGP Satellite Measurements

Intensity error of the satellite rainfall measurements can be estimated from the difference between the NOWRAD measurement and the aligned satellite measurements. Suppose that at each pixel the aligned measurement (*y*) is related to the corresponding NOWRAD data (*x*) via the equation:

$$y = hx + v \tag{2.1}$$

where *h* is a constant, and *v* is the measurement intensity error, which we assume to have a Gaussian distribution with zero mean and variance σ_v^2. Since we are directly observing the rainfall intensity, the constant *h* should be equal to 1.0. In the USGP project, we can estimate the constant *h* by performing a linear regression between the NOWRAD (on the x-axis) and the aligned satellite measurements (on the y-axis). Because there are many pairs (x,y) near zero, the analysis will be more robust if we weigh the regression by the average measurement value (e.g., *0.5(x+y)*). Figures 2-16, 2-17, and 2-18 present examples of the TRMM, SSM/I, and AMSU scatter plots and weighted regression analysis with NOWRAD data, respectively. In each figure, the image on the left represents the scatter plot with NOWRAD and satellite measurements before the alignment, while the image on the right presents the scatter plot between NOWRAD and satellite measurement after the alignment. The R-squared statistics indicate that the regression using aligned measurement data is much better than using raw measurements before alignment. Note that we coarsened NOWRAD measurements to the same resolution as satellite-based measurements prior to perform the regression analyses.

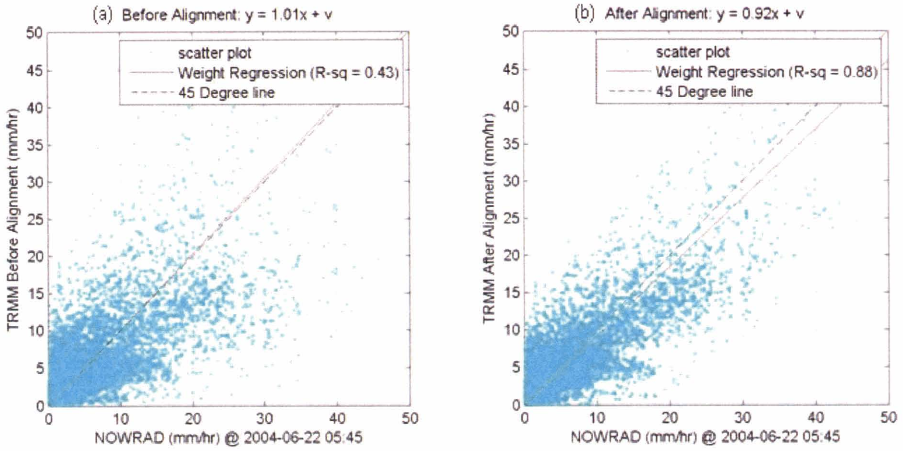


Figure 2-16: The scatter plots and regression analyses of TRMM versus NOWRAD (a) before, and (b) after aligning TRMM with NOWRAD on 2004-06-22 05:45 GMT

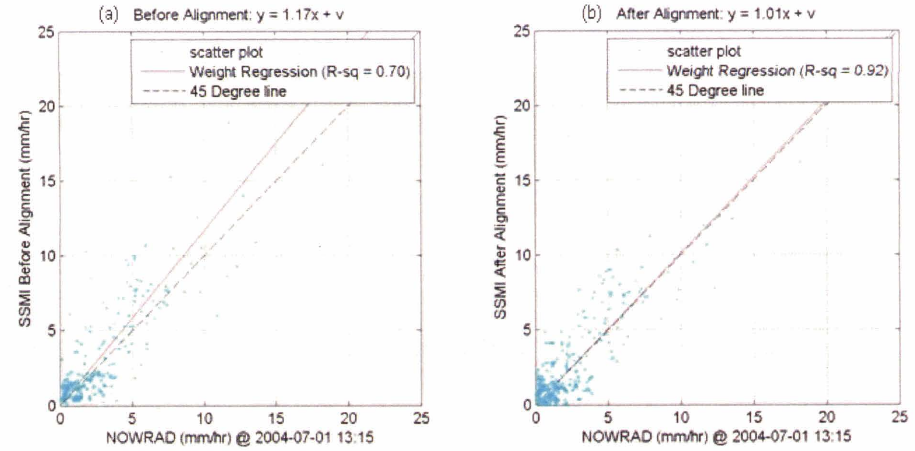


Figure 2-17: The scatter plots and regression analyses of SSM/I versus NOWRAD (a) before, and (b) after aligning SSM/I with NOWRAD on 2004-07-01 13:15 GMT

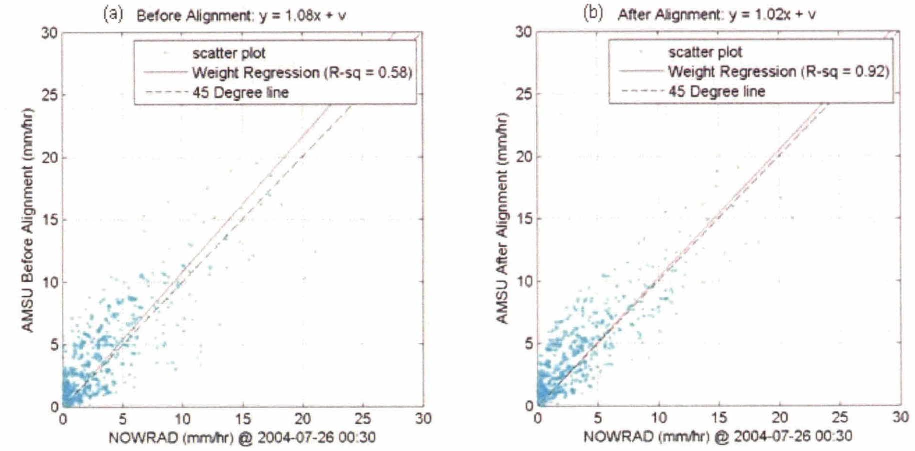


Figure 2-18: The scatter plots and regression analyses of AMSU-B versus NOWRAD (a) before, and (b) after aligning AMSU-B with NOWRAD on 2004-07-26 00:30 GMT

We repeat the regression analysis for many storm events and find the slope of the regression between NOWRAD (x) and aligned satellite measurements (y). These storm events are selected from the USGP case study between June 1st and August 31st, 2004 when there is significant amounts of rainfall. For each satellite source, we select roughly 200 measurements. The distributions of the slope h of TRMM, SSM/I, and AMSU are given in Figure 2-19. The red line in Figure 2-19 represents the Gaussian distribution with mean and variance equal to the sample mean and variance of h. We can perform the hypothesis test and conclude that with 95% confidence the constant h in (2.1) is equal to 1.0 for all satellite sources. Note that we choose to use $h = 1.0$ in order to simplify the problem. It would have been possible to use the mean values from the histograms of each measurement source to represent the constant h for that measurement type.

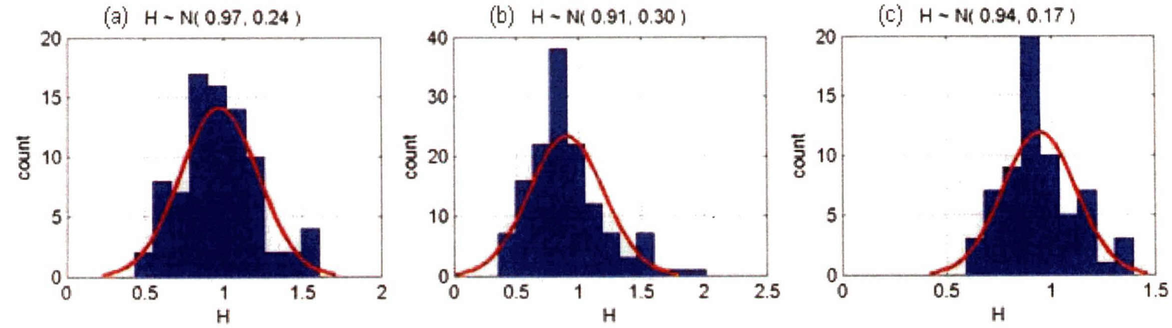


Figure 2-19: Histograms of the slope of regression $y = hx$ with x representing NOWRAD data and y representing (a) TRMM, (b) SSM/I, and (c) AMSU rainfall measurement

In addition to the constant h, we need to define the measurement noise variance σ_v^2 in order to obtain the complete description of the measurements in (2.1). We assume that the standard deviation σ_v is given by

$$\sigma_v = c_1 y + c_2 \tag{2.2}$$

where both c_1 and c_2 are constant. Equation (2.2) implies that the measurement magnitude uncertainty is proportional to the measurement value itself.

To estimate these constants, we obtain the residual from the difference between the aligned satellite-based rainfall measurement and NOWRAD rainrate, e.g. $r = y - x$. Then we bin these residuals according to the value of the satellite measurement rounded to a nearest integer. We repeat the same method over many storm events (the same storm events used to estimate the position errors and the constant *h*). Finally, we perform another regression analysis between the bin center taken from the satellite measurement *(y)* and the standard deviation of the residual (i.e., $\sigma_v = std(r)$). By estimating the slope and intersection of the regression, we can estimate the constants c_1 and c_2 for each measurement source. This regression analysis is used to obtain the constants c_1 and c_2 for TRMM, SSM/I, and AMSU, as shown in Figure 2-20. The analysis implies that the standard deviation of the measurement error is roughly around ¼ of the measurement value given.

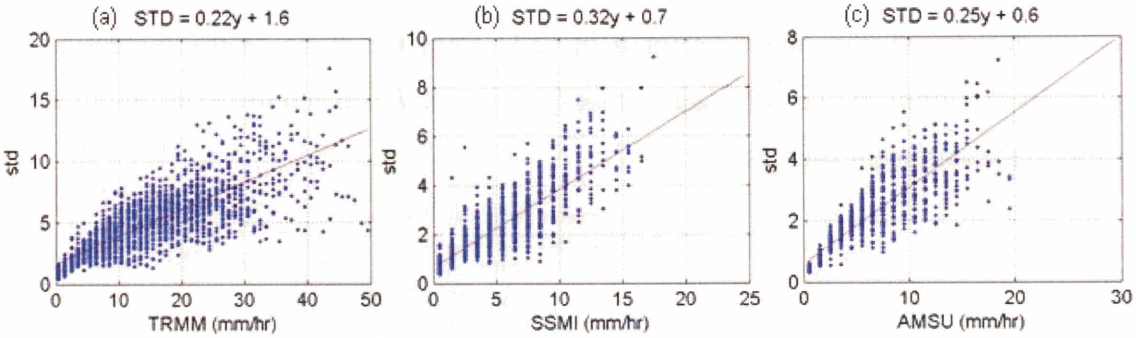


Figure 2-20: The regression analysis for estimating the constants c_1 and c_2 for (a) TRMM, (b) SSM/I, and (c) AMSU measurement

2.3.4 Error Statistics of ASOS Measurement

The cumulative rainfall measurements from the ASOS station network are too sparse to obtain error statistics using the approach applied to the satellite measurement. There are roughly around 300 gauge measurements at each hour from all stations in the rectangular domain engulfing the USGP region. Since NOWRAD is not accurate in the mountainous regions, we only consider the ASOS stations within the USGP boundary. Consequently, the number of rainfall data at each hour reduces to less than 200, and we cannot perform the analysis at each measurement time as we have done for the satellite data.

We aggregate all measurement data from all measurement times in order to have enough measurement points for the regression. The scattered plot of the ASOS and the NOWRAD measurements during June 1st to August 31st, 2004 and the regression analysis are shown in Figure 2-21. We force the regression to pass through the origin. The R-square statistics in Figure 2-21 is very low. This implies that there is no significant correlation to estimate the constant *h* in (2.1) from the slope of the regression. To simplify the problem, we will assume that the constant *h* is equal to 1.0, as with the other rainfall measurement sources.

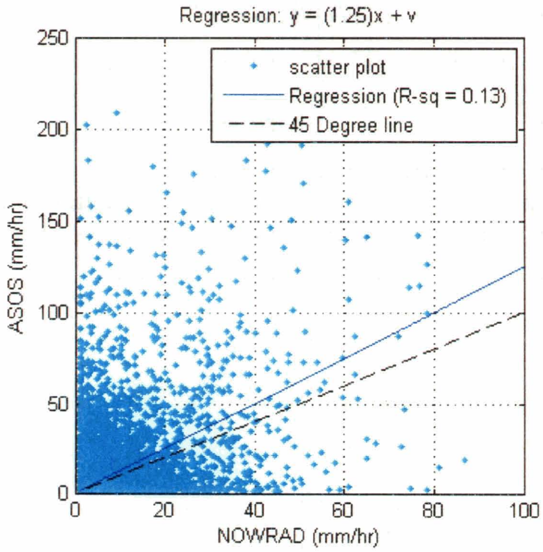


Figure 2-21: The scatter plot and regression analysis of ASOS versus NOWRAD from an hourly measurement during June 1st – August 31st, 2004 over the USGP region

With the constant h given, we only need to estimate the constant c_1 and c_2 used to specify the measurement error standard deviation in equation (2.2). We propose using the same analysis as for the satellite-based measurements. However, we do not have enough data to perform the analysis at each measurement time. Therefore, we must use the aggregate 3 months of data of ASOS in the regression analysis. First, we find the residual between ASOS and NOWRAD and bin the residual the nearest rounded ASOS integer. Finally, we perform the regression analysis between the standard deviation of the residual in each bin and the bin center, which is the ASOS measurement. Since we merge all 3 months data together, we will only have one standard deviation at each bin center. The regression analysis result and the estimated c_1 and c_2 for the ASOS rainfall data are given in Figure 2-22. The constant c_1 for ASOS is relatively smaller than the satellite based measurements. However, the constant c_2 is significantly higher. The high value of c_2 may be due to the position error and the assumption of $h = 1.0$. The values of c_1 and c_2 for ASOS measurements imply that ASOS measurements are not reliable for a low-intensity rainrate, but are more relatively more accurate at a high intensity rainrate.

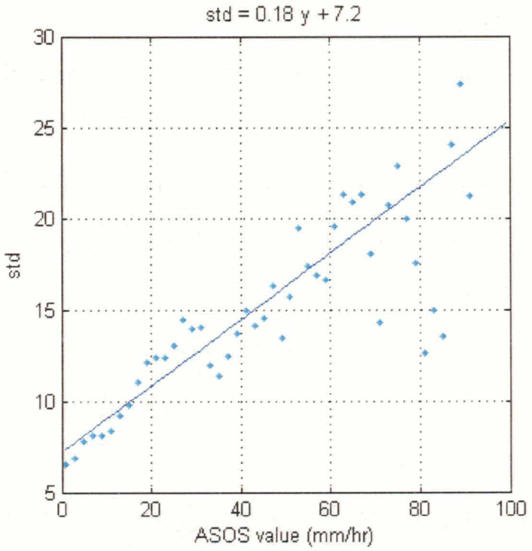


Figure 2-22: The regression analysis for estimating the constants c_1 and c_2 of ASOS measurement

2.4 Conclusions

In this chapter, we introduced the atmospheric forcing and rainfall measurements in the United States Great Plains region during the months of June – August 2004. The datasets consist of GOES cloud-top temperature, rainfall rate from NOWRAD ground-based radar stations, cumulative rainrate from ASOS rain-gauge stations, and satellite-based instantaneous rainrate measurement from TRMM, SSM/I, and AMSU-B. We would use these datasets to illustrate rainfall data assimilation over a large region later in this thesis. Details of each data source and its role for the USGP rainfall assimilation case study are summarized in Table 2.1.

We also presented techniques to estimate the error statistics of these rainfall data by validating it with NOWRAD rainfall data. For satellite-base rainfall measurement, we assumed that there are two independent types of measurement errors: the position error and the intensity error. To estimate the position error, we used the field alignment algorithm to align the satellite-based rainfall measurements with NOWRAD data and kept the average displacement to represent position error statistics in the x- and y-direction. For the intensity error, we related the measurements after correcting for position error with the true rainfall using (2.1). We showed that we can confidently assume the constant h to have an expected value of 1.0. Finally, we assumed that the measurement intensity error variance is related to the measurement value by (2.2) and estimated the constants c_1 and c_2 using regression analysis.

As for the ASOS gauge data, we cannot efficiently estimate the position error because the measurement point is too scattered in space. Therefore, we assumed that it only has intensity error. Similarly, we related the measurements to the true state using (2.1), with the constant $h = 1.0$. Finally, we perform the regression analysis to obtain the constants c_1 and c_2, which are needed to calculate the intensity measurement error variance in equation (2.2). Error statistics of rainfall measurement for the USGP case study are summarized in Table 2.2.

Table 2.2: Position and intensity error statistics of the USGP rainfall measurements

Measurement	**Position Error Statistics (degree)**				**Intensity Error Statistics**	
	x-displacement		**y-displacement**			
	mean	**std**	**mean**	**std**	c_1	c_2
1. ASOS	-	-	-	-	0.18	7.2
2. TRMM	-0.03	0.05	-0.03	0.04	0.22	1.6
3. SSM/I	-0.07	0.20	-0.03	0.21	0.32	0.7
4. AMSU	-0.02	0.11	-0.06	0.10	0.25	0.6

Conclusively, we obtained all measurements for the USGP case study in the consistent format. We estimated the error statistics of rainfall measurements. Thus, we can now focus on rainfall model and data assimilation technique in the following Chapters.

Chapter 3

Dynamic Rainfall Model

3.1 Introduction

In this chapter, we propose the recursive cluster-point rainfall (RCR) model for propagating rainfall information through space and time. The RCR model is computationally efficient while capable of simulating reliable spatial and temporal structures of rainfall. The model is modified from the spatiotemporal stochastic rainfall model using the cluster-point process [25, 93, 94, 112]. It utilizes cloud-top temperature to improve model accuracy and to deal with rainfall intermittency. In addition, the rainfall model combines the multi-resolution alignment (MRA) algorithm to estimate from cloud-top temperature data the velocity field, which is used to propagate rainfall. The recursive form of the model fits well with the sequential data assimilation framework, and its low computation cost is ideal for the ensemble approaches.

The organization of this chapter is as follows. In Section 3.3 we will focus on the stochastic model, present the spatiotemporal cluster-point process model, and introduce the RCR model. The early version of the RCR model is purely stochastic and cannot deal with rainfall intermittency. In Section 3.4, we introduce the use of GOES cloud-top temperature to cope with the intermittency problem, and use the MRA algorithm to estimate the velocity field. Then we revise the RCR model by adding GOES as the input forcing data. We implement the RCR model for the United States Great Plains (USGP) project in Section 3.5. Finally, Section 3.6 provides the summary of the chapter.

3.2 Spatiotemporal Stochastic Rainfall Model

We are interested in using a spatiotemporal stochastic rainfall model to propagate rainfall through space and time and provide comprehensive reanalysis of rainfall information over a large area. Even though it may not be as accurate as meteorological models, the stochastic model is much simpler and demands significantly less computation resources. In addition, by conditioning on past measurements using data assimilation, a stochastic model should be sufficient for providing reliable characteristics of rainfall space and time at a particular resolution. With limitations of current technology, we believe that the best way to obtain good short-term rainfall estimation and reanalysis data is to combine relatively simple but physically credible models with carefully designed observational strategies. Thus, we focus our interest on modifying and developing an efficient but accurate spatiotemporal stochastic rainfall. We will then combine the model with a data assimilation framework to provide a complete description of rainfall in space and time conditioned on the available measurements.

There have been a fair number of spatiotemporal stochastic rainfall models proposed in the last couple of decades. Most are based on hierarchical clustering of rainfall structure and make use of the cluster-point process to model rainfall in space and time [26, 93, 112-114]. Among these models, the cluster-point process rainfall model proposed by Rodriguez and Eagleson [112] (the RE model) is directly applicable to the rainfall data assimilation problem. This model is fast and simple to implement, yet capable of providing reasonable spatial and temporal rainfall structures. The following section will provide details in generating rainfall using the RE model and develop an efficient recursive form for it.

3.2.1 The Rodriguez and Eagleson Cluster-Point Process Model

The cluster-point process rainfall model by Rodriguez and Eagleson [112] is a spatiotemporal stochastic rainfall model. It describes a rainstorm event in space and time with the point-process method [94] and the hierarchical clustering structure [101, 114, 131]. This model is used primarily to provide a descriptive characteristic of rainfall intensity and cumulative rainfall processes. The RE model is a relatively simple stochastic description of the rainfall process. It uses a small number of parameters that can be estimated from historical rainfall data. Secondly, geometry and kinematics of the model is suited to the structure and organization of tropical cloud cluster as described by [60]. Moreover, the covariance function derived from the model satisfies approximately a Taylor frozen turbulence hypothesis for turbulent flows [53], which is well suited to the analysis scheme using first and second moment statistics such as the Kalman filtering algorithm [83]. Finally, the model can easily be adjusted and modified to fit many data assimilation and estimation frameworks.

The hierarchical structure of the rainfall field in the cluster model is based on the cluster-point process, first introduced by [72]. The idea is empirically supported by radar and gauge measurements [7, 59]. According to the cluster model assumption, rainfall occurs over a large region called a large mesoscale area (LMSA). Inside a LMSE, there are clustering regions of more intense rainfall called small mesoscale areas (SMSA).

To obtain the rainfall field from the R-E model, the first step is to obtain the locations of the *cluster* centers. The model assumes that the rain cells are born according to the Neyman-Scott process, where cluster centers are randomly distributed in 2-dimensional space according to the Poisson process with parameter λ_c, e.g. cluster per length^2. These cluster centers do not have rainfall intensity directly associated with them, but they contain a random number or rain cells. Figure 3-1 shows the concept of cluster-point model with cell centers placed around each cluster center. These cluster centers generated from the Poisson process in space will last for an entire storm event.

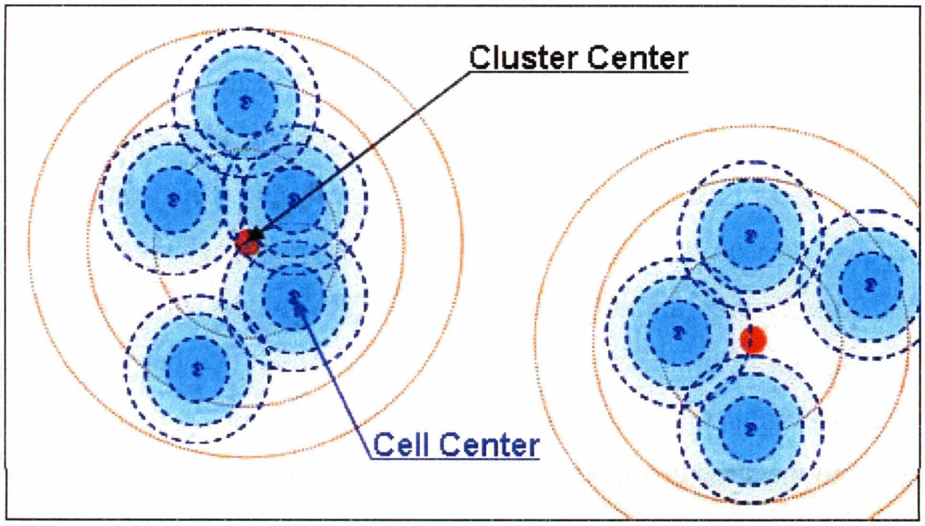


Figure 3-1: A spatial diagram of the RE cluster point rainfall process showing cluster centers and rain cell centers

After the cluster centers are located, we need to obtain the locations of *rain cell* centers. The probability of occurrence of a rain cell centered at *(x,y)* and time *t* after the storm origin around each cluster centered at *(x_c,y_c)* is given by

$$f_c(x,y,t) = f_1(t)\, f_2(x,y) \tag{3.1}$$

where

$$f_1(t) = \beta e^{-\beta \cdot t}, \; t>0 \tag{3.2}$$

$$f_2(x,y) = \frac{1}{2\pi\sigma_c^2}\exp\left\{-\frac{(x-x_c)^2+(y-y_c)^2}{2\sigma_c^2}\right\} \tag{3.3}$$

Equation (3.2) assumes that rain cells are born stochastically in time according to an exponential distribution with parameter β (e.g., the chance of generating a new rain cell decreases as the time from the storm origin increases.) Equation (3.3) assumes a symmetric Gaussian distribution of rain cells around the cluster center with the spatial decay constant σ_c (e.g., the chance of generating a new rain cell decreases with the distance from the cluster center.)

The number of rain cells in each cluster is an independent and identically distributed random variable with mean *v* and are independent of the Poisson process, which governs the recurrence of cluster centers. The probability of a rain cell to occur at the Euclidian grid point (x,y) after time *t* from the storm origin is given by

$$p(x,y,t) = v \cdot \sum_c f_c(x,y,t) \tag{3.4}$$

where the summation is over all cluster centers in the domain. Note that the expected number of rain cells in the storm (e.g., λ cells per length^2) is given by

$$\lambda = \lambda_c \cdot v \tag{3.5}$$

Once rain cells are born, they last throughout a whole storm event. However, the rainfall intensity generated from the rain cells will dissipate in space and time.

After we obtain the locations of all rain cells in the storm events, we can obtain the rainfall intensity field from the following procedure. First, we draw a random birth time and an intensity at the cell center at birth for each rain cell from exponential distributions with mean $1/\beta$ hour and *E[i₀]* mm/hr, respectively. Second, we assume that rainfall intensity exponentially decreases with the age of the cell and the distance from the cell center. Thus, a rain cell *j* centered at (x_j,y_j) born at time t_j has a rainfall intensity at location *(x,y)* and time *t* (e.g., $r_j(x,y,t)$) via

$$r_j(x,y,t) = i_0^j \cdot g_1^j(t,t_j) \cdot g_2^j(x,y,x_j,y_j) \tag{3.6}$$

where i_0^j is a random rainfall intensity of rain cell *j*, and g_1^j and g_2^j are given by

$$g_1^j(t,t_j)=\begin{cases}\exp\left\{-\alpha(t-t_j)\right\} & ,t\geq t_j\\ 0 & ,t<t_j\end{cases} \tag{3.7}$$

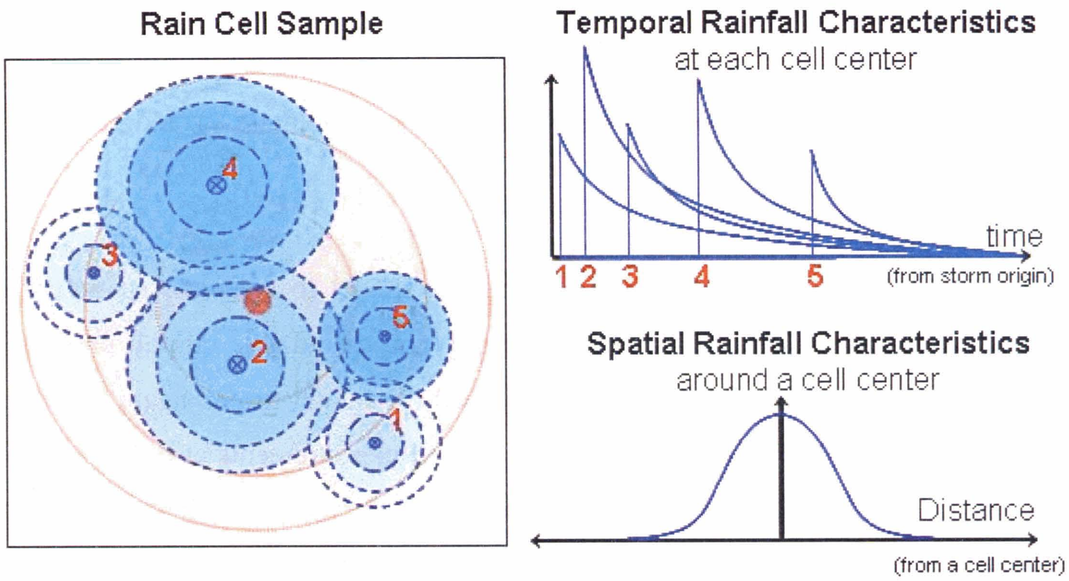
$$g_2^j(x,y,x_j,y_j)=\exp\left\{\frac{(x-x_j)^2+(y-y_j)^2}{2\sigma^2}\right\} \tag{3.8}$$

Equation (3.7) assumes that rainfall is dissipating in time exponentially with parameter α (e.g., temporal decay constant.) In addition, there is no rainfall at the time before the cell was born (e.g., $t < t_j$). Equation (3.8) assumes that rainfall is also dissipating in space. The dissipation depends on the distance from the cell center and the parameter σ (e.g., cell spatial decay constant).

Finally, rainfall intensity at the Euclidian grid point (x,y) at time *t* is the summation of the contribution from all rain cells in the domain, e.g.

$$r(x,y,t)=\sum_j r_j(x,y,t) \tag{3.9}$$

where the summation is over all rain cells which are born before the current time *t*. The model can incorporate cell movement with common velocity *(u,v)* in the x-and y-direction if desired. Figure 3-2 illustrates the spatial and temporal characteristic of the rainfall field generated from the RE cluster rainfall model. The temporal characteristic diagram shows that rainfall at each cell center exponentially decays from its birth intensity. The spatial characteristic diagram shows that at each time instance, rainfall intensity exponentially decays with distance from the rain cell center.


Rain Cell Sample

Temporal Rainfall Characteristics at each cell center

Spatial Rainfall Characteristics around a cell center

Figure 3-2: The spatial and temporal characteristics of rainfall field generated from the RE cluster point rainfall model

The procedure for obtaining a rainfall field on a two-dimension Euclidian grid is summarized as follows:

1) Locate the rain cluster centers inside the domain according to the Poisson process
   - Create independent uniform random numbers (i.e., *RAND₁*) between zero and one at each pixel on the Euclidian grid domain.
   - Select any pixel with the random number $RAND_1 < \lambda_c / n_x \cdot n_y$ to become a cluster center where n_x and n_y is the grid dimension in x- and y- direction.
2) Locate the rain cell centers inside the domain according to the Neyman-Scott process
   - Create independent uniform random numbers (i.e., $RAND_2$) between zero and one at each pixel on the Euclidian grid domain.
   - Select any pixel with the random number $RAND_2 <$ p(x,y,t) given in equation (3.4). At this point the cluster centers can be neglected.
3) Assign a random birth time from the exponential distribution with mean value $1/\beta$ to each rain cell center.

4) Assign a random initial rainfall intensity from the exponential distribution with mean value $E[i_0]$ to each rain cell center.
5) Calculate the rainfall field at the Euclidian grid point (x,y) and time t from equation (3.9).

Rodriguez and Eagleson [112] applied this rainfall model to calculate spatiotemporal mean and covariance functions of rainfall intensity and cumulative rainfall processes at any given time and location of interest. [83] used the model to propagate rainfall and applied the Kalman filter algorithm using all rainfall measurements at once. However, their method is not practical for real-time problems, especially over a large area because the rainfall model depends on the origin time of the storm. First defining the original time of the storm event is subjective and vague especially over a large region where rainfstorms usually advect and overlap one another. Second, the absolute time reference scheme is costly and time consuming to re-evaluate at every time step when a new measurement becomes available.

From the original R-E model, we propose a recursive form using the Markov property. The recursive form allows us to disregard information in the past in order to save storage and computation time. Moreover, we can employ a sequential data assimilation scheme to efficiently update new measurement in real-time.

3.2.2 The Recursive Cluster-Point Rainfall (RCR) Model

The Recursive Cluster-Point Rainfall (RCR) model is a modified version of the original RE cluster-point rainfall model introduced in the previous section. It assumes the Markov property and defines a rainfall process over a time interval instead of over a whole storm event. Let *r(x,y,t)* denote a rainfall intensity at the Euclidian grid location *(x,y)* at time *t*. The recursive rainfall model can be written as

$$r(x,y,t+dt) = F\left\{r(x',y',t)\right\} + w(x,y,t) \tag{3.10}$$

where *dt* is the time interval, $F\{\cdot\}$ is the dissipation-advection function and *w(x,y,t)* is the process noise. The dissipation-advection function accounts for a temporal rainfall dissipation and a two-dimensional rainfall advection from the beginning of the time interval to the end. The process noise is a non-negative random but spatially correlated field. It represents new rainfall randomly generated during the time interval. Note that the cluster-point process is introduced only in the process noise *w(x,y,t)* when the additional new rainfall is generated. The rainfall clusters and cells in the RCR model have different meanings than those in the original model, in which they are used to represent all rainfall fields of a storm event. Details of the dissipation-advection term (e.g., $F\left\{r(x',y',t)\right\}$) and the process noise term (e.g., $w(x,y,t)$) are given in the following sections.

(a) The Dissipation-Advection Term: $F\left\{r(x',y',t)\right\}$

The dissipation-advection function, $F\{\cdot\}$, describes the dissipation and advection of an existing rainfall field. The function is separated into two components: (1) the temporal rainfall dissipation component and (2) the spatial two-dimension advection component.

With regard to the temporal dissipation, we assume that existing rainfall intensity field exponentially dissipates over time with a dissipation constant α per hour. Physically, this constant implies that rainfall will reduce to about one-half of its existing value within approximately $^{0.7}\!/_{\alpha}$ hours. This dissipation represents the temporal decay of rainfall of rainfall. It is the only mechanism in the model to reduce the amount of rainfall. This constant is equally applied to a whole domain. Thus, given a current rainrate (e.g., $r(x,y,t)$), the rainrate at the next *dt* hour is given by

$$r(x,y,t+dt) = r(x,y,t)\cdot e^{-\alpha\cdot dt} \tag{3.11}$$

In addition to temporal dissipation, rainfall may spatially advect in two-dimensional space. We assume that rainfall advection follows the Lagrangian persistent framework [44] over a short forecasting interval, *dt*. It states that the forecast variable at time *t+dt* over a position *(x,y)* comes from the existing variable at position (x_0, y_0) at time *t*.

$$r(x,y,t+dt) = r(x_0,y_0,t) \tag{3.12}$$

The position *(x,y)* is related to (x_0, y_0) by the velocity field *(u,v)* in the following form:

$$x = x_0 + dt\cdot u \tag{3.13}$$

$$y = y_0 + dt\cdot v \tag{3.14}$$

This velocity field *(u,v)* is commonly used in many rainfall model to advect rainfall data through space and time. Acquiring a comprehensive velocity field *(u,v)* over a whole domain at a spatial and temporal resolution of interest is a challenging task. There are many studies that focus entirely on the estimation of this velocity field [106, 127, 128]. For now, we will assume that the velocity field is available at the same temporal and spatial resolution as the rainfall field.

The order of applying the temporal dissipation and spatial advection to rainfall field is not important. Moreover, there is no randomness associated with the dissipation-advection term in the form given in this section. It is possible to introduce some noise into the parameter α, as well as the velocity fields *(u,v)* to introduce uncertainty. In this case, we should be careful not to use noise that is too large, especially for generating the random velocity field. A velocity field that is too scattered can disaggregate rainfall features and causes the dissipation-advection term to be unrealistic.

(b) The Process Noise Term

The process noise *w(x,y,t)* in (3.10) is a non-negative random but spatially correlated field. It represents new rainfall generated during the time interval *dt* of interest. This new random rainfall field is constructed using the concept of cluster-point process from the original RE rainfall model. However, rain clusters and rain cells are defined over the time interval instead of over a whole storm event beginning at an ambiguous origin time. In other words, we assume that the birth times are uniformly distributed over the time interval instead of exponentially distributed over the storm event. This assumption is applicable to any interval *dt*. However, we recommend the time step size between 15 minutes to a few hours. We will discuss about selecting a suitable time step later. The model modification produces similar rainfall features, but has different statistics (e.g., the mean and covariance functions of rainfall intensity.)

Six parameters needed to be specified in order to generate a new rainfall field over the time interval *dt*. These parameters and their definitions are summarized in Table 3.1. Note that the last two cluster parameters are used to relate the cluster properties to the cell properties, e.g.

$$\beta_c = {}^{\beta}\!/_{\nu} \tag{3.15}$$

$$\sigma_c = \sigma \cdot \rho \tag{3.16}$$

where β_c and σ_c are the cluster birth probability and the cluster spatial dissipation constant, respectively. The cluster birth probability governs the number of clusters to be born during the time interval *dt*, according to the two-dimensional Poisson process. This parameter is defined as λ_c in the original model. The cluster spatial dissipation constant governs the distribution of rain cells within the domain during the time interval *dt* according to the following probability

$$p(x,y) = \nu \cdot \sum_k \frac{1}{2\pi\sigma_c^2} \exp\left\{ -\frac{(x-x_c^k)^2 + (y-y_c^k)^2}{2\sigma_c^2} \right\} \tag{3.17}$$

where the summation is over all clusters, and $\left(x_c^k, y_c^k\right)$ is the center of cluster *k*. Equation (3.15) is adapted from equations (3.1)-(3.4), but it neglects the $f_1(t)$ term because the birth time is uniformly distributed over the time interval.

Table 3.1: Parameters of the Recursive Cluster-Point Rainfall (RCR) model

Parameter Name	Unit	Description
1. Cell Birth Probability, β	$km^{-2} \cdot hr^{-1}$	Probability of a rain cell to be born inside a unit area over a time interval
2. Temporal Decay Constant, α	hr^{-1}	An exponential decay constant for temporal dissipation of rainfall field
3. Cell Spatial Decay Constant, σ	km	A Gaussian decay constant for spatial dissipation of rainfall from each rain cell center
4. Initial Mean Rainrate at Cell Center, *E[i₀]*	mm / hr	An expected rainrate at rain cell centers when they are first born
5. Mean Cell Density, ν	cells / cluster	An expected number of rain cells per rain cluster
6. Spatial Dissipation Ratio, ρ	-	A ratio between the cluster and the cell spatial dissipation constant

Once we locate all rain cells born, we assign a random age and a random initial intensity at birth to each cell. The random age is drawn from an independent and identical uniform distribution between zero and *dt*. The random initial intensity at a center is drawn from an independent and identical exponential distribution with the mean $E[i_0]$. The new rainfall intensity field presented at the end of time interval *dt* (e.g., denoted by $w(x,y,t)$) is given by

$$w(x,y,t)=\sum_j i_0^j \cdot \exp\{-\alpha \cdot a_j\} \cdot \exp\left\{-\frac{(x-x_j)^2+(y-y_j)^2}{2\sigma^2}\right\} \tag{3.18}$$

where $\left(x_j, y_j\right)$, a_j, and i_0^j represent the center, the random age, and the random initial intensity of the j-th cell, respectively. The summation in equation (3.16) is over all rain cells in the domain. Finally, the rainfall field at the next time step *t+dt* is given by the summation of the dissipation-advection term and this new rainfall according to the recursive equation (3.10).

The procedure for obtaining the rainfall field at time *t+dt* using the RCR model are summarized as follows:

1) Decay and advect existing rainfall field using (3.11) – (3.14) to obtain the dissipation-advection term
2) Obtain new rainfall field born during time interval *dt* by the following:

2.1) Locate the cluster centers inside the domain according to the Poisson process

- Create independent uniform random numbers (i.e., *RAND$_1$*) between zero and one at each pixel on the Euclidian grid domain.
- Select any pixel with the random number *RAND$_1$* < $\beta_c \cdot dt = {}^{\beta}\!/_{V} \cdot dt$ to become a cluster center.

2.2) Locate rain cell centers inside the domain

- Create independent uniform random numbers (i.e., *RAND$_2$*) between zero and one at each pixel on the Euclidian grid domain.

- Select any pixel with the random number $RAND_2 < p(x,y) \cdot dt$ where $p(x,y)$ is given in (3.17). At this point, we can neglect the cluster centers.

3) Assign a random age by drawing from the uniform distribution between 0 and dt to each rain cell center.
4) Assign to each rain cell center a random initial rainfall intensity from an exponential distribution with mean value $E[i_0]$.
5) Calculate new rainfall field generated during time interval dt (e.g., $w(x,y,t)$) from (3.18).
6) Calculate total rainfall field at time $t+dt$ from (3.10).

There are some remarks on the recursive rainfall model that need emphasis. First, the spatiotemporal mean and covariance of rainfall intensity and cumulative rainfall derived in [112] cannot be applied to the RCR model. This is because we change the temporal structures of rainfall and include the advection by the velocity field (u,v). These probabilistic characterizations of rainfall are essential in data assimilation and many other applications. However, we can use the Monte Carlo method to estimate these spatiotemporal statistics numerically. It is fast and simple to generate many replicates with the RCR model. By using ensemble approaches, we have more flexibility in the model that does not depend on fixed analytical statistics.

Secondly, all terms in the recursive rainfall model in (3.10) are non-negative values since they are all representing rainrate intensity. The process noise term (e.g., $w(x,y,t)$) will only add more rain to the model. The only mechanism in the RCR model that decreases rainfall intensity is with the exponential decay with the parameter α. Rainfall generated from the RCR will abruptly increase because of new rain cells but always slowly and continuously decreases by the temporal dissipation. Theoretically, rainrate will never reach zero with this approach, but we can set a minimum detection threshold and force rainfall to zero if one desires.

Thirdly, the new rainfall field $w(x, y, t)$ is generated by assuming that the locations of all cell centers are defined at the end of time interval *t+dt*. This assumption simplifies the algorithm because there is no need to advect each cell center using the velocity field *(u,v)*. It is possible to define them at the beginning of the time step and advect them using the Lagrangian persistent framework as well.

Finally, we assume that rainfall spreading from a cell center and the distribution of rain cells around a cluster center are isotropic (i.e., have circular shapes). We could move from circular to elliptical shapes by specifying the cell and cluster spatial dissipation constant in x- and y-direction separately. We can also rotate the ellipse shapes to any desired angle (e.g., in order to better fit a frontal storm system). However, such a modification adds more parameters and increases computation cost and complexity of the model.

3.3 The RCR Model with GOES Forcing

The RCR model provides a simple yet efficient way to propagate rainfall fields through time and provide spatial and temporal characterization of rainfall. However, the model has two important drawbacks. First, the RCR model cannot efficiently deal with rainfall intermittency (i.e. zero rainrate), especially over a large-scale problem. Instead, it generates scattered clusters of rainfall everywhere in the domain according to the Poisson process. Therefore, rainfall may occur at an inappropriate location. Moreover, once rainfall is generated, it is difficult to remove using just a temporal decay function. Secondly, the RCR model requires a velocity field (e.g., *(u,v)*) to advect rainfall field. This velocity field can be difficult to acquire. To cope with these problems, we propose using the GOES cloud-top temperature as a forcing input for the RCR model. This atmospheric forcing will help adjust the location and amount of new cells. In addition, we can employ the multi-resolution alignment algorithm to estimate the velocity field from two consecutive GOES images and use it to advect the rainfall field.

3.3.1 Rainfall Intermittency and GOES Usage

When the RCR model is utilized on a large-scale rainfall problem, it generates clusters of rainfall at random locations. Although the characteristics of each rain cluster can be realistic, their locations are too sparse and are not consistent with real rainfall event. For example, Figure 3-3(a) shows rainfall measurements from NOWRAD observations on June 2nd, 2004 at 0:00GMT. If we use the RCR model to generate rainfall, we will obtain a rainfall event that looks similar to Figure 3-3(b). Each individual cluster of rainfall from the RCR model is relatively realistic, but there are so many clusters scattering everywhere in the domain. Since the RCR model generates new rainfall stochastically, it has no information about where rain clusters should be or should not be placed.

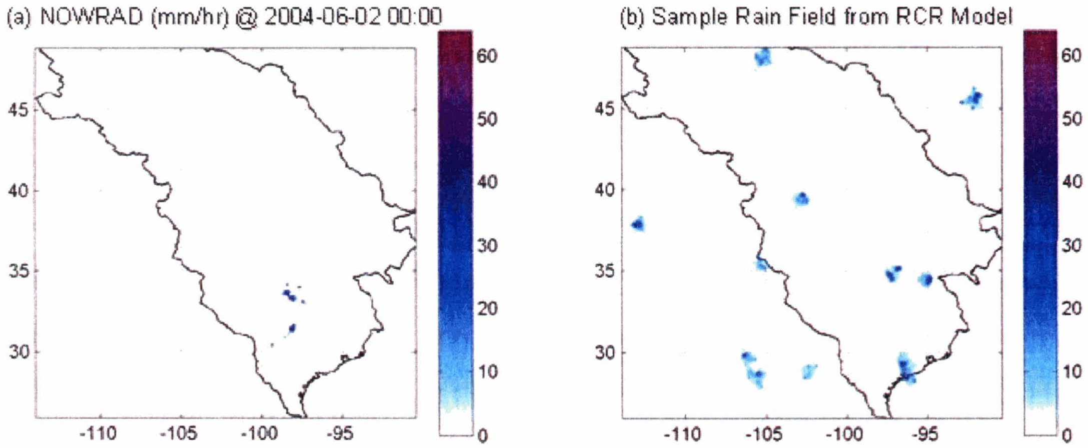

(a) NOWRAD (mm/hr) @ 2004-06-02 00:00
45
40
35
30
-110
-105
-100
-95
60
50
40
30
20
10
0
(b) Sample Rain Field from RCR Model
45
40
35
30
-110
-105
-100
-95
60
50
40
30
20
10
0

Figure 3-3: Intermittency problem in the RCR model when apply to a large-scale problem - (a) NOWRAD rainfall intensity in the USGP study region on June 2nd, 2004 at 0:00GMT and (b) sample rainfall field generated from the RCR model

To solve the rainfall intermittency problem, we propose incorporating real-time atmospheric forcing variables into the RCR model. There are many variables and many methods to incorporate those variables into the model. However, since we are seeking a simple and efficient rainfall model, we prefer to minimize the number of forcing variables and keep the concepts as straightforward as possible. In addition, the forcing variables should be relatively easy to acquire at the desired spatial and temporal resolution.

Among the many weather-related variables, cloud-top temperature from the infrared channel seems to be the most suitable one. It is relatively easy to acquire and available at high spatial and temporal resolution. In the USGP project, the cloud-top temperature is obtained from GOES infrared data described in Chapter 2. Cloud-top temperature can help the model give more accurate forecasts in many ways. Various studies attempt to use it directly to estimate rainfall intensity [1, 46, 80, 130], but none of these methods provides acceptable results. Others attempt to use cloud-top temperature in combination with other data sources to approximate the location of rainfall regions [61, 62, 90, 111, 125, 135]. The later method of employing cloud-top temperature usually provides better results. These studies agree that most convective thunderstorms are characterized by very low cloud-top temperatures. In addition, areas with little or no clouds (e.g., higher cloud-top temperature) usually contain zero rainrate. However, not all deep cloud regions have rainfall. For illustration purposes, the comparison between GOES cloud-top temperature and NOWRAD rainfall rate is given in Figure 3-4. It is apparent that low temperature clouds usually cover large areas where only a small portion coincides with the rainy region.

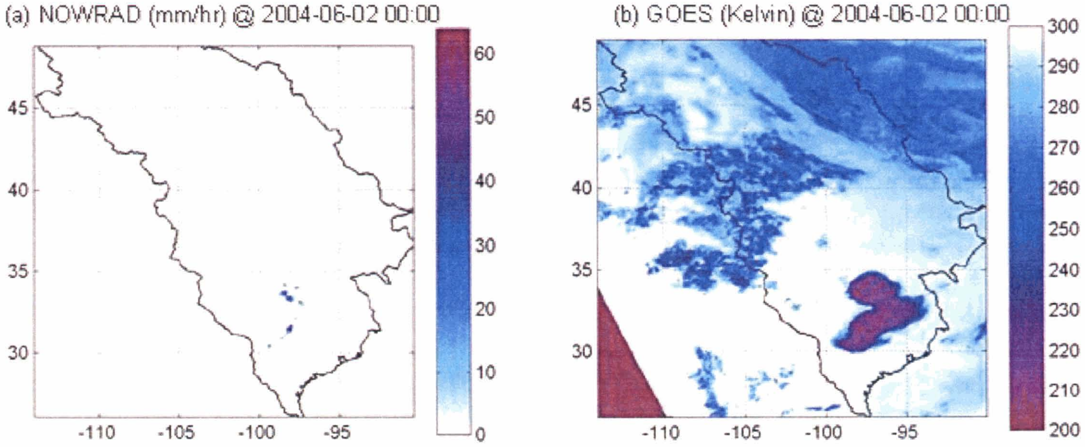


Figure 3-4: A comparison between (a) NOWRAD rainfall rate in mm/hr, and (b) the GOES cloud-top temperature in degree Kelvin over the U.S. great plain on 2004/06/01 at 04:00 GMT

In the USGP project, we propose a simple approach to condition the location and amount of clusters and rain cells in RCR model using a GOES cloud-top temperature threshold. We define two cloud-top temperature thresholds: the genesis threshold (T_G) and the rainy threshold (T_R). The genesis or birth threshold is the maximum cloud-top temperature for which new clusters and new rain cells can be born. In other words, a new rainfall field will only be generated inside the deep cloud region with cloud top temperature lower than T_G. The rainy threshold is the maximum cloud-top temperature for which rainfall is allowed. In other words, any rainfall in a pixel with cloud-top temperature greater than T_R will be suppressed. An example in which GOES cloud-top temperature was used to condition the rainfall field generated by the RCR model is given in Figure 3-5. The blue dotted line represents the boundary of GOES temperatures lower than $T_R = 290^o$ Kelvin and red dotted line represents the boundary of GOES temperatures lower than $T_G = 220^o$ Kelvin. It is clear that the RCR model with GOES input produces much more realistic rainfall output and solves the intermittency issue in both space and time.

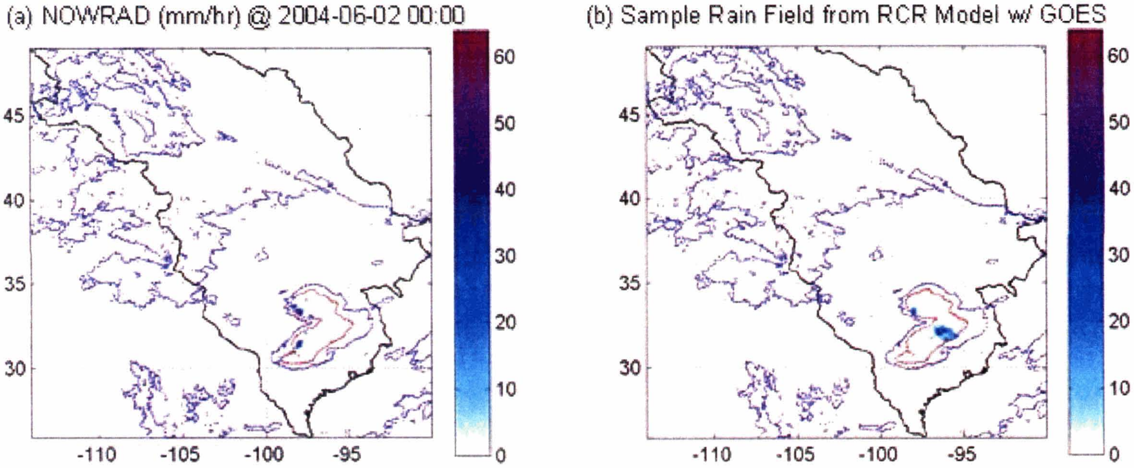

(a) NOWRAD (mm/hr) @ 2004-06-02 00:00
45
40
35
30
-110
-105
-100
-95
60
50
40
30
20
10
0
(b) Sample Rain Field from RCR Model w/ GOES
45
40
35
30
-110
-105
-100
-95
60
50
40
30
20
10
0

Figure 3-5: Improvement when using RCR model with GOES forcing – (a) NOWRAD rainrate in mm/hr, (b) sample rainfall field from RCR model with GOES forcing; blue and red boundary representing region with GOES temperature lower than T_R and T_G, respectively

The genesis threshold (T_G) and the rainy threshold (T_R) can be estimated from the statistics relating NOWRAD rain data and GOES temperature. The scatter plots of NOWRAD data and GOES cloud-top temperature during deep convective storms usually show a strong cloud-top temperature barrier. It separates high intensity rainrate from low intensity rainrate regions. We will use this barrier temperature to represent the genesis threshold (T_G). In addition, the scatter plot also shows a maximum cloud-top temperature for rainy pixels, which we will use as the rainy threshold (T_R). For example, the scattered plots of NOWRAD rainrate and GOES cloud-top temperature on June 1st, 2004 from 00:00-12:00 GMT are shown in Figure 3-6. Each row represents a time step at 00:00, 04:00, 08:00 and 12:00 GMT, while the left column shows NOWRAD rainrate (mm/hr), the middle column shows GOES cloud-top temperature (° K), and the right column shows the scatter plots from pixels inside the USGP regions only. Although the characteristics of the scatter plots vary with time and age of the storm, the plots exhibit a strong barrier for high intensity rainrate at around 220 degrees Kelvin. The maximum cloud-top temperature for which rainfall exists is at around 290 degrees Kelvin. We selected approximately 200 pairs of GOES and NOWRAD images from our case study from June 1st to August 31st, 2004, and the scatter plots for these measurements possess similar characteristics. Thus, we propose that for the USGP case study, the genesis threshold (T_G) is 220 degrees Kelvin and the rain threshold (T_R) is 290 degrees Kelvin.

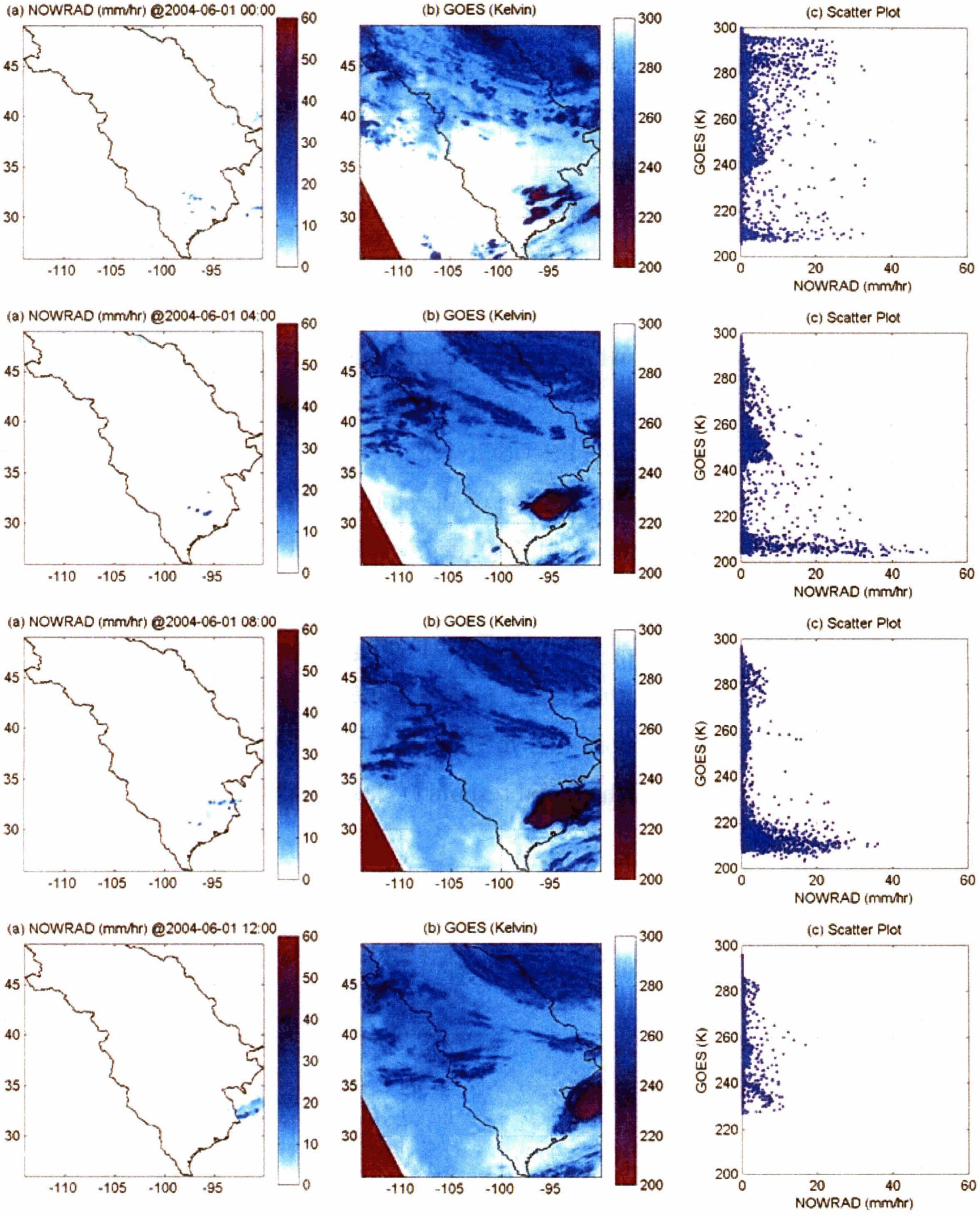


Figure 3-6: A scatter plot of NOWRAD rainrate versus GOES cloud-top temperature on June 1st, 2004 from 00:00-04:00 GMT over the USGP case study

3.3.2 Velocity Field from Consecutive GOES Images

The velocity field in the x- and y-direction denoted by *(u,v)* is a major component in the RCR model. It is required at every time step in order to propagate rainfall spatially. The velocity field can be obtained from many sources and at many elevations (e.g., direct wind measurement at the weather station, displacement of cloud or other substances in the atmosphere). However, it is normally very difficult to obtain a comprehensive velocity field in space and time at a specific resolution of interest. Since we would like to have the most effective and simple rainfall model possible, acquiring new forcing data seems to contradict our objective. Therefore, we propose to obtain the velocity field from the movement implied by two consecutive GOES cloud-top images. Because the RCR model uses GOES cloud-top temperature to help locate the rain cells, also using GOES to determine the velocity field minimizes the number of input forcing needed, thereby making the model easier to use.

To obtain a velocity field, we employ the multi-resolution alignment (MRA) algorithm to estimate the displacement that produces the minimal misfit between two consecutive GOES cloud-top images. The original alignment algorithm is proposed by [107] to deal with position error adjustment in the data assimilation framework. It iteratively searches for a displacement field that aligns one image to the other and minimizes the local constraint with regard to the misfit between those two images. Ravela and Chatdarong used the MRA algorithm to estimate the velocity field from GOES [106] and compare results with CIMSS derived wind [127, 128], which used a correlation-based algorithm. Details of the MRA algorithm are provided in Appendix A.

When using the velocity field derived from the movement of GOES cloud-top images, we assume that the movement of the cloud is equal to the movement of the rainfall feature itself. To test this assumption, we use the MRA algorithm to estimate the displacement field from one time step to the next using two consecutive NOWRAD rainrate images. This is then compared with the velocity field obtained from two consecutive GOES images. From our experience, the velocity fields derived from NOWRAD and GOES data look similar over a region with high rainfall intensity. Thus, using the velocity field derived from GOES cloud images to advect rainfall in the RCR model should give reasonable results. Figure 3-7 shows the comparison between two displacements obtained from two consecutive NOWRAD images and GOES images on June 1st, 2004 at 8:00 – 9:00 GMT over the USGP region. Images (a) and (b) show NOWRAD rainrate in mm/hr at 8:00 and 9:00 GMT, while images (d) and (e) show GOES cloud-top temperature in Kelvin at the same time periods. Images (c) and (f) show the displacement magnitudes and directions obtained from the MRA algorithm using NOWRAD and GOES, respectively. The displacement field presents the distance in degree needed to move the image at 8:00 GMT in order to align well with the image at 9:00 GMT. It is obvious that over the region with deep convective rainfall, the velocity field from GOES and NOWRAD are similar.

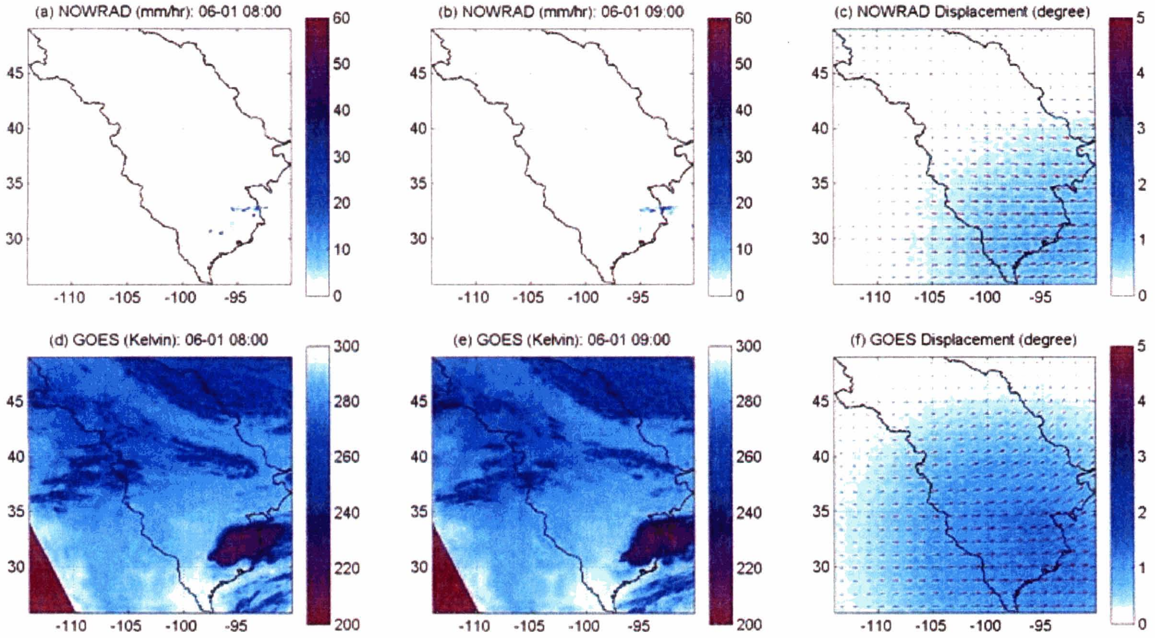

(a) NOWRAD (mm/hr): 06-01 08:00
(b) NOWRAD (mm/hr): 06-01 09:00
(c) NOWRAD Displacement (degree)
(d) GOES (Kelvin): 06-01 08:00
(e) GOES (Kelvin): 06-01 09:00
(f) GOES Displacement (degree)
60 50 40 30 20 10 0
5 4 3 2 1 0
300 280 260 240 220 200
45 40 35 30
-110 -105 -100 -95

Figure 3-7: Comparison between NOWRAD and GOES displacements obtained from MRA algorithm on June 1st, 2004 from 8:00 – 9:00 GMT over the USGP region

We repeated this same experiment on approximately 200 consecutive images of GOES and NOWRAD measurements at times during June 1st – August 31st, 2004 when there are significant storms. The magnitude differences between GOES and NOWRAD velocity fields are less than 0.5 degrees (distance degree) with 95% confidence. The angle differences between GOES and NOWRAD velocity are less than 10 degrees (angle degree) with 95% confidence. These differences are insignificant, and thus it is reasonable to use GOES velocity field to advect rainfall in the RCR model.

3.3.3 RCR Model with GOES Forcing

With the use of GOES cloud-top temperature, the RCR model is now practical for propagating two-dimensional rainfall through time without experiencing serious intermittency issues. In addition, the RCR model is fast, efficient and practical for approximate rainfall dynamics, even for a very large scale application. It only requires one atmospheric forcing source that is relatively easy to obtain at the spatial and temporal resolution of interest. The procedure to execute the RCR model and tips to make the algorithm faster and more efficient are as follows.

1) Define the RCR model parameters listed in Table 3.1 as well as the T_G and T_R thresholds mentioned in section 3.4.1. In general, these parameters are assumed to be constant throughout the simulation period.
2) Beginning at initial time t, define a matrix $\mathbf{X}_t \in \Re^{n_1 \times n_2}$ representing the initial two-dimensional rainfall field of dimension n_1 x n_2. $\mathbf{X}_t$ can be a zero matrix if we do not have prior knowledge about the initial rainfall field.
3) Obtain GOES cloud-top temperature at time t and $t+dt$ and estimate the velocity field using the Scaling Field Alignment algorithm presented in Appendix A.
4) Propagate $\mathbf{X}_t$ through to the next time step, e.g. $t+dt$, using the dissipation-advection function $F\{\cdot\}$ with velocity field obtained from step 3) and obtain $F\{\mathbf{X}_t\}$

5) Create $\mathbf{W} \in \Re^{n_1 \times n_2}$, the new rainfall matrix during the time interval *dt* by the following:

   5.1) Assign rain clusters on the study domain with the amounts and locations given by the two-dimension Poisson process over the time interval *dt*

      i) Generate a random matrix $RAND_1 \in \Re^{n_1 \times n_2}$ between zero and one from a uniform distribution, with the dimension equal to the two-dimensional grid.

      ii) Mark the location of cluster centers where $RAND_1 < \beta_c \cdot dt$

   5.2) Assign rain cells on the study domain based on the locations and amounts of rain clusters in the domain.

      i) From all locations of rain clusters, generate the cell birth probability field, denoted by $PROB_B$ from (3.15)

      ii) From GOES cloud-top temperature at time *t+dt*, obtain a screening matrix $MASK_B = 1$ for all pixel where GOES < T_G, and zero otherwise.

      iii) Generate a random matrix $RAND_2 \in \Re^{n_1 \times n_2}$ between zero and one from a uniform distribution.

      iv) Mark the location of the cell centers at the pixel where $RAND_2 < PROB_B \cdot MASK_B \cdot dt$

   5.3) Assign random cell ages, denoted by "*a*", to each rain cell by drawing a random number between *0* and *dt* from a uniform distribution.

   5.4) Assign initial rainfall intensity at the cell center at birth, denoted by "i_0", to each cell center by drawing from an exponential distribution with mean $E[i_0]$.

   5.5) Obtain the new rainfall field **W** at location *(x,y)* at the end of time interval *t+dt*, by summing up the rainfall from all rain cells using using equation (3.16).

6) Acquire rainfall at the next time step using (3.8), e.g. $\mathbf{X}_{t+dt} = F\{\mathbf{X}_t\} + \mathbf{W}$.

7) Suppress $\mathbf{X}_{t+dt}$ at the pixels where GOES at time $t+dt$ is greater than T_R.

8) Increment the time step and repeat starting from step 3

It is important to make sure the time step *dt* and the units of the parameters in Table 3.1 are consistent. In general, the time step is defined over one hour, e.g. $dt = 1$ hr. There is no restriction on the time step; however, making the time step too large may affect the accuracy of the velocity field obtaining from the SFA algorithm. Moreover, if the time step is too large, rainfall locations suggested by using the GOES cloud-top temperature at the end of the time step will be inaccurate. Therefore, it is recommended to define the time step *dt* to be approximately 1 hour or less.

Finally, there are a few guidelines for speeding up the algorithm, especially when working with a large domain. First, we recommend evaluating the probability function in equation (3.15) and (3.16) only up to a distance of 3-5 times the spatial dissipation constant σ_c or σ. At further distances, the function will be very close to zero and will become insignificant. Second, the summation in equation (3.15) and (3.16) may be time consuming if we have to run through each cluster center or cell center individually. We highly recommend creating a two-dimensional Gaussian surface matrix and performing two-dimensional calculations with a delta function centered at the cluster center or cell center. With this modification, the RCR model can efficiently generate and propagate rainfall through time.

3.4 Implementation of the RCR Model to the USGP Project

In this section, we will illustrate that the RCR model with only GOES forcing should be able to provide reasonable rainfall characteristics over the USGP. In addition, the model can be used to generate rainfall ensembles that have the anticipated correlation in space and time. By following the procedures of the RCR model with GOES input in the previous section, we should be able to efficiently propagate rainfall through time. The speed of the model is very fast even for a large problem. The simulation on MATLAB using ©Pentium-4 2.8 GHz processor with 2GB RAM for 500x500 domain is roughly about 0.5 second for each time step, regardless of the parameters used. When propagating the ensemble of rainfall through time, the computation time is linearly proportional to the ensemble size, e.g. it takes about 50 seconds to propagate 100 ensemble rain members through one time step on the same computer. The limitation in our simulation usually comes from insufficient memory for storing the rainfall ensemble, not computation time.

To demonstrate capability of the RCR model, we perform a simple experiment over the USGP region to propagate the rainfall field on June 1st, 2004 with a time step of 1 hour at a resolution of $L = 0.05$ degree. The parameters used in the simulation are as followed: $\beta=0.05\ L^{-2}\ hr^{-1}$, $\alpha = 0.6\ hr^{-1}$, $\sigma = 1.0\ L$, $E[i_0] = 5\ mm/hr$, $\nu = 50$ *cells/cluster*, and $\rho=2.5$. We use the GOES temperature threshold; $T_G = 220$ K, and $T_R = 290$ K. Figure 3-8 shows the simulation results at times 5:00 – 8:00 GMT. The left column represents the GOES cloud-top temperature (the only input given to the RCR model), and the middle column shows rainfall output from the RCR model. NOWRAD rainfall images are also given in the right column for comparison purposes. It is evident that the RCR model with GOES forcing is a simple, fast and efficient algorithm, while capable of providing rainfall fields that are realistic, especially for convective storm systems.

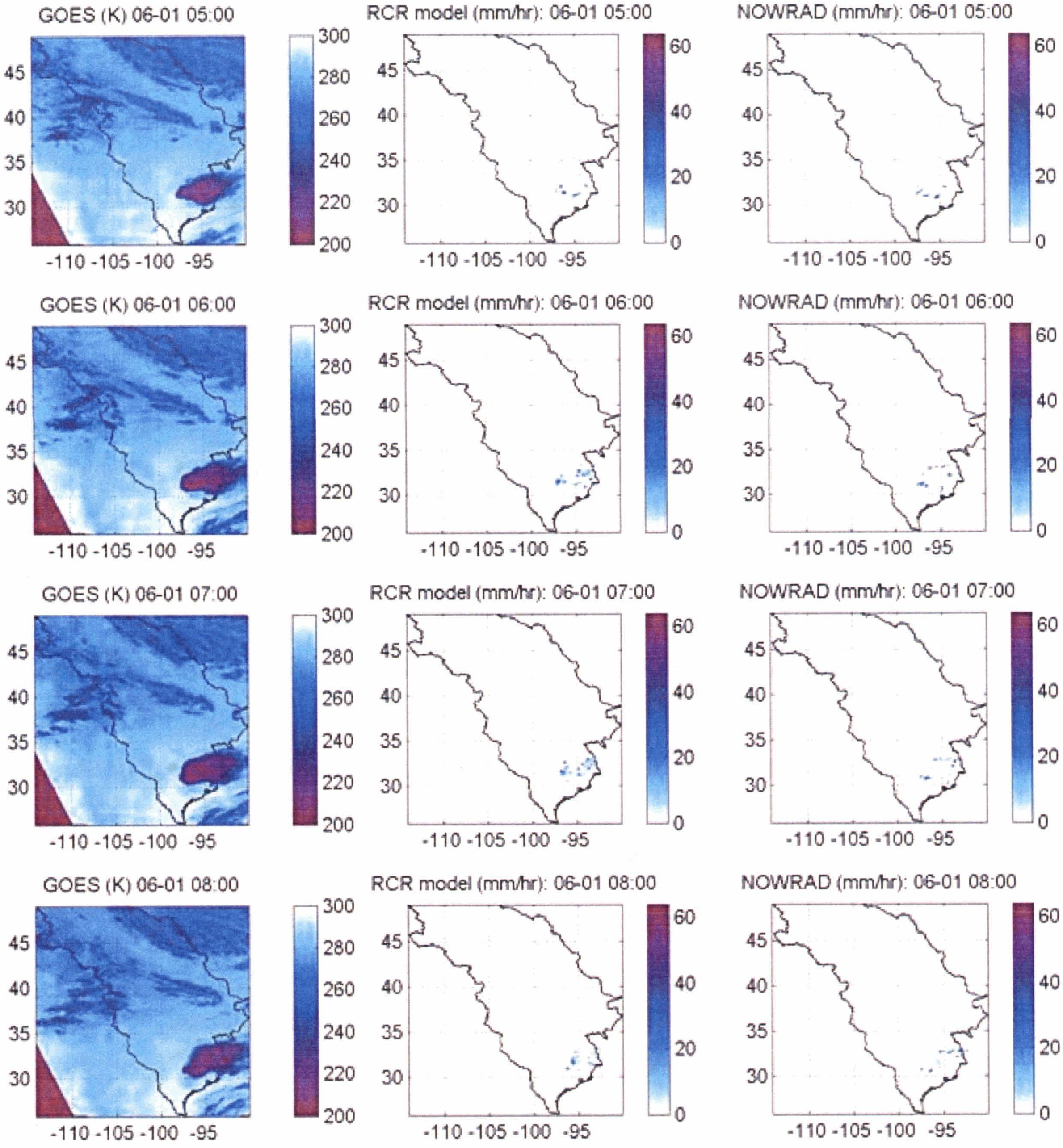

GOES (K) 06-01 05:00
RCR model (mm/hr): 06-01 05:00
NOWRAD (mm/hr): 06-01 05:00
GOES (K) 06-01 06:00
RCR model (mm/hr): 06-01 06:00
NOWRAD (mm/hr): 06-01 06:00
GOES (K) 06-01 07:00
RCR model (mm/hr): 06-01 07:00
NOWRAD (mm/hr): 06-01 07:00
GOES (K) 06-01 08:00
RCR model (mm/hr): 06-01 08:00
NOWRAD (mm/hr): 06-01 08:00
300
280
260
240
220
200
60
40
20
0
45
40
35
30
-110 -105 -100 -95

Figure 3-8: An implementation of the RCR model with GOES input over the USGP region from 05:00 – 08:00 GMT

3.5 Conclusions

In this chapter, we have presented the simple but efficient Recursive Cluster-Point Rainfall (RCR) model, which dynamically propagates rainfall in space and time. The RCR rainfall model is based on the spatiotemporal cluster-point process model first introduced in [112] as a descriptive representation of rainfall statistics in space and time. We imposed the Markov property on the original rainfall model to arrive at the RCR model. The recursive model eliminates the need to store all past rainfall history or evaluate rainfall from the storm origin. In addition, measurement information can be incorporated into the rainfall model in real-time using the sequential data assimilation framework.

To handle rainfall intermittency, we used cloud-top temperatures to precondition the probability of birth for rainfall cells. The cloud-top temperature is relatively easy to acquire. For example, in our USGP project, the cloud-top temperature is obtained from the GOES infrared dataset. The RCR model with GOES input can deal with rainfall intermittency reasonably well and produce much more realistic rainfall fields. In addition, the same cloud-top temperature data will be used to obtain the velocity field required by the RCR model at each time step. We use the Multi-Resolution Alignment (MRA) algorithm to estimate displacement field from two consecutive GOES images and use this velocity to advect rainfall in space. The experiments show that velocity fields derived from GOES cloud-top temperature is consistent with velocity fields obtained directly from two consecutive NOWRAD rainfall images, especially over deep convective rainfall regions. Therefore, it is appropriate to use the velocity fields from GOES data to propagate rainfall fields.

The RCR model with GOES input is capable of efficiently propagating rainfall features through space and time. In the next chapter, we will present detailed discussion of the data assimilation framework and techniques. By combining data assimilation with the RCR model, we can merge multiple sources of rainfall measurements to provide comprehensive reanalysis of rainfall data and ensemble rainfall fields, the ultimate goal of our thesis.

Chapter 4

Dynamic Rainfall Data Assimilation

4.1 Introduction

In this chapter, we introduce the methodology of ensemble sequential data assimilation and apply it with the Recursive Clustered Rainfall (RCR) model presented in Chapter 3 to estimate comprehensive rainfall ensembles. The organization of Chapter 4 is as follows. Section 4.2 will provide the background on sequential data assimilation and introduce the Ensemble Kalman Filter (EnKF) algorithm. We then propose a more stable and efficient algorithm for the EnKF as well as introduce the Ensemble Kalman Smoothing (EnKS) algorithm, which is an extension of the EnKF algorithm. In Section 4.3, we introduce the state-augmentation technique for estimating the parameters of the RCR model to be used with the United States Great Plains (USGP) case study. We also revisit the rainfall measurement sources and their error statistics, which were described in Chapter 2. These parameters and statistics are required to perform the ensemble data assimilation correctly. The implementation of the dynamic rainfall data assimilation over the USGP project is carried out in Section 4.4. The comprehensive rainfall ensemble is then validated with the NOWRAD rainfall data to assess the accuracy of the algorithm. Finally, we discuss the rainfall data assimilation and conclude the chapter in Section 4.5.

4.2 Sequential Data Assimilation for Non-linear Dynamic Systems

Sequential data assimilation is a technique that efficiently characterizes the variables of interest, known as the state variables, and produces the analysis state from all relevant information (e.g., forecasts and measurements) in a recursive fashion. In the data assimilation framework, information is divided into two categories, forecasts and measurements. Forecasts, which include the state variables and their statistics, are obtained recursively from a dynamic model. A major role of the dynamic model is to flow local information to all states in the domain of interest spatially, temporally or both. Measurements are the information obtained from field observations, which may be directly, or indirectly related to the state. In most cases, neither forecasts nor measurements are perfect, but each contains some valuable information. The key idea of data assimilation is to blend the two sources of imperfect information in order to obtain the statistically optimal characterization of the system state.

In the Earth science community, well-known data assimilation techniques include 3DVAR and 4DVAR [23, 24, 123], the representer method [9, 10], the approximated Grid-Based methods [35, 86, 100, 109], the Kalman Filter [43], the Extended Kalman Filter [35, 86], the Ensemble Kalman Filter [16, 36, 38, 39], and the Partilcle Filer [6, 132]. Among the many techniques mentioned, the Ensemble Kalman Filter (EnKF) algorithm has rapidly gained popularity and has been utilized in numerous applications [38, 74, 77, 108, 110]. The EnKF-based algorithms are attractive for various reasons. First, their sequential structures are convenient and efficient for processing measurements in real-time. Second, it uses an ensemble characterization of the state, which provides the distributional information and uncertainty. Third, it is relatively easy to implement and is applicable to a wide range of dynamic models without the need to derive analytical forms of the state or its statistics. Finally, there is no restriction on the form of process noise (e.g., noise can be non-additive and correlated to the state.)

4.2.1 Ensemble Kalman Filter (EnKF)

The Ensemble Kalman filter (EnKF) is a forward sequential data assimilation method based on the Monte Carlo technique. It is first introduced in [36] and later clarified in [16]. The EnKF algorithm integrates an ensemble of model states and propagates them forward through time using the dynamic model. Therefore, for a large enough ensemble size, it is possible to construct the probability density from the ensemble and calculate any necessary statistics. The forecast state ensemble obtained from the dynamic model can then be recursively updated with new measurement data by using the Kalman Filter analysis scheme [43]. Having an ensemble of the states eliminates the need to propagate the covariance matrix analytically through time, which is difficult for many non-linear dynamic models. For this reason, the EnKF algorithm is applicable to a wide range of models, easy to implement, relatively effective to compute and accurate enough for many applications.

The propagation, or forecast, is made through the state or dynamic equation:

$$x_{t+1} = f_t\left(x_t, u_t, w_t\right) \tag{4.1}$$

where $x_t \in \Re^{n\times 1}$ is the state vector of dimension *n* at time *t*, $f_t(\cdot)$ is a dynamic function, u_t , and w_t are the forcing variables and the process noise at time *t*, respectively. The state ensemble is propagated to the time step where a new measurement becomes available. We assume that the measurement is related to the state via the measurement equation,

$$y_t = h_t\left(x_t\right) + v_t \tag{4.2}$$

where $y_t \in \Re^{m\times 1}$ is the measurement vector of dimension *m* at time *t*, $h_t(\cdot): \Re^{n\times 1} \to \Re^{m\times 1}$ is a measurement transformation matrix, and $v_t \in \Re^{m\times 1}$ is independent measurement noise with zero mean and covariance matrix **R**. At this measurement time, we employ the analysis (update or filter) step to incorporate new measurements into the forecast ensemble and produce the analysis ensemble. The analysis is done by using the first two moments of the prior density and the measurement likelihood function, which can be numerically estimated from the ensemble. Finally after the analysis ensemble is obtained, we again apply the forecast step to propagate the ensemble forward to the next measurement time step and perform the analysis step; a conceptual diagram of the EnKF is shown in Figure 4-1. The process repeats until the ensemble reaches the final time step.

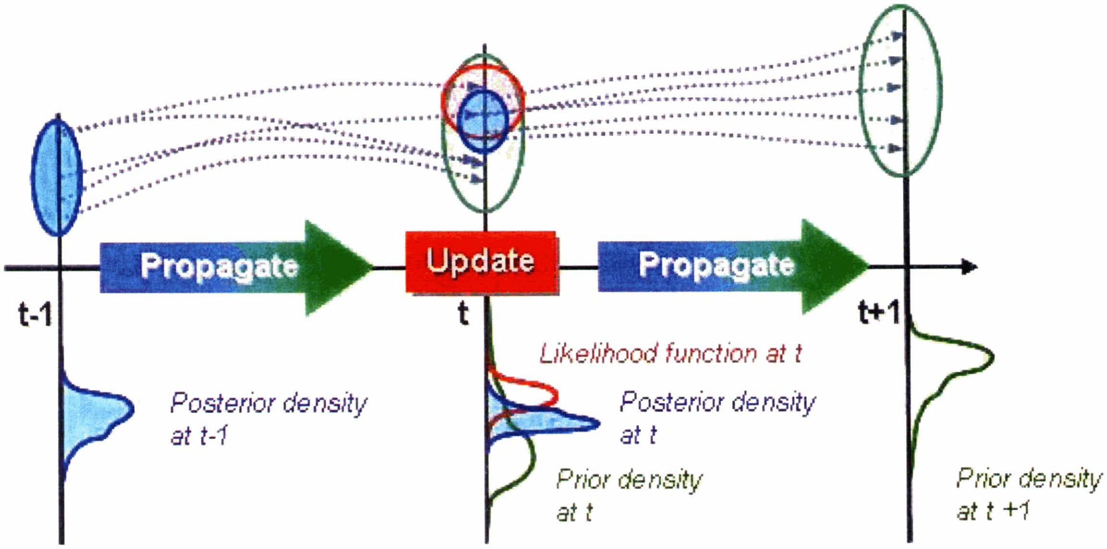


Figure 4-1: The conceptual diagram of the Ensemble Kalman Filter

The analysis step in EnKF is similar to that of the Kalman Filter [43]. It uses the mean and covariance matrix to calculate Bayes' linear least square estimator. The EnKF incorporates the current measurement y_t into each forecast state member $x_i^f \equiv x^i(t\,|\,t-1)$ and produces the analysis state $x_i^a \equiv x^i(t\,|\,t)$ from the analysis or update equation,

$$x_i^a = x_i^f + \mathbf{K}\left(\tilde{y}_i - h(x_i^f)\right) \tag{4.3}$$

where $\tilde{y}_i$ represents a perturbed measurements obtained by adding a random measurement noise v_t to the real measurement y_t, e.g. $\tilde{y}_i = y + v_i$. The Kalman gain matrix K is given by

$$\mathbf{K} = \mathbf{P}^{x,h(x)}\left(\mathbf{P}^{h(x)} + \mathbf{R}\right)^{-1} \tag{4.4}$$

where $\mathbf{P}^{x,h(x)} \in \Re^{n\times m}$ is the cross covariance between the forecast state x^f and the measurement prediction $h(x^f)$, and $\mathbf{P}^{h(x)} \in \Re^{m\times m}$ is the covariance matrix of the measurement prediction. These covariance and cross-covariance matrices are defined theoretically in terms of the true state x^t as

$$\mathbf{P}^{x,h(x)} = \overline{\left(x^f - x^t\right)\left(h(x^f) - h(x^t)\right)^T} \tag{4.5}$$

$$\mathbf{P}^{h(x)} = \overline{\left(h(x^f) - h(x^t)\right)\left(h(x^f) - h(x^t)\right)^T} \tag{4.6}$$

where $\overline{\bullet}$ represents the mean value. However, since the true state is usually unknown, the EnKF approximates the covariance and cross-covariance matrices in (4.4) with the ensemble covariance and cross-covariance matrices, taken around the ensemble mean $\bar{x}$ and $\overline{h(x)}$, e.g.

$$\mathbf{P}^{x,h(x)} \approx \mathbf{P}_e^{x,h(x)} = \tfrac{1}{N}\sum_{i=1}^{N}\left(x_i^f - \overline{x}^f\right)\left(h(x_i^f) - \overline{h(x^f)}\right)^T \tag{4.7}$$

$$\mathbf{P}^{h(x)} \approx \mathbf{P}_e^{h(x)} = \tfrac{1}{N}\sum_{i=1}^{N}\left(h(x_i^f) - \overline{h(x^f)}\right)\left(h(x_i^f) - \overline{h(x^f)}\right)^T \tag{4.8}$$

where N is the ensemble size. Thus, we can numerically calculate the Kalman gain in (4.4) by using the ensemble covariance and finally update the state ensemble using the analysis (4.3). Using the ensemble covariance matrices implies that we view the ensemble mean as the best estimate of the state and the spreading of the ensemble around the mean as an error in the ensemble mean.

There are several caveats about the analysis scheme of the EnKF. First, it is important that the update equation (4.3) uses the randomly perturbed measurement $\tilde{y}_i$ in order to retain the correct posterior covariance or the ensemble spread. If the real measurement value y is used, the posterior ensemble covariance will be underestimated [16]. Secondly, all of the statistics needed to calculate Kalman gain matrix, **K**, can be calculated directly from ensemble spreads. However, explicit calculation of the inverted term in (4.4) is computationally expensive. Furthermore, the inversion can be numerically ill-conditioned when the state dimension is large.

Evensen proposed a pseudo-inversion technique to be used with the EnKF algorithm [38]. The technique is based on singular value decomposition avoids explicit inversion of the full covariance matrix and reduces the computation cost from the order of m^2 to mN (i.e., m is the measurement dimension and N is the ensemble size). It makes the EnKF algorithm even more appealing and practical. However, when the ensemble size is smaller than the measurement dimension (e.g., $N < m$), this pseudo-inverse technique will be rank-deficient. The ensemble also collapses to a single member in the common situation in which the number of observations is more than twice the number of ensemble members (e.g., $N < 2m$) [66]. This is a serious issue in many large-scale data assimilation problems where the ensemble size is limited.

The next important issue is the sub-optimality of the EnKF algorithm. When using large numbers of ensemble members, the EnKF algorithm still produces a sub-optimal analysis ensemble because it ignores higher moments in the analysis step. The update (4.3) only uses the mean and covariance to update the forecast ensemble. Thus, the analysis ensemble is the linear least-squares estimator, and it is optimal only when all underlying distributions are Gaussian. In general, the importance of the higher moments varies with the processes of interest and their underlying dynamics. The first two moments usually contain the most important characteristics of a process. In addition, it is possible to transform the state in order to make the first two moments dominant over others (e.g., logarithmic transformation.). Consequently, updating the forecast ensemble using the 1st and the 2nd moment statistic can be sufficient and accurate enough in many applications.

The final consideration is the ensemble size. The EnKF uses the ensemble to calculate all necessary statistics in the analysis scheme. Using an insufficient ensemble size leads to sampling error and may cause instability in the analysis step. An appropriate ensemble size depends on many factors, including behavior of the dynamic model, and dimension and resolution of the states and measurements [14, 15, 30, 87]. Theoretically, the ensemble size should be as large as possible relative to the state dimension to minimize the sampling error and accurately approximate all statistics of interest. However, the ensemble size is normally limited by computation time, computation cost, and storage. Commonly, an ensemble size is far less than the state or the measurement dimension. In this case, the ensemble generated from the EnKF is prone to collapsing [66] or the update can become unreliable. The stability issue of the EnKF can be minimized by using a proper sampling scheme or performing a more stable pseudo-inverse technique. A stable pseudo-inverse technique is presented in the next section. The traditional SVD scheme used for minimizing sampling errors is included in Appendix C. It is highly recommended to consider the stable pseudo-inverse technique when one chooses to use the EnKF with a limited ensemble size.

4.2.2 The Stable Pseudo-Inversion Technique and the Stable EnKF

The EnKF algorithm is a simple and effective sequential analysis scheme that can be used in many applications. However, the Kalman gain matrix in (4.4) requires the inversion of the $m \times m$ matrix$\left(\mathbf{P}^{h(x)} + \mathbf{R}\right)$, which can be very computationally expensive for a large problem. Moreover, the inversion of a large matrix can become numerically ill-conditioned. In this section, we introduce a stable pseudo-inverse technique initially introduced in [39] that offers a significant improvement over the pseudo-inverse technique proposed in [38]. This stable pseudo-inverse technique is initially used in the square-root analysis scheme [39], which is an alternative algorithm to the EnKF. However, for our problem, the original formulation of the EnKF is preferable over the square-root formulation because it is easier to extend to the smoother form, which can be beneficial in the reanalysis problem. Moreover, the EnKF updates each individual ensemble member, while the square-root analysis scheme updates the ensemble mean and later adds a random perturbation. Although the random perturbation is guaranteed to match the theoretical value, its spatial pattern may be too random and scattered, and thus not suitable for our rainfall application. More information about the square-root analysis scheme is given in [39].

To derive the pseudo-inverse technique, we define the ensemble matrix $\mathbf{X} \in \Re^{n \times N}$ that holds all the forecast members $x_i^f \in \Re^{n \times 1}$ as

$$\mathbf{X} = \left[x_1^f, x_2^f, \ldots, x_N^f\right] \tag{4.9}$$

where n is the state dimension, and N represents the ensemble size. In addition, we define the ensemble mean $\overline{\mathbf{X}} \in \Re^{n \times N}$ and the ensemble perturbation $\tilde{\mathbf{X}} \in \Re^{n \times N}$ as

$$\overline{\mathbf{X}} = \mathbf{X}\mathbf{1}_N \tag{4.10}$$

$$\tilde{\mathbf{X}} = \mathbf{X} - \bar{\mathbf{X}} = \mathbf{X}(\mathbf{I} - \mathbf{1}_N) \tag{4.11}$$

where $\mathbf{1}_N \in R^{N \times N}$ is the averaging matrix with each element and is equal to *1/N*. Then, the forecast ensemble covariance matrix $\mathbf{P}_e \triangleq \mathbf{P}_e^f \in \Re^{n \times n}$ can be defined as,

$$\mathbf{P}_e = \frac{\tilde{\mathbf{X}}\tilde{\mathbf{X}}^T}{N-1} \tag{4.12}$$

From (4.2), we define the matrix $\mathbf{Y} \in \Re^{m \times N}$ that holds all the perturbed measurement, and the matrix $\tilde{\mathbf{Y}} \in \Re^{m \times N}$ that holds all the measurement perturbation v_i by

$$\mathbf{Y} = [y + v_1, y + v_2, \ldots, y + v_N] \tag{4.13}$$

$$\tilde{\mathbf{Y}} = \mathbf{Y} - \bar{\mathbf{Y}} = [v_1, v_2, \ldots, v_N] \tag{4.14}$$

where m is the measurement dimension, and y is a measurement vector at the analysis time. Moreover, we use the ensemble representation of the measurement error covariance $\mathbf{R}_e \in \Re^{m \times m}$ to approximate the true measurement error covariance $\mathbf{R}$ by

$$\mathbf{R}_e = \frac{\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}^T}{N-1} \tag{4.15}$$

Finally, we define $\mathbf{S} \in \Re^{m \times N}$ to be the matrix that holds the prediction of measurements given the ensemble state, and its perturbation from the mean $\tilde{\mathbf{S}} \in \Re^{m \times N}$ as

$$\mathbf{S} = \left[h\left(x_1^f\right), h\left(x_2^f\right), \ldots, h\left(x_N^f\right)\right] \tag{4.16}$$

$$\tilde{\mathbf{S}} = \mathbf{S} - \bar{\mathbf{S}} = \mathbf{S}(\mathbf{I} - \mathbf{1}_N) \tag{4.17}$$

Using these definitions, the Kalman gain given in (4.4) can be approximated by the ensemble covariance matrix from (4.7) and (4.8), e.g.

$$\mathbf{K} \approx \mathbf{P}_e^{x,h(x)}\left(\mathbf{P}_e^{h(x)} + \mathbf{R}_e\right)^{-1} \tag{4.18}$$

$$= \frac{\tilde{\mathbf{X}}\tilde{\mathbf{S}}^T}{N-1}\left[\frac{\tilde{\mathbf{S}}\tilde{\mathbf{S}}^T + \tilde{\mathbf{Y}}\tilde{\mathbf{Y}}^T}{N-1}\right]^{-1} \tag{4.19}$$

$$= \tilde{\mathbf{X}}\tilde{\mathbf{S}}^T\mathbf{C}^{-1} \tag{4.20}$$

with,

$$\mathbf{C} = \tilde{\mathbf{S}}\tilde{\mathbf{S}}^T + \tilde{\mathbf{Y}}\tilde{\mathbf{Y}}^T \tag{4.21}$$

We seek the pseudo-inverse $\mathbf{C}^+ \in \Re^{m\times m}$ to approximate the inverse matrix $\mathbf{C}^{-1}$. The stable pseudo-inverse technique projects $\tilde{\mathbf{Y}}$ onto the first *N-1* singular vectors of $\tilde{\mathbf{S}}$. Thus, we only account for the measurement variance contained in the subspace and reject all possible contributions in the null space. It is this property of the pseudo-inverse which avoids the rank deficient issue and prevents the ensemble from collapsing [66]. We begin by taking the full-sized singular value decomposition of $\tilde{\mathbf{S}}$, whose rank equals *N-1*,

$$\mathbf{U}_0\Sigma_0\mathbf{V}_0^T = \tilde{\mathbf{S}} \tag{4.22}$$

where $\mathbf{U}_0 \in \Re^{m\times m}$, $\mathbf{\Sigma}_0 \in \Re^{m\times N}$, and $\mathbf{V}_0 \in \Re^{N\times N}$. By definition, the product of any matrix with its pseudo-inverse equals the identity matrix with the first q elements equal to one and the others equal to zero; q represents the shorter dimension of the matrix. The pseudo-inverse of $\tilde{\mathbf{S}}$ is given by

$$\tilde{\mathbf{S}}^+ = \mathbf{V}_0\Sigma_0^+\mathbf{U}_0^T \tag{4.23}$$

$\mathbf{\Sigma}_0^+ \in \Re^{N\times m}$ is a diagonal matrix whose first *N-1* diagonal elements are the inverse of $\mathbf{\Sigma}_0$ and the remaining elements are zeros:

$$\mathbf{\Sigma}_0^+ = \begin{bmatrix} \sigma_1^{-1} & & & \\ & \ddots & & \\ & & \sigma_{N-1}^{-1} & \\ & & & \underline{\mathbf{0}}_{p\times p} \end{bmatrix} \tag{4.24}$$

where $p = m\text{-}(N+1)$. We can then express $\mathbf{C}$ from (4.21) as

$$\mathbf{C} = \mathbf{U}_0\mathbf{\Sigma}_0\mathbf{\Sigma}_0^T\mathbf{U}_0^T + \tilde{\mathbf{Y}}\tilde{\mathbf{Y}}^T \tag{4.25}$$

$$= \mathbf{U}_0\left(\mathbf{\Sigma}_0\mathbf{\Sigma}_0^T + \mathbf{U}_0^T\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}^T\mathbf{U}_0\right)\mathbf{U}_0^T \tag{4.26}$$

$$\approx \mathbf{U}_0\mathbf{\Sigma}_0\left(\mathbf{I} + \mathbf{\Sigma}_0^+\mathbf{U}_0^T\tilde{\mathbf{Y}}\tilde{\mathbf{Y}}^T\mathbf{U}_0\mathbf{\Sigma}_0^{+T}\right)\mathbf{\Sigma}_0^T\mathbf{U}_0^T \tag{4.27}$$

$$= \mathbf{U}_0\mathbf{\Sigma}_0\left(\mathbf{I} + \mathbf{K}_1\mathbf{K}_1^T\right)\mathbf{\Sigma}_0^T\mathbf{U}_0^T \tag{4.28}$$

where $\mathbf{K}_1 \in \Re^{N\times N}$ is given by

$$\mathbf{K}_1 = \mathbf{\Sigma}_0^+\mathbf{U}_0^T\tilde{\mathbf{Y}} \tag{4.29}$$

We then take the singular value decomposition of $\mathbf{K}_1$

$$\mathbf{U}_1\Sigma_1\mathbf{V}_1^T = \mathbf{K}_1 \tag{4.30}$$

with all matrices having dimension *N*×*N*. Finally, we substitute (4.30) into (4.28) and obtain the pseudo-inverse of $\mathbf{C}$ in the following form:

$$\mathbf{C}^+ = \left[\mathbf{U}_0\mathbf{\Sigma}_0\left(\mathbf{I} + \mathbf{U}_1\Sigma_1^2\mathbf{U}_1^T\right)\mathbf{\Sigma}_0^T\mathbf{U}_0^T\right]^+ \tag{4.31}$$

$$= \left[\mathbf{U}_0 \mathbf{\Sigma}_0 \mathbf{U}_1 \left(\mathbf{I} + \Sigma_1^2 \right) \mathbf{U}_1^T \mathbf{\Sigma}_0^T \mathbf{U}_0^T \right]^+ \tag{4.32}$$

$$= \left(\mathbf{U}_1^T \mathbf{\Sigma}_0^T \mathbf{U}_0^T \right)^{+T} \left(\mathbf{I} + \Sigma_1^2 \right)^{+T} \left(\mathbf{U}_0 \mathbf{\Sigma}_0 \mathbf{U}_1 \right)^{+T} \tag{4.33}$$

$$= \left(\mathbf{U}_0 \mathbf{\Sigma}_0^{+T} \mathbf{U}_1 \right) \left(\mathbf{I} + \Sigma_1^2 \right)^{-1} \left(\mathbf{U}_0 \mathbf{\Sigma}_0^{+T} \mathbf{U}_1 \right)^T \tag{4.34}$$

$$= \mathbf{K}_2 \left(\mathbf{I} + \Sigma_1^2 \right)^{-1} \mathbf{K}_2^T \tag{4.35}$$

where $\mathbf{K}_2 \in \Re^{m \times N}$ of rank $N\text{-}1$ is defined by

$$\mathbf{K}_2 = \mathbf{U}_0 \mathbf{\Sigma}_0^{+T} \mathbf{U}_1 \tag{4.37}$$

Note that (4.27) requires that $\mathbf{U}_0 \mathbf{U}_0^T = \mathbf{I}$, which is true only if we perform the full-sized singular value decomposition of $\tilde{\mathbf{S}}$ as given in (4.22). However, for $\mathbf{K_1}$ and $\mathbf{K_2}$ in (4.28) and (4.29), we can neglect the last N-m singular vectors in $\mathbf{U}_0$ because of the multiplication with $\mathbf{\Sigma}_0^+$. Thus, the reduced-sized singular value decomposition of $\tilde{\mathbf{S}}$, e.g. $\mathbf{U}_0 \Sigma_0 \mathbf{V}_0^T = \tilde{\mathbf{S}}$ with $\mathbf{U}_0 \in \Re^{m \times N}$, $\mathbf{\Sigma}_0 \in \Re^{N \times N}$, and $\mathbf{V}_0 \in \Re^{N \times N}$, can be used without loss of generality. The reduce-sized SVD will significantly speed up the computation time.

Now that the stable pseudo-inverse $\mathbf{C}^+$ matrix is used to approximate $\left(\mathbf{P}^{h(x)} + \mathbf{R} \right)^{-1}$, we can compute the Kalman gain matrix from equation (4.4) and update the forecast ensemble matrix $\mathbf{X}$ using the analysis equation (4.3). Let the matrix $\mathbf{X}^a \in \Re^{n \times N}$ hold all the ensemble analysis members $x_i^a \in \Re^{n \times 1}$ after the update, e.g.

$$\mathbf{X}^a = \left[x_1^a, x_2^a, \ldots, x_N^a \right] \tag{4.38}$$

The analysis (4.3) can be written in the ensemble form as

$$\mathbf{X}^a = \mathbf{X} + \tilde{\mathbf{X}}\tilde{\mathbf{S}}^T\mathbf{C}^+\left(\mathbf{Y}-\mathbf{S}\right) \tag{4.39}$$

$$= \mathbf{X} + \tilde{\mathbf{X}}\mathbf{K}_3 \tag{4.40}$$

$$= \mathbf{X} + \mathbf{X}\left(\mathbf{I}-\mathbf{1}_N\right)\mathbf{K}_3 \tag{4.41}$$

$$= \mathbf{X}\left(\mathbf{I}+(\mathbf{I}-\mathbf{1}_N)\mathbf{K}_3\right) \tag{4.42}$$

$$= \mathbf{X}\mathbf{K}_4 \tag{4.43}$$

where the matrices **K**$_3$ and **K**$_4$ are given by

$$\mathbf{K}_3 = \tilde{\mathbf{S}}^T\mathbf{C}^+\left(\mathbf{Y}-\mathbf{S}\right) \tag{4.44}$$

$$\mathbf{K}_4 = \mathbf{I}+(\mathbf{I}-\mathbf{1}_N)\mathbf{K}_3 \tag{4.45}$$

Note that $\mathbf{1}_N\mathbf{K}_3$ gives the row average of **K**$_3$. The analysis equation of the stable EnKF algorithm in the form of (4.43) implies that the analysis ensemble is a weakly non-linear combination of the forecast ensemble. Each column of the update matrix **K**$_4$ represents weights from each forecast member and is given by the projection of measurement onto the forecast ensemble space. In order for the estimate to be unbiased, the sum of each column of **K**$_4$ should be one. In addition, the diagonal elements of **K**$_4$ should be dominant because they hold the weight for the first-guess ensemble member, while off-diagonal elements introduce correlations imposed by the measurement.

Conclusively, the analysis ensemble $\mathbf{X}^{\mathbf{a}}$ can be obtained from the stable EnKF by the following steps:

1) At the analysis time step, construct ensemble matrices: **X**, $\tilde{\mathbf{X}}$, **S**, $\tilde{\mathbf{S}}$, **Y**, and $\tilde{\mathbf{Y}}$.
2) Compute the reduced-sized SVD: $\mathbf{U}_0\Sigma_0\mathbf{V}_0^T = \tilde{\mathbf{S}}$ where, $\mathbf{U}_0 \in \Re^{m\times p}$, $\Sigma_0 \in \Re^{p\times N}$, $\mathbf{V}_0 \in \Re^{N\times N}$, and $p = \min\left(m, N\right)$

3) Form the pseudo-inverse diagonal matrix $\mathbf{\Sigma}_0^+ \in \Re^{p\times p}$ by inverting the first p diagonal elements of $\mathbf{\Sigma}_0$, i.e. $diag\left(\mathbf{\Sigma}_0^+\right)=\left(\sigma_1^{-1},\sigma_2^{-1},...,\sigma_p^{-1}\right)$
4) Compute the matrix product: $\mathbf{K}_1=\mathbf{\Sigma}_0^+\mathbf{U}_0^T\tilde{\mathbf{Y}}$ where $\mathbf{K}_1 \in \Re^{p\times N}$
5) Compute the reduced-sized SVD: $\mathbf{U}_1\mathbf{\Sigma}_1\mathbf{V}_1^T=\mathbf{K}_1$ where $\mathbf{U}_1 \in \Re^{p\times p}$, $\mathbf{\Sigma}_1 \in \Re^{p\times N}$, and $\mathbf{V}_1 \in \Re^{N\times N}$.
6) Form the matrix product: $\mathbf{K}_2=\mathbf{U}_0\mathbf{\Sigma}_0^{+T}\mathbf{U}_1$ where $\mathbf{K}_2 \in \Re^{m\times p}$.
7) Compute the pseudo-inverse: $\mathbf{C}^+=\mathbf{K}_2\left(\mathbf{I}+\Sigma_1^2\right)^{-1}\mathbf{K}_2^T$ where $\mathbf{C}^+ \in \Re^{m\times m}$
8) Compute the matrix product: $\mathbf{K}_3=\tilde{\mathbf{S}}^T\mathbf{C}^+\left(\mathbf{Y}-\mathbf{S}\right)$ where $\mathbf{K}_3 \in \Re^{N\times N}$
9) Compute the update matrix: $\mathbf{K}_4=\mathbf{I}+(\mathbf{I}-\mathbf{1}_N)\mathbf{K}_3$ where $\mathbf{K}_4 \in \Re^{N\times N}$
10) Finally, obtain the analysis ensemble matrix: $\mathbf{X}^a=\mathbf{X}\mathbf{K}_4$

The benefit of using the stable EnKF algorithm with the stable pseudo-inverse is significant when the ensemble size is relatively small in comparison to the state or the measurement dimension, e.g. *N* << *m, n.* To illustrate the advantage of the stable EnKF algorithm, a sample experiment is shown in Figure 4-2. The forecast ensemble is obtained by perturbing the mean and variance of the one-dimensional Gaussian function centered at zero with a variance of one. In this sample, the size of the state dimension (*n*) is *200,* the size of the measurement dimension (*m*) is 20, and the size of the ensemble (*N*) is 10. Figure 4-2a shows the forecast ensemble mean as blue dots, the truth as a black solid line and the measurements with red circles. Figures 4-2b and 4-2c show the analysis ensemble mean obtained from the original EnKF [38] and the stable EnKF, respectively. Figures 4-2d, 4-2e, and 4-2f show the ensemble covariance of the forecast, the analysis from the original EnKF, and the analysis from the stable EnKF, respectively. It is evident in this example that the original EnKF diverges when the ensemble size is relatively much smaller than the state size and measurement size. In contrast, the stable EnKF performs well and produces reliable analysis mean and covariance.

Figure 4-3 shows a sample experiment result when the ensemble size is relatively large. In this sample, the size of the state dimension (*n*) is *200*, the size of the measurement dimension (*m*) is 20, and the size of the ensemble (*N*) is 500. The analysis mean and covariance from the original EnKF and the stable EnKF (e.g., Figure 4-3b vs. 4-3c, and Figure 4-3f vs. 4-3g) are very similar. In this case, both algorithms provide accurate analysis ensemble.

In our rainfall data assimilation over the USGP region, we would like to obtain the rainfall estimate at the resolution of 0.05 degree. This means we are propagating and updating the state at 475x475 pixel^2, which translates to a state dimension of around a quarter million. It is impractical to use an ensemble with a size close to the state or measurement dimension. Hence, we should always use the stable EnKF, and from this point forward, we will refer to the stable EnKF as the EnKF algorithm.

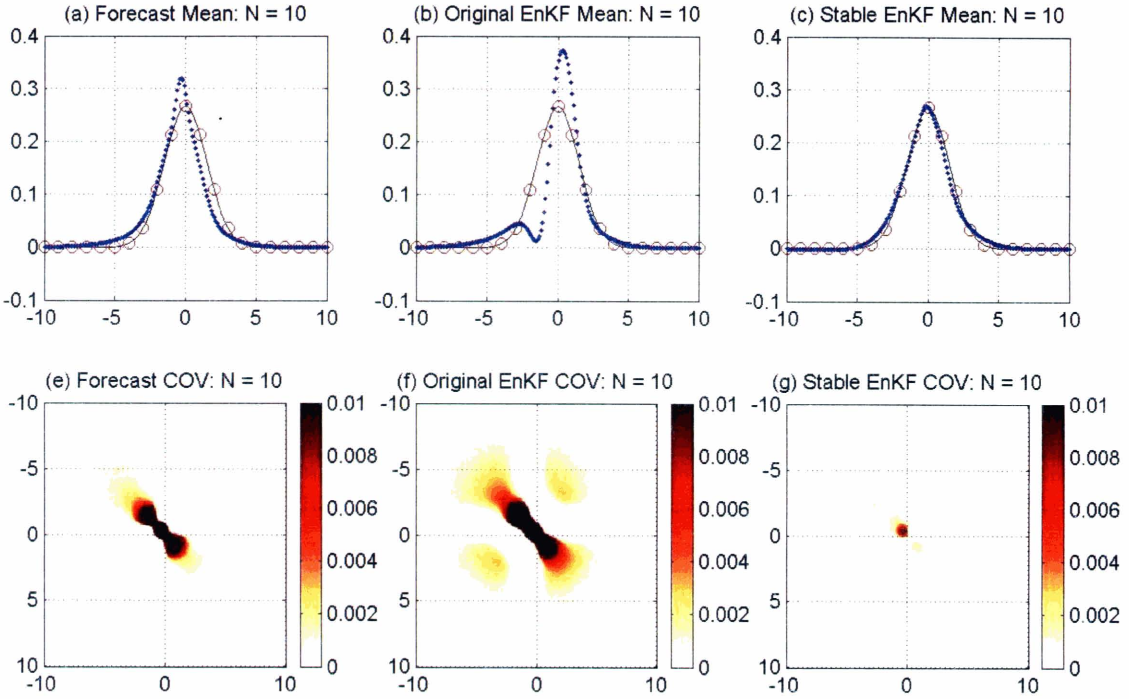


Figure 4-2: Performances of the original EnKF [38] and the stable EnKF when the ensemble size is small (e.g., *n = 200, m = 20, N = 10*)

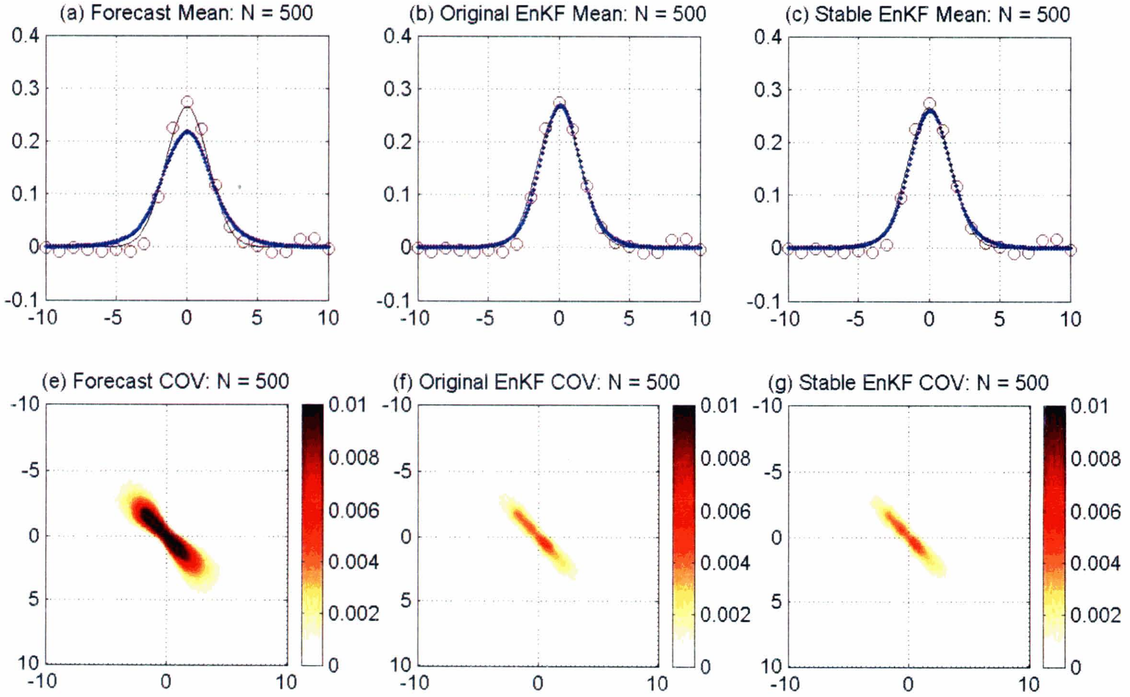


Figure 4-3: Performances of the original EnKF [38] and the stable EnKF when the ensemble size is large (e.g., *n = 200, m = 20, N = 500*)

4.2.3 Ensemble Kalman Smoother (EnKS)

The EnKF algorithm was presented previously is a forward sequential algorithm. As a state ensemble progresses forward through time, the EnKF algorithm updates it with a new measurement and provides an analysis ensemble. This analysis ensemble is then propagated to the next time step. At any moment, the ensemble only accounts for measurements before and at the estimation time. It cannot incorporate measurements after the estimation time. Thus, the EnKF algorithm is suitable for a real-time data assimilation problem where we are seeking the most recent state. For a reanalysis problem where the state of interest may be in the past, it would be more beneficial to account for all the measurements inside a given period in which we are interested.

To incorporate future measurement with a sequential algorithm, we first need to employ a filter algorithm that moves forward from an initial to a final time. Then we propagate information backward to the time of interest. This forward-backward sequential scheme is call a fixed-interval smoothing algorithm, since the state will be conditioned on all measurements in the fixed-time interval between the initial and the final time step. For a process with short memory, any measurement far away from the estimation time is less likely to affect the state at the current time. In other words, the improvement provided by the smoother is related to the system memory (e.g., the temporal persistence of the state). There is no need to apply the smoother over a time interval greater than the system memory. Therefore, it is sufficient to account only for a fixed amount of measurements after the estimate time, which should correspond to the system memory. This method is called a fixed-lag smoothing algorithm. A fixed-lag smoothing algorithm can become useful for a near-real-time problem where we can wait for some more measurements in the future to help with conditioning the state, or when we would like to minimize the storage. The concept of a filtering algorithm, a fixed-interval smoothing algorithm and a fixed-lag smoothing algorithm can be illustrated with a temporal diagram in Figure 4-4.

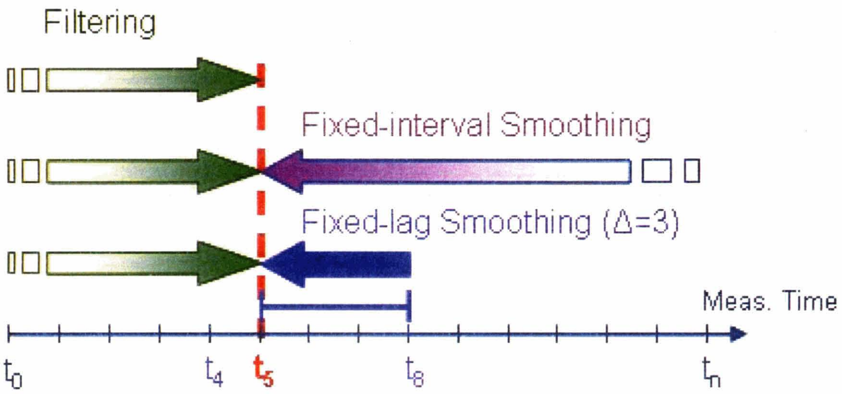


Figure 4-4: The temporal diagram of a filing, a fixed-interval smoothing, and a fixed-lag smoothing scheme

Theoretically, it is difficult to propagate information backward through time because the dynamic function $f_t(\cdot)$ in (4.1) is not necessarily invertible. However, it is possible to calculate a sub-optimal smoothed state ensemble by using the ensemble covariance in space and time, similar to the EnKF methodology. This smoothing algorithm is called Ensemble Kalman Smoother (EnKS), which is a straightforward extension of the EnKF [37, 38, 40]. The EnKS provides the analysis ensemble at time t' from measurements available at a later time t_i as

$$\mathbf{X}^a(t') = \mathbf{X}(t') + \tilde{\mathbf{X}}(t')\tilde{\mathbf{S}}^T(t_i)\mathbf{C}^{-1}(t_i)\left[\mathbf{Y}(t_i) - \mathbf{S}(t_i)\right] \tag{4.46}$$

where $\mathbf{Y}(t_i)$ from (4.13), $\mathbf{S}(t_i)$ from (4.16), $\tilde{\mathbf{S}}(t_i)$ from (4.17), and $\mathbf{C}(t_i)$ from (4.21) are evaluated using the ensemble and measurement at the future time t_i. From (4.46), it is obvious that the update at time t' uses the same combination of ensemble members as defined by $\mathbf{K}_4$ in (4.45) in the EnKF. Thus, the fixed-interval smoothing ensemble $\mathbf{X}_T^S(t') \in \Re^{n \times N}$ at time t' where $t_{i-1} \le t' < t_i < t_T$ is given by

$$\mathbf{X}_T^S(t') = \mathbf{X}^F(t')\prod_{j=i}^{T}\mathbf{K}_4(t_j) \tag{4.47}$$

where t_T is the final time step, $\mathbf{K}_4(t_j) \in \Re^{N \times N}$ is the filter update matrix from (4.45) evaluated at time t_j, and $\mathbf{X}^F(t') \in \Re^{n \times N}$ is the forward ensemble matrix from the EnKF algorithm at time t'. The forward ensemble matrix is equal to the analysis ensemble matrix $\mathbf{X}^a$ if there is an update at time t'; otherwise, it will equal the forecast ensemble matrix **X**.

Likewise, the fixed-lag smoothing ensemble $\mathbf{X}_\Delta^S(t') \in \Re^{n \times N}$ with a lag Δ at time t' where $t_{i-1} \le t' < t_i < t_{i+\Delta}$ is given by

$$\mathbf{X}_\Delta^S(t') = \mathbf{X}^F(t') \prod_{j=i}^{i+\Delta} \mathbf{K}_4(t_j) \tag{4.48}$$

As long as these filter update matrices $\mathbf{K}_4$'s during the period of interest are stored, and the columns of the filter ensemble $\mathbf{X}^F$ have not been shuffled, it is fast and straight-forward obtain the smoothed ensemble. It is also possible to store only some rows of $\mathbf{X}^F$ that represent particular state variables of interest, and apply the EnKS algorithm without storing the full ensemble matrix.

The post multiplication of the update matrix $\mathbf{K}_4$ will always result in a new ensemble with a different mean and a smaller variance. Consecutive smoothing will lead to a slight reduction of the variance and a slight change in the mean value. Despite the reduction in ensemble spread, the EnKS does not guarantee that the smoothed ensemble will be more accurate than the forecast in estimating the true state. Similar to the EnKF, the EnKS is sub-optimal because it only uses the first two moments and ignores higher moment information. Moreover, the EnKS tends to smooth abrupt changes in the state temporally. Thus, we expect that the EnKS to work best with a process that is temporally smooth or a process that has been updating very frequently.

4.2.4 Implementation of EnKF and EnKS on Synthetic Rainfall Problems

In this section, the EnKF and the EnKS are utilized to merge multiple sources of synthetic rainfall measurements and provide a comprehensive rainfall ensemble. We choose to perform a synthetic experiment before implementing the data assimilation on the USGP region for several reasons. First, the dynamic model is guaranteed to be correct because we generate our true state from the dynamic model used in the data assimilation. Second, the model parameters are known and controllable. Finally, we can generate synthetic experiments for any scenario we would like to test (e.g., missing in space and time.) By performing the synthetic experiments, we can pay full attention to the difference in the data assimilation results (e.g., forecast, filter vs. smoother) and minimize any uncertainty in the problem.

We use the recursive clustered rainfall (RCR) model proposed in Chapter 4 to propagate rainfall ensemble forward in time. In this synthetic experiment, we use the GOES cloud-top temperature and GOES velocity field from the USGP region as forcing variables for the RCR model. The synthetic study domain is 40 x 40 pixel^2 at a spatial resolution of 0.05 degree at a location and time where a deep convective storm is occurring. The true synthetic rainrate is generated from the RCR model with parameters β*=0.08 pixel*$^{-2}$ *hr*$^{-1}$, α *= 0.6 hr*$^{-1}$, σ *= 1.75 pixel*, *E[i*$_0$*] = 5 mm/hr*, *v = 25 cells/cluster*, and ρ*=2.* The GOES cloud-top temperature thresholds, T_G = 220 K, and T_R = 290 K, were obtained in Chapter 4 are are used to alleviate the rainfall intermittency problem.

We generate two types of measurements: (1) scattered but fine-scaled measurements, and (2) coarse-scaled measurement. These synthetic measurements are intended to duplicate the characteristics of rain-gauge and satellite measurements, respectively. Their values are taken from the true synthetic rainfall, but are perturbed by random noise as in equation (4.2). The covariance matrix $\mathbf{R} \in \Re^{m \times m}$ of the measurement noise is assumed to be a diagonal matrix and each element on the diagonal is given by

$$\mathbf{R}_i = \left(c_1 x_i + c_2\right)^2 \tag{4.53}$$

where x_i is the true rainfall at pixel i, c_1 and c_2 are 0.1 and 1, respectively. The true synthetic rainfall and the rainfall measurements from time t_1 to t_6 are illustrated in Figure 4-5. Note that the red boundary over the true rainfall represents the region with GOES cloud-top temperature lower than T_B, i.e. the region where new rain cells can be born.

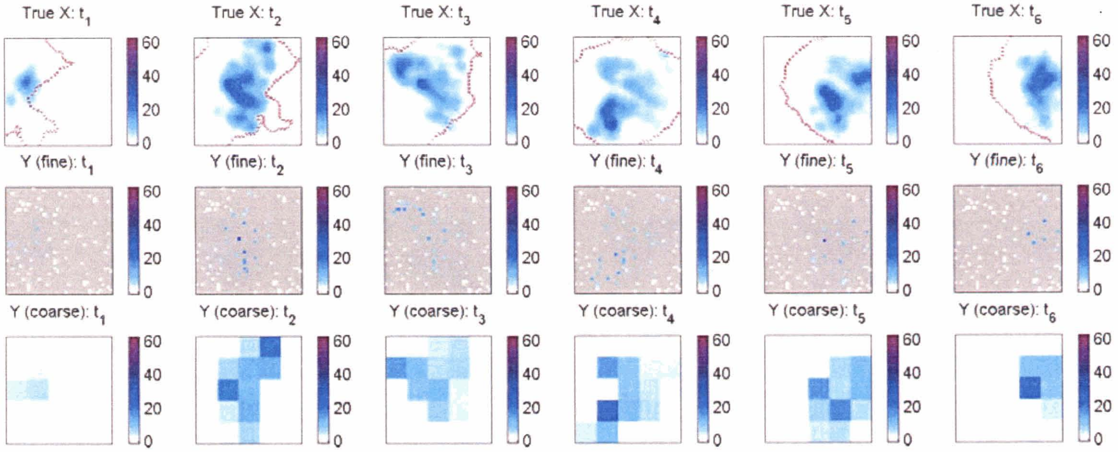


Figure 4-5: True synthetic rainfall and rainfall measurements to be used in experiment #1

In the first synthetic experiment (experiment #1), we take all measurement inputs given in Figure 4-5 and try to estimate the true state. We begin at time t_0 by generating 1000 realizations of zero rainfall fields and propagate them forward using the RCR model with all parameters known. At each time step, we update the forecast ensemble with the measurements using the EnKF algorithm. The propagation step and the update step are repeated until the update ensemble at the final time t_6 is acquired. Then we propagate the ensemble backward using the EnKS algorithm and calculate the fixed-interval smoothing. The results from the first experiment, which are the ensemble mean, the ensemble standard derivation (ensemble spread), and the root mean square error (RMSE) from the true synthetic rainfall, are given in Figures 4-6, 4-7, and 4-8, respectively. In Figure 4-5 the first row shows the synthetic true rainfall field, the 2nd row shows the forecast mean, the 3rd row shows the filter mean, and the 4th row shows the smoother mean. In Figures 4-7 and 4-8, the 1st row shows the forecast results, the 2nd row shows the update results, and the 3rd row shows the smoother results. Each column of Figures 4-6, 4-7, and 4-8 represents the time step from t_1 to t_6. In addition, the values in the parenthesis in Figures 4-7 and 4-8 represent the spatial average of ensemble spread and RMSE, respectively.

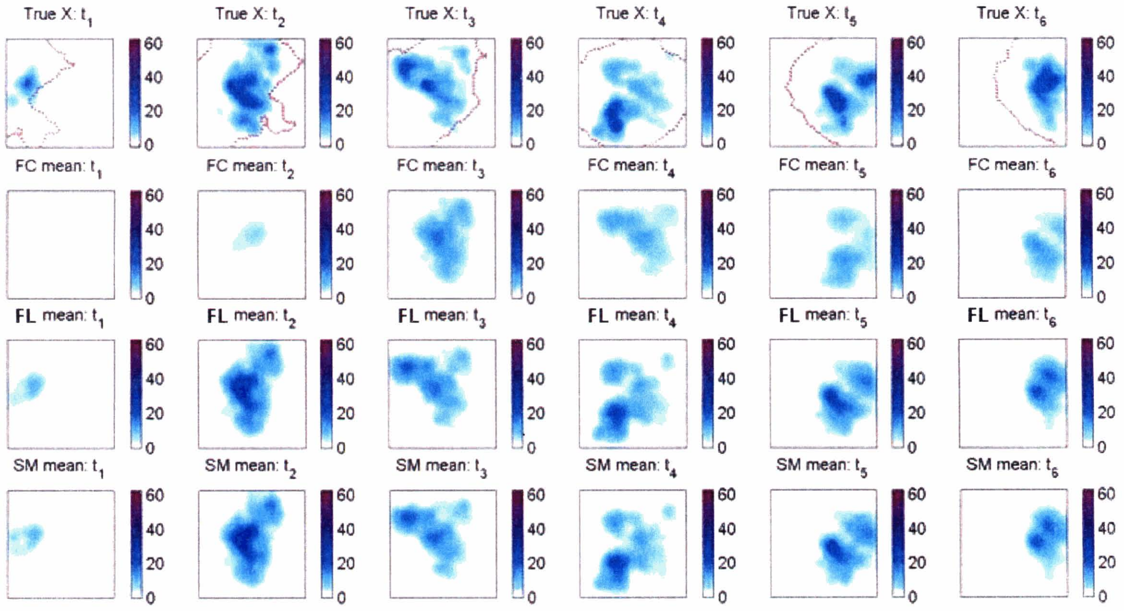


Figure 4-6: The synthetic truth and the ensemble mean of the forecast (FC), filter (FL), and smoother (SM) ensemble from experiment #1 from time t_1 to t_6

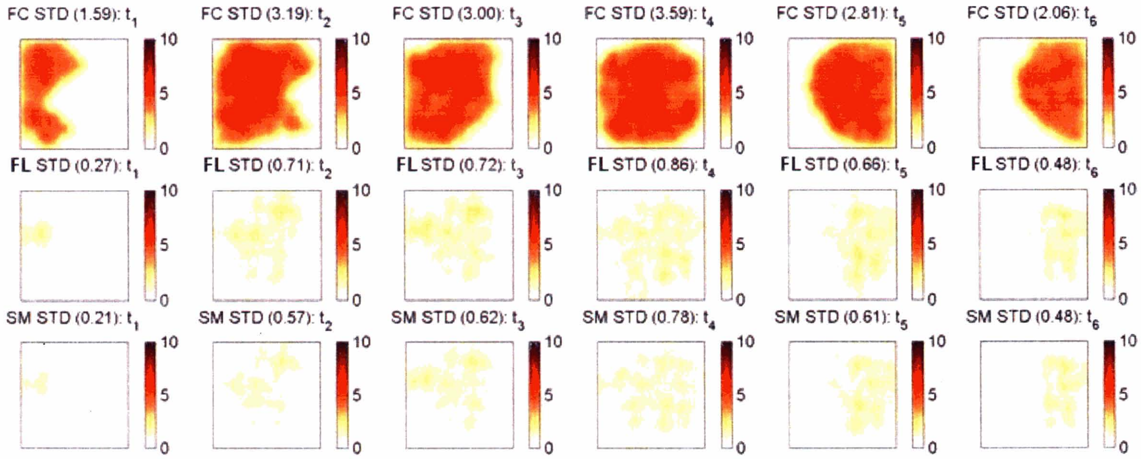


Figure 4-7: Ensemble standard deviation of the forecast (FC), filter (FL), and smoother (SM) ensemble from experiment #1 from time t_1 to t_6

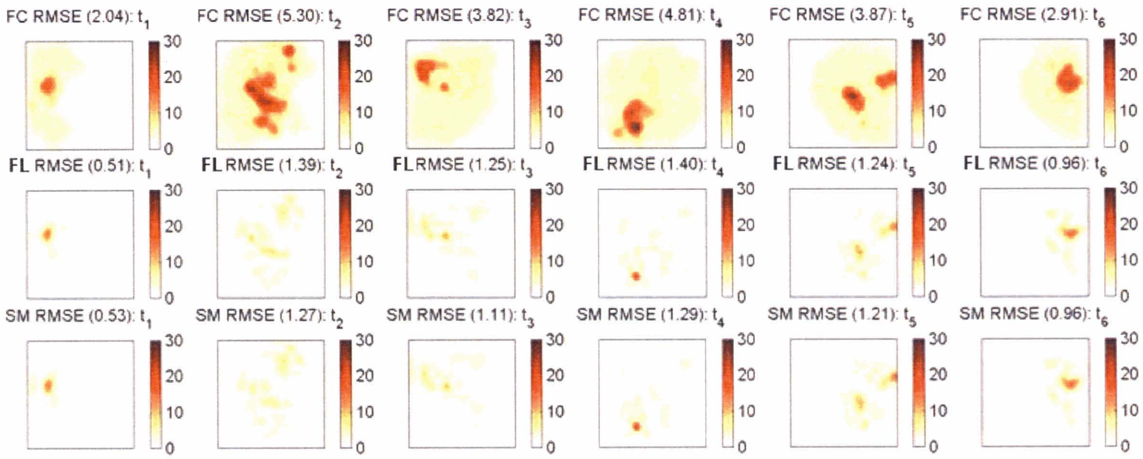


Figure 4-8: Root mean square error of the forecast (FC), filter (FL), and smoother (SM) ensemble from experiment #1 from time t_1 to t_6

The results show that we can use the EnKF and EnKS to merge multiple sources of rainfall measurements, and provide reliable results. In the first experiment, the EnKS provides only a slight improvement over the EnKF. This is because there is plenty of information provided at each time step, and thus, knowledge from other time steps are less significant to further improve the ensemble.

To see the benefit of the EnKS over the EnKF algorithm more clearly, we must reduce the amount of measurement information provided at each time step. We set up the second experiment (experiment #2) by withholding all measurements at time steps t_1, t_2, t_4, and t_5. Figure 4-9 shows the true synthetic rainfall and new measurements to be used in the second experiment. Because there is less information provided at time steps t_1, t_2, t_4, and t_5, we expect to see significant differences between the EnKF and EnKS results. At these time steps, information from their neighboring time steps become more significant for improving the ensemble.

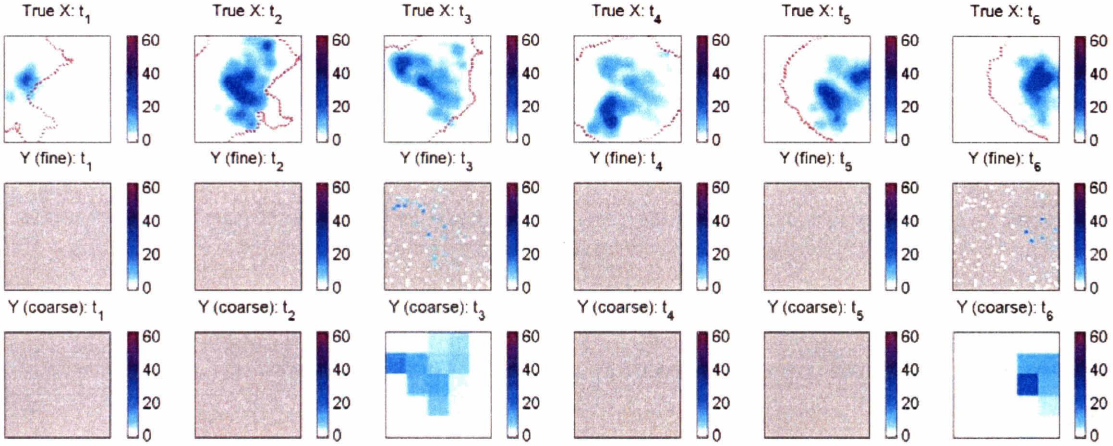


Figure 4-9: True synthetic rainfall and rainfall measurements in experiment #2

The results from the second experiment, which include the ensemble mean, the ensemble standard derivation (e.g., the ensemble spread), and the ensemble RMSE, are presented in Figure 4-10, 4-11, and 4-12, respectively. The format of these results is similar to that used for the first experiment (e.g., Figure 4-6, 4-7, and 4-8).

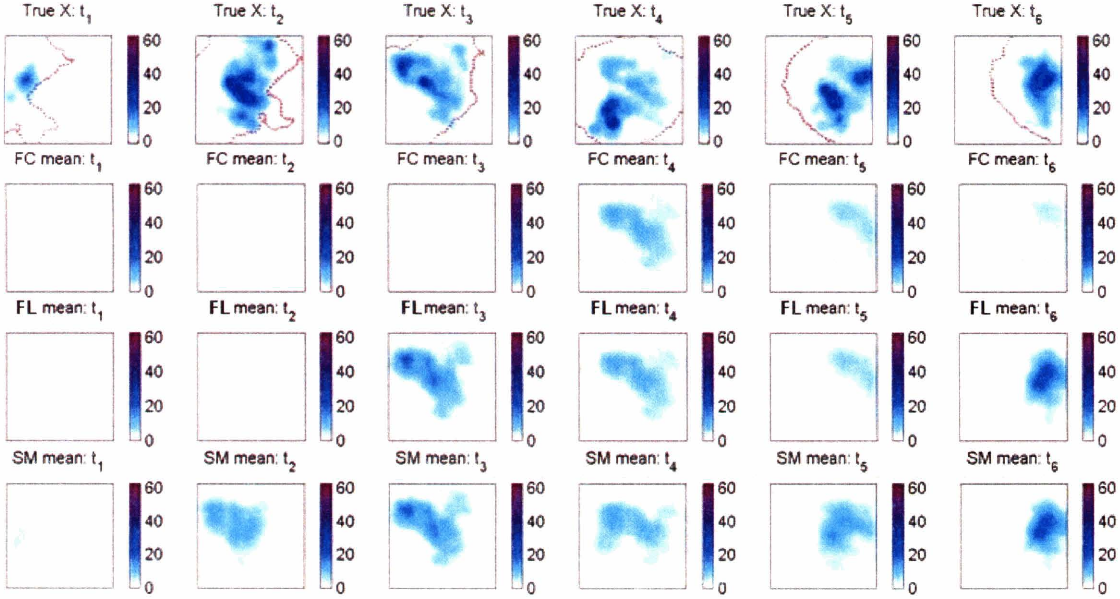


Figure 4-10: The synthetic truth and the ensemble mean of the forecast (FC), filter (FL), and smoother (SM) ensemble from experiment #2 from time t_1 to t_6

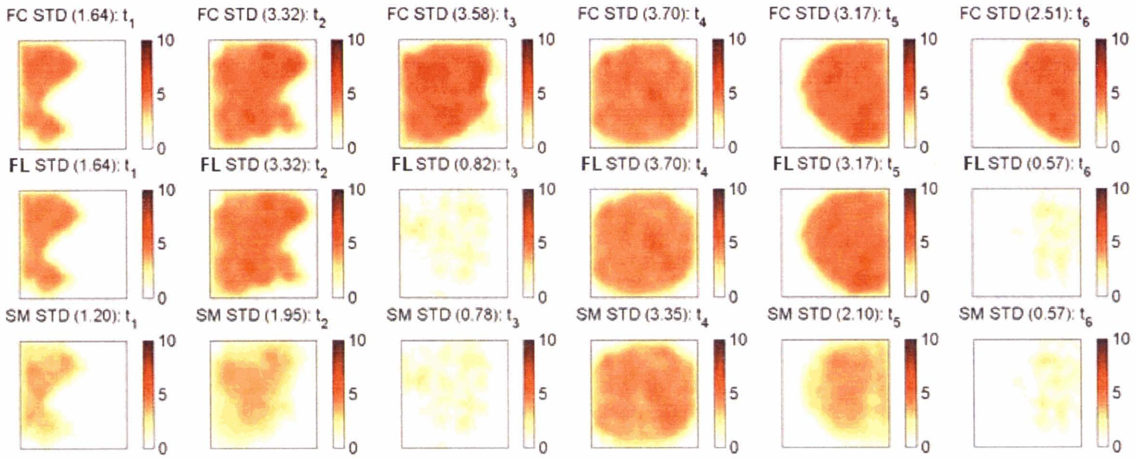


Figure 4-11: Ensemble standard deviation of the forecast (FC), filter (FL), and smoother (SM) ensemble from experiment #2 from time t_1 to t_6

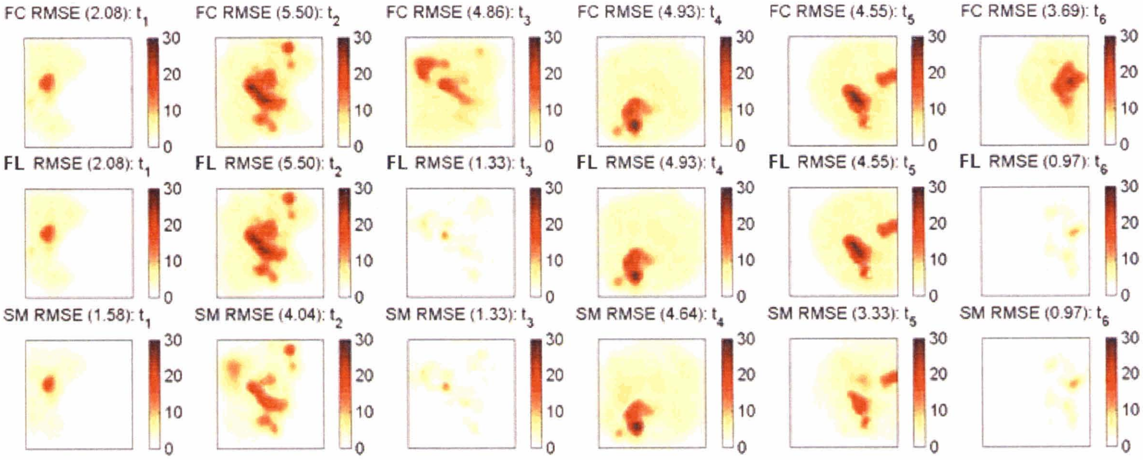


Figure 4-12: Root mean square error of the forecast (FC), filter (FL), and smoother (SM) ensemble from experiment #2 from time t_1 to t_6

From Figures 4-11 and 4-12, the ensemble spread and the RMSE of the smoother are less than those of the filter. The mean ensemble of the smoother is also closer to the synthetic true rainfall than the filter. It is evident that the smoother ensemble from the EnKS is superior to the filter ensemble from the EnKF. The differences are significant at times t_2 and t_4. These time steps are located before the next measurement times, and the smooth ensemble receives full benefit from propagating information backward. The contribution from the EnKS decreases as the measurement time is farther away. It thus seems that the improvement from EnKS is minimal when a significant amount of measurement information is given, e.g. at t_3. Note that the smoother and the filter ensemble at the final time will always be the same.

The smoother ensemble mean is not always closer to the true solution than the filter. The EnKS uses correlations between the current states and the later measurements. It implicitly use a dynamic model to implicitly propagate information temporally forward instead of explicitly propagate information backward using the inversion of the dynamic models. Moreover, the EnKS only uses the 1st and the 2nd moments and ignores higher moments in the same manner as the EnKF. Regardless, the smoother ensemble always has smaller ensemble spread than the filter ensemble.

Another important characteristic of the forecast generated from the RCR model is that the ensemble mean is usually smooth and has very low intensity. This is also observed in the updated ensemble mean when there is not sufficient measurement information to update the ensemble. This characteristic is closely related to the GOES cloud-top temperature forcing variable used to define the new rainfall region. The ensemble mean is usually too smooth to represent rainfall features because of the variation in rainfall position in each ensemble member. Therefore, it may be more realistic to use an ensemble member instead of the mean to represent the rainfall field.

4.3 Model Parameter and Measurement Error Estimation

The two key ingredients in a sequential data assimilation scheme are a recursive dynamic model for propagating the state of interest, and an analysis algorithm for incorporating new measurement information to the state. Since we have both of these key elements, the RCR model and the EnKS algorithm, we should theoretically be ready to merge multiple rainfall measurements and provide compressive rainfall ensemble for the USGP project. Unfortunately, both the rainfall model and the assimilation techniques contain some unknown parameters whose values can greatly alter the accuracy and reliability of the results. These parameters must be estimated prior to performing the rainfall data assimilation algorithm.

The unknown parameters in our rainfall data assimilation algorithm can be categorized into two groups. They are the model parameters and the measurement error statistics. The model parameters are required by the RCR model to propagate rainfall ensemble forward through time and to produce a reasonable forecast ensemble, while the measurement errors statistics are required to weight the uncertainty in the analysis algorithm. The following section will provide the parameter estimation technique to estimate model parameters in the RCR model; we will also summarize the measurement error statistics obtained in Chapter 2 for the USGP rainfall measurements. After all the parameters and statistics have been obtained, we can then apply the rainfall data assimilation algorithm and calculate the comprehensive rainfall ensemble in the USGP case study.

4.3.1 Model Parameter Estimation by State-Augmentation

Many of the hydrologic dynamic models, including our RCR model, conceptualize complex characteristics and behaviors of variables-of-interest with simple parameterizations. Generally, the associated parameters cannot be directly or easily measured and they must be inferred by indirect methods. There are varieties of parameter estimation techniques commonly used. Popular examples include manual and automatic model calibration techniques with historical data [13, 32], direct perturbation techniques [3, 69], adjoint methods [34, 91, 122, 123], and the state augmentation with an ensemble analysis algorithm [4, 40, 68, 89].

For the USGP case study, we choose to use the state-augmentation algorithm. State augmentation is a fast and straight-forward technique. It can be easily applied with the EnKF and EnKS update algorithm. The main idea of the state-augmentation method is that the model parameters are considered parts of the model state. They can be updated alongside the state by including these parameters in the state vector and using an ensemble analysis scheme. This technique allows model parameters to be time-variant and updated in real-time in the similar manner as the state. The state-augmentation technique with the ensemble Kalman filter-based algorithm has been proven to work successfully in many hydrological and Earth climate systems studies [4, 68, 89].

Some studies suggest that a combination of the state-augmentation technique and the particle filter is better for estimating the model parameters [68, 120]. Unlike the ensemble Kalman filter-based algorithm, where the higher moments are neglected, the particle filter uses the full probability density function from the ensemble in the analysis step. Hence, it gives more accurate and reliable result, especially when the update variable is non-linearly related to the measurement data. This is exactly the case for estimating the parameters by the state-augmentation technique. However, the particle filter algorithm is usually much more computationally expensive. It can be impractical to employ for a high dimensional system such as our rainfall problem.

It is important to note that when the parameters are time-invariant, the estimated parameters from the state-augmentation technique using the fixed-interval EnKS will always equal to the parameters at the final time step obtained from the EnKF. This can be easily verified from (4.52) with the parameter ensemble matrix persisting in time. However, to prevent the ensemble from collapsing, independent and identically distributed random noise is usually added to the parameter ensemble before we propagate the state ensemble. Therefore, the parameters may vary from one time to another, and the parameter estimated from the EnKS will no longer be a constant.

To ensure that the state-augmentation can be used to estimate our RCR model parameters, we first test the technique with synthetic data where the parameters are known. The synthetic rainfall are taken from the experiment in Section 4.2.4 with parameters $\beta=0.08\ pixel^{-2}\ hr^{-1}$, $\alpha = 0.6\ hr^{-1}$, $\sigma = 1.75\ pixel$, $E[i_0] = 5\ mm/hr$, $\nu = 25\ cells/cluster$, and $\rho=2$. Instead of using the scattered fine-scaled and coarse-scaled measurement, we use full fine-scale measurement with low measurement noise, e.g. $\sigma_v^2=0.1$, and make the data available at every time step. The ensemble size is increased to 2000. The state and parameters are rescaled so that they have approximately the same order of magnitude. After each update time, the ensemble spread decreases. Because the parameters are assumed to be persistent in time, there is no process noise introduced when the ensemble of parameters is propagated. Consequently, the ensemble variation of the forecast will be too small to be effected by the update. Therefore, we subjectively increase the parameter ensemble spread at each time before propagating the rainfall ensemble forward by adding independent random noises. These noises are drawn from the Gaussian distribution with zero mean and standard derivation equal to 1% of the parameter values.

An example of the RCR model parameters estimated from the state-augmentation technique is shown in Figure 4-13. In each plot, the blue line represents the ensemble mean from the EnKF algorithm, the red line represents the ensemble mean from the

EnKS algorithm, and the black constant line represents the true parameter values. The two values in the parentheses are the temporal average from the EnKS results and the true parameter value, respectively. In this example, all six parameters in the RCR model are treated as unknown and are augmented to the state ensemble when performing the EnKF and EnKS analysis. The initial condition of each parameter is assumed to follow a Gaussian distribution with the mean value shown at time t_0 and the standard derivation equal to ¼ of the mean value. At any time step, if a parameter becomes less than zero, it will be resampled from the initial condition. This minimum boundary is crucial to prevent the RCR model from becoming unstable.

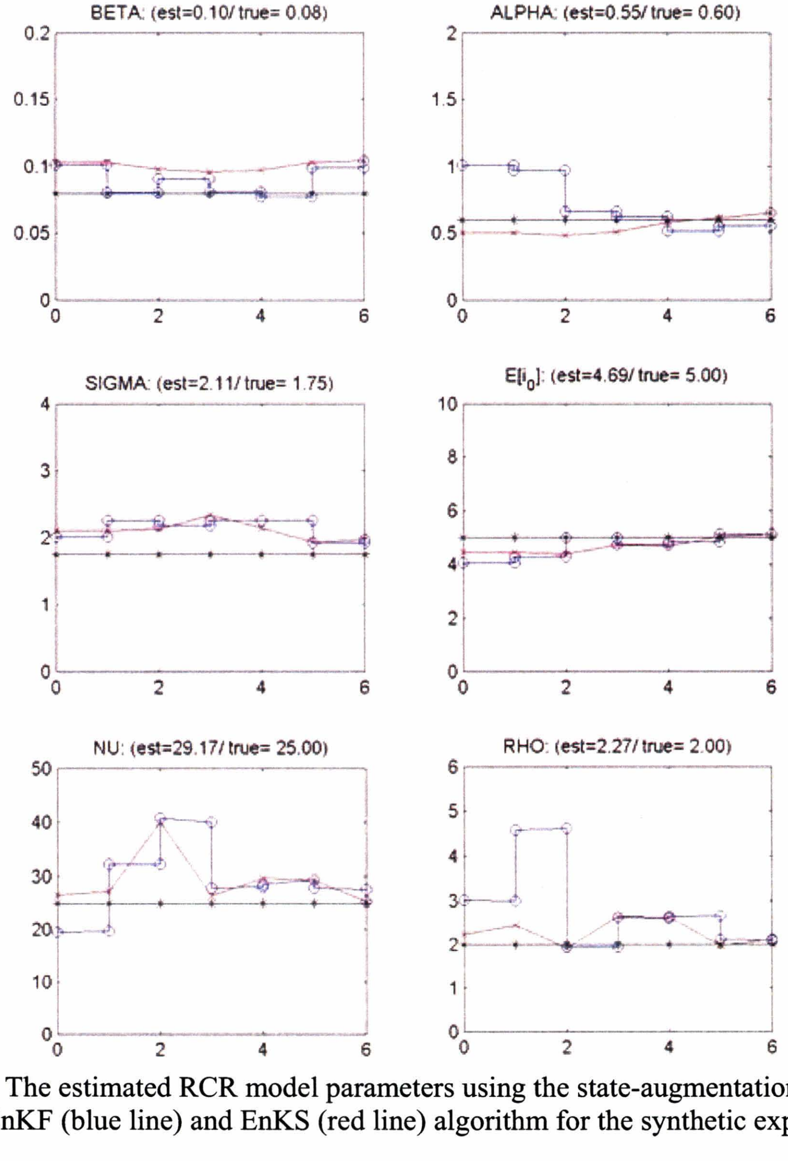


Figure 4-13: The estimated RCR model parameters using the state-augmentation technique with the EnKF (blue line) and EnKS (red line) algorithm for the synthetic experiment

The results in Figure 4-13 show that the state-augmentation technique can be used to estimate all unknown parameters in the RCR model. The state-augmentation technique requires a large ensemble size to estimate the RCR model parameters with the EnKF and the EnKS algorithm. In our experiment, we can obtain reliable results if we use 2000 or more ensemble members. Moreover, we obtain more accurate results when we update with good quality measurements (e.g., comprehensive over space and time and with small error variance). In addition, the estimated parameters from the state-augmentation usually cluster around the initial condition values. To obtain a reasonable initial condition, we can iteratively perform the state-augmentation technique many times, each time using results from the previous iteration as the initial condition. Finally, the state-augmentation is significantly more reliable if we fix some parameters to the true value.

We have attempted to estimate these parameters by various methods including the manual calibration, direct perturbation, and minimization of many different objective functions. However, none of these algorithms provides more credible results than the state augmentation technique. In estimating the RCR model parameters, the state-augmentation technique performs relatively well when the parameter values are well constrained and the ensemble size is large. Furthermore, it is significantly faster and more convenient to utilize than the other approaches tested. Thus, we propose to use the state-augmentation technique to estimate the RCR model parameters for the USGP case study.

4.3.2 Estimated RCR Parameters for the USGP Case Study

In this section, we apply the state-augmentation technique to estimate the RCR model parameters for the USGP project. To make the algorithm stable and reliable, it is necessary to generate a large number of ensemble replicates and use good quality rainfall measurements. These criteria cannot be met in the USGP project. First, there are roughly 500×500 pixel^2 over the USGP region at resolution of 0.05°. It is impractical to use a large ensemble size. Second, the rainfall measurements to be used in the USGP rainfall data assimilation case study, ASOS, SSM/I, TRMM and AMSU, are too sparse in space or time. It may not be possible to estimate the RCR model parameters using the state-augmentation technique in real-time. Therefore, we propose estimating the RCR parameters off-line. We will estimate the RCR model parameters over many different rainfall events over a small region. We can then analyze the estimated parameters. If the estimated parameters agree relatively well, we can use the average parameters to perform rainfall data assimilation in the USGP case study.

We first choose several storm events for estimating the RCR model parameters. Each storm candidate should be contained in a small region, e.g. 2.5°×2.5°, in order to employ a large ensemble size. It should consist of regions with very low cloud-top temperatures, e.g. GOES < T_B, so that the RCR model can generate new rain cells. Finally, it should last long enough for the state-augmentation to converge. The storm events selected for estimating the model parameters for the USGP project are summarized in Table 4.1.

Table 4.1: Storm events chosen to estimate the RCR model parameters

No.	Start Date	Start Time	Stop Date	Stop Time	Latitude (°N)	Longitude (°W)
1.	2004/06/01	22:00 GTM	2004/06/02	04:00 GTM	32.5 / 35.0	99.0 / 96.5
2.	2004/06/06	22:00 GTM	2004/06/07	08:00 GTM	32.5 / 35.0	99.0 / 96.5
3.	2004/06/12	05:00 GTM	2004/06/12	11:00 GTM	37.0 / 39.5	96.0 / 93.5
4.	2004/07/01	00:00 GTM	2004/07/01	06:00 GTM	34.0 / 36.0	101 / 98.0
5.	2004/07/15	04:00 GTM	2004/07/15	11:00 GTM	33.5 / 36.0	96.0 / 93.5
6.	2004/07/23	20:00 GTM	2004/07/24	04:00 GTM	35.0 / 37.5	98.0 / 95.5
7.	2004/08/06	22:00 GTM	2004/08/07	04:00 GTM	33.0 / 35.5	104 / 101.5
8.	2004/08/09	20:00 GTM	2004/08/10	03:00 GTM	38.0 / 40.5	98.5 / 96.0
9.	2004/08/21	03:00 GTM	2004/08/21	08:00 GTM	30.0 / 32.5	97.5 / 95.0

For each storm event in Table 4.1, we use the state-augmentation technique with EnKS to estimate the RCR model parameters. The ensemble size of 1000 is used and we repeat the experiment several times to ensure that the results are relatively reliable. From our experience, the state-augmentation technique does not work well when applied with real measurement observations. We believe that this is because the RCR model cannot truly represent the real physics of the rainfall process, especially intermittency. However, we found out that if we slightly smooth the NOWRAD data and provide some constraints on the parameters, the state-augmentation algorithm will provide more consistent results.

It is possible to constrain our parameters because they have physical meaning, but the estimated results can be subjective. We choose to minimize the bias by constraining only a few parameters that have the most direct physical meaning. Hence, we fix the parameter σ, which represents the size of rain cells, to have a value of 1.50 L, where L equals to the pixel length of 0.05°. This corresponds to approximately 7.5 km when the cell is first born. This value is chosen so that the model can reproduce small features of a storm. In addition, we strictly bound the parameter α to be less than 1.5 hr^{-1} to prevent existing rainfall from dissipating too quickly. Figure 4-13 provides an example of the parameter estimated from the storm event case 1 of Table 4.1 when all the restrictions applied.

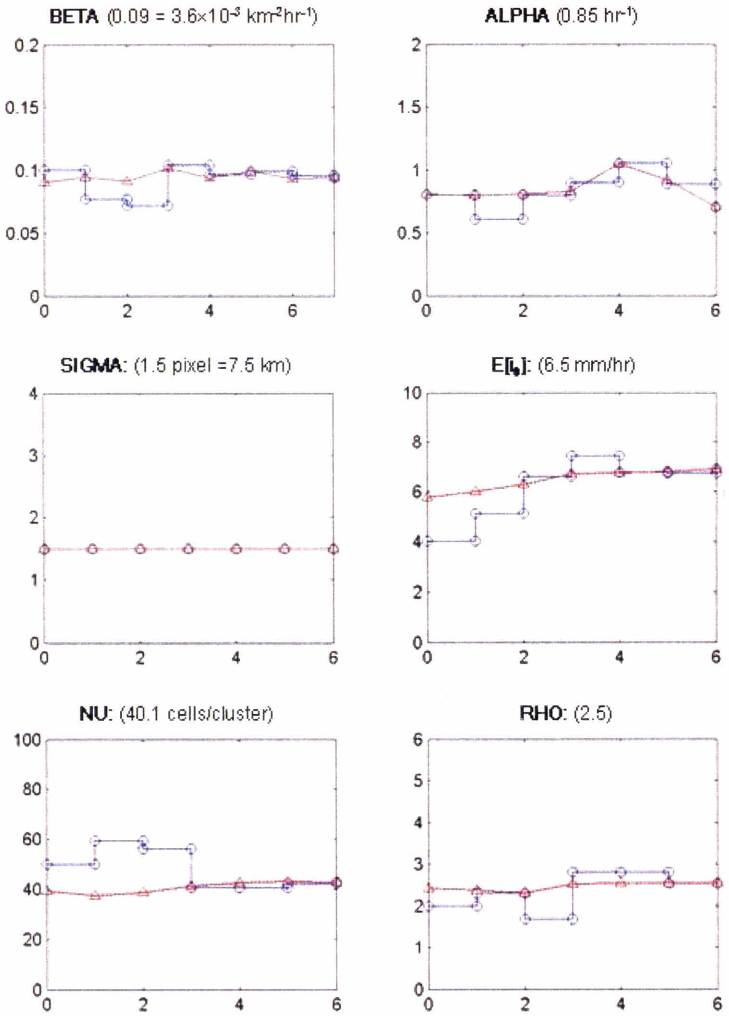


Figure 4-14: The estimated RCR model parameters from the storm event #1 in Table 4.1

We then repeat the experiment many times over each storm event. The initial conditions and the maximum and minimum bounds for the RCR parameters are given in Table 4.2. The average RCR model parameters estimated from the state augmentation technique using the EnKS algorithm are given in Table 4.3. The resulting mean parameters are then chosen for the rainfall data assimilation implementation during June 1st to August 31st, 2004 over the USGP region. Note that these parameter values are to be used as guidelines. Altering a parameter value by a reasonable amount will not drastically impact the overall assimilation result.

Table 4.2: Constraints on the RCR model parameters for the USGP project

Parameter	**Initial Values**	**Max. Bound**	**Min. Bound**
β *(km^{-2} hr^{-1})*	5.0×10^{-3}	5.0×10^{-3}	1.0×10^{-3}
α *(hr^{-1})*	0.8	1.2	0.4
σ *(km)*	7.5	7.5	7.5
$E[i_0]$ *(mm/hr)*	5	10	1
N *(cells/cluster)*	50	100	10
ρ	2	4	1

Table 4.3: The average RCR model parameters from the state-augmentation using the EnKS over the storm events listed in Table 4.1

Event No.	**β** (km^{-2} · hr^{-1})	**α** (hr^{-1})	**σ** (km)	**E[i$_0$]** (mm/hr)	**ν** (cell/cluster)	**ρ**
1.	3.6×10^{-3}	0.9	7.5	6.5	41	2.5
2.	3.2×10^{-3}	0.6	7.5	5.1	45	2.8
3.	4.2×10^{-3}	1.0	7.5	4.8	32	2.7
4.	3.5×10^{-3}	0.7	7.5	5.8	39	2.4
5.	3.4×10^{-3}	0.6	7.5	6.7	40	2.0
6.	3.0×10^{-3}	0.5	7.5	7.1	52	2.1
7.	3.3×10^{-3}	0.5	7.5	6.8	42	2.4
8.	3.6×10^{-3}	0.7	7.5	6.1	35	2.6
9.	3.1×10^{-3}	0.6	7.5	7.0	42	2.8
Ave.	**3.4×10^{-3}**	**0.7**	**7.5**	**6.2**	**41**	**2.5**

4.3.3 Generating Position and Intensity Perturbed Measurements

In order to use the EnKF and the EnKS, we must obtain the perturbed measurement matrix $\tilde{\mathbf{Y}}$ in (4.13) and the measurement perturbation matrix $\mathbf{Y}$ in (4.14). These matrices are essential for calculating the update matrix $\mathbf{K}_4$ in equation (4.45) and updating the forecast ensemble. We first generate the perturbed measurement matrix $\mathbf{Y}$ by taking each real measurement images and shifting them with a random displacement (q_x,q_y). This random displacement is drawn from the Gaussian distribution with position error statistics from Table 2.2. We recommend using one random displacement to shift a whole image. If we choose to randomly and independently shift each pixel, the measurement realization may become too scattered. We repeat the process and generate position-perturbed ensemble.

Next, we perturb the rainfall intensity of the ensemble by adding an independent random noise to each pixel. This random noise will be drawn from the Gaussian distribution with zero mean and the standard derivation given by equation (2.2). The constants c_1 and c_2 in equation (2.2) are also given in Table 2.2. At this point, we have obtained the perturbed measurement matrix $\mathbf{Y}$. The measurement perturbation matrix $\tilde{\mathbf{Y}}$ is easily calculated by subtracting the matrix $\mathbf{Y}$ from its mean, as given in equation (4.14). With the measurement perturbation matrix generated in this form, the theoretical covariance matrix $\mathbf{R}$ will no longer be diagonal.

4.4 Implementation of the Rainfall Data Assimilation on the USGP Case Study

At this point, we have obtained all the key components needed for the data assimilation algorithm, and we are ready to implement it on the USGP project. These three components are (1) the RCR rainfall model with GOES input and known parameters, (2) rainfall measurements from ASOS, TRMM, SSM/I and AMSU-B with known error statistics, and (3) the analysis scheme, which consists of the EnKF and the EnKS. The USGP project consists of roughly a 500×500 pixel^2 domain. Because of limitations in CPU and memory, the maximum ensemble size allowed is approximately equal to 200. Despite the small ensemble size used, the analysis ensemble calculated from the EnKF and the EnKS algorithm is usually stable. Samples of the smoother ensemble from the USGP project are given in Figures 4-15. In the figure, each row represents a 3-hour time step beginning on July 24th, 2004 at 19:00 GMT. The first column represents NOWRAD rainrate (e.g., validation data), the middle column represents the ensemble mean from the EnKS algorithm, and the right column represents the different between the first two columns. The smoother ensembles are obtained from the fixed-lag EnKS algorithm with lag $\Delta = 1.5$ hours.

The ensemble results in Figures 4-15 provide reasonable rainfall fields over a coarser scale that match the measurement data from NOWRAD. The results have correlation in space and time. However, it cannot provide fine scale features of rainfall, nor can it be used to replace NORAD rainfall product. During this time interval, the data assimilation merges scattered ASOS rain gauge measurements at every hour and infrequent satellites images as shown in Figures 4-16. In addition, the RCR model only includes the GOES cloud-top temperature data at every hour and approximate dynamics of rainfall using relatively simple equations. Therefore, it is not fair to expect this data assimilation procedure to match fine scale rainfall of the NOWRAD data. The approach should provide better results at coarser scales where position errors are less important.

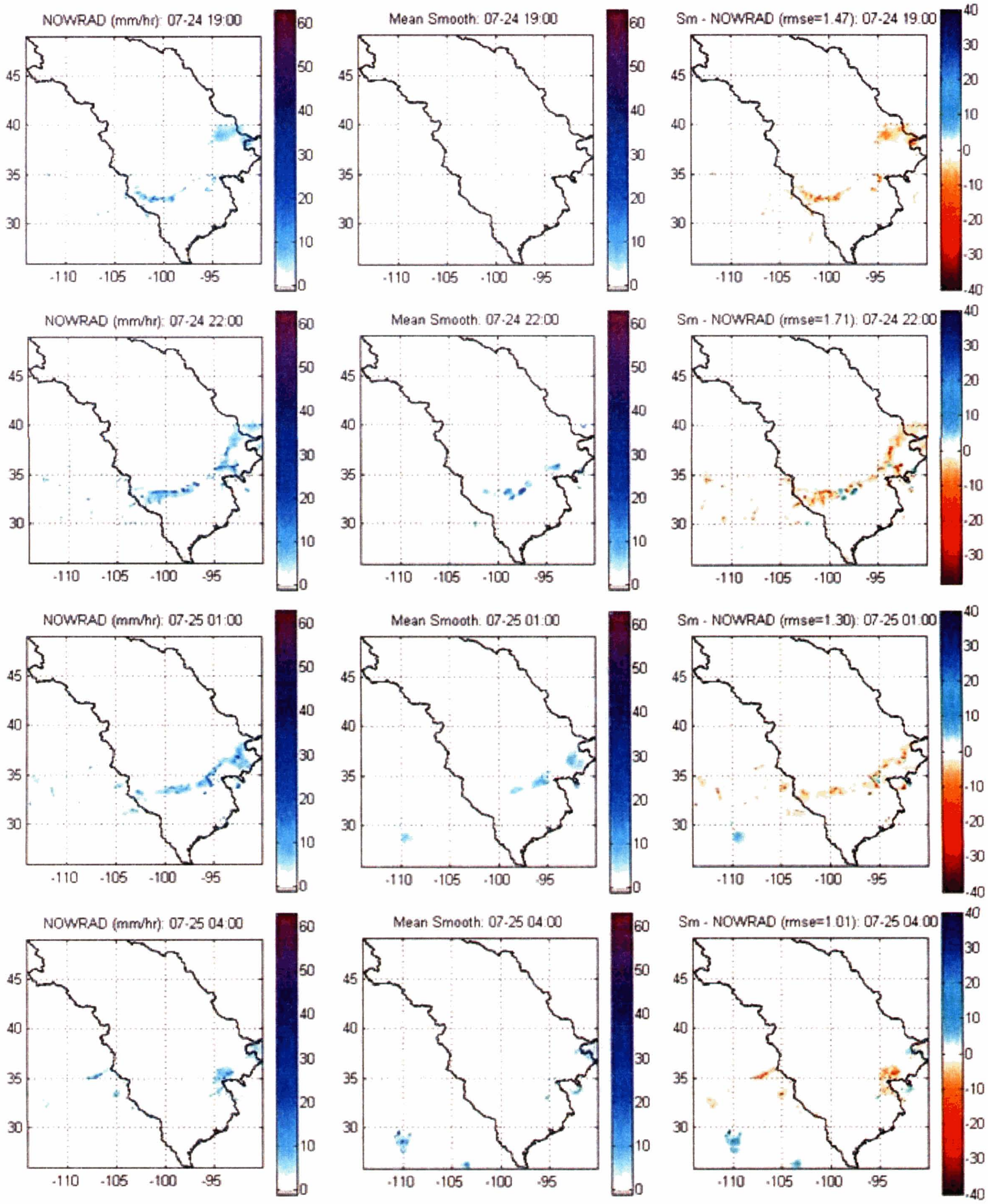

NOWRAD (mm/hr): 07-24 19:00
Mean Smooth: 07-24 19:00
Sm - NOWRAD (rmse=1.47): 07-24 19:00
NOWRAD (mm/hr): 07-24 22:00
Mean Smooth: 07-24 22:00
Sm - NOWRAD (rmse=1.71): 07-24 22:00
NOWRAD (mm/hr): 07-25 01:00
Mean Smooth: 07-25 01:00
Sm - NOWRAD (rmse=1.30): 07-25 01:00
NOWRAD (mm/hr): 07-25 04:00
Mean Smooth: 07-25 04:00
Sm - NOWRAD (rmse=1.01): 07-25 04:00
45
40
35
30
-110
-105
-100
-95
0
10
20
30
40
50
60
-10
-20
-30
-40

Figure 4-15: The NOWRAD rainrate (left), the smoother ensemble mean (middle), and the differences of rainrate over the USGP region at every 3 hours during July 24th, 2004 19:00 GMT to July 25th, 2004 4:00 GMT

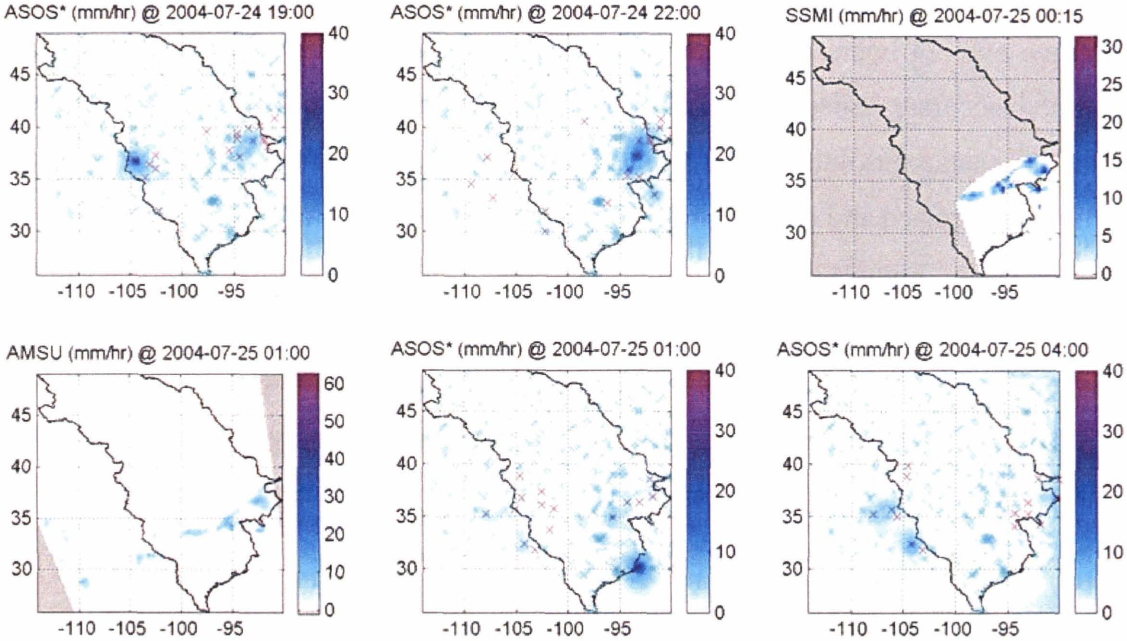

ASOS* (mm/hr) @ 2004-07-24 19:00
ASOS* (mm/hr) @ 2004-07-24 22:00
SSMI (mm/hr) @ 2004-07-25 00:15
AMSU (mm/hr) @ 2004-07-25 01:00
ASOS* (mm/hr) @ 2004-07-25 01:00
ASOS* (mm/hr) @ 2004-07-25 04:00
45
40
35
30
-110 -105 -100 -95
0
10
20
30
40
5
15
25
50
60

Figure 4-15: Samples of the USGP raw data available during July 24th, 2004 19:00 GMT to July 25th, 2004 4:00 GMT

There are some aspects about the rainfall data assimilation project that should be emphasized. First, the RCR model cannot perform well where there is a large area of deep clouds (i.e., large region where GOES < T_B). In this case, the rainfall ensemble generated with the RCR model is usually too scattered (e.g., Figure 4-17). The rainfall ensemble will not look realistic unless there are good quality measurements to update the ensemble. On the contrary, the RCR model will not generate new rain cells over the area with GOES cloud-top temperatures higher than T_G. Hence, it usually underestimates low intensity rainfall (e.g., stratiform rainfall with low rainfall intensity and high cloud-top temperature). This problem has to do with the physics of the rainfall model and can be minimized to improve the RCR model.

Secondly, the root mean square errors (RMSE) of the open-loop result directly from the RCR model, the forecast (i.e. state values that account for past information but not the current), the filter from the EnKF, and the fixed-lag smoother from the EnKS are illustrated in Figures 4-17. These RMSE's are evaluated with regard to NOWRAD rainrate data after aggregating to the 0.25 degree resolution and performing the moving average over 24 hours. Note that the RMSE is not a perfect performance indicator. It can be inaccurate, especially when there are significant position errors. By aggregating data to a coarser resolution (e.g., from 0.05 degree to 0.25 degree resolution), the position error is minimized and the RMSE provide a more reliable indicator of the accuracy at the coarser scale Figures 4-17 shows the open-loop result is always the most inaccurate. In general, the smoother ensemble provides more accurate results than the filter and the forecast. All RMSE's decrease during the dry down period, and increase when there is significant rainfall. The filter and smoother perform significantly better than the forecast when there are significant amounts of rainfall over a large region and there are good quality measurements available. The smoother can be worse than the filter or the forecast results, especially if the RMSE is evaluated at finer resolution. The measurements also play an important role in the accuracy of the EnKF and the EnKS. In general, during the period when there are many satellite-based measurements, the accuracy of the EnKF and EnKS improve significantly.

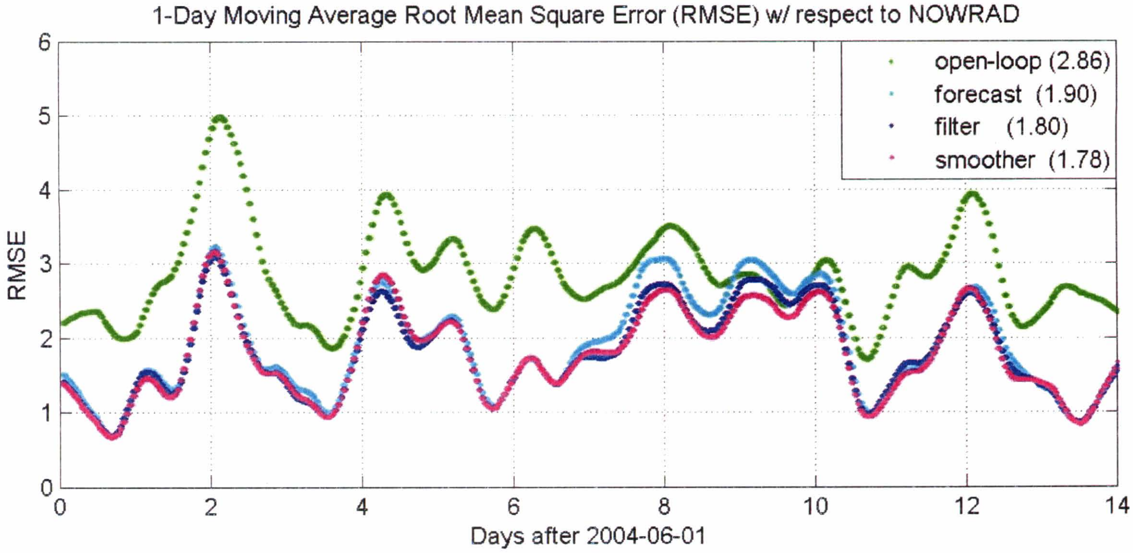


Figure 4-17: The 1-day moving average RMSE of the open-loop, forecast, filter and smoother mean ensemble with regard to NOWRAD measurement over a 0.25 degree resolution

Finally, Figures 4-18 show examples of the filter results and the smoother results, and emphasize the ensemble spreading (i.e., standard derivation of the ensemble.) These properties are consistent with the synthetic experiment described in Section 4.3.4. In Figures 4-18, the left column (a,c,e) shows the update result and the right column shows the smoothing result. The first row represents the ensemble mean from the USGP project on June 1st, 2006 at 00:15 GMT. The second and the third row provide the associated RMSE and the ensemble spread, respectively. Lastly, the value in the parentheses on the title is the spatial average of the RMSE and the ensemble spread over the USGP region. Unlike the RMSE, the ensemble spread of the smoother is always significantly less than that of the update. Therefore, if the variation in the ensemble is of interest or the rainfall ensemble is to be used as input in a land-surface model, the filtering result (e.g., EnKF result) will provide more suitable ensemble than the smoothing result (e.g., EnKS result.)

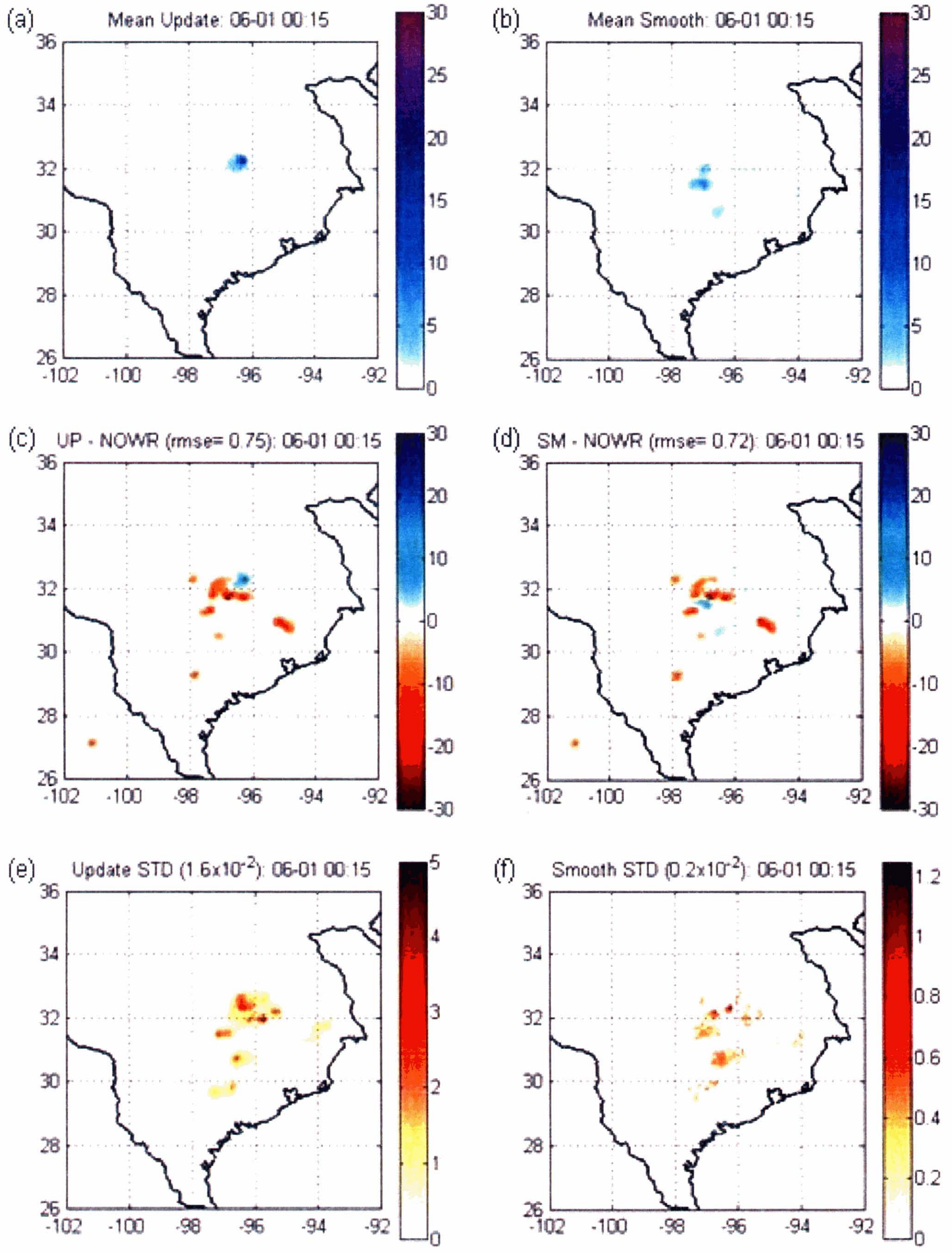


Figure 4-17: Effects of the EnKS smooth ensemble over the EnKF update ensemble

4.5 Conclusions

This chapter focused on the ensemble data assimilation for merging multiple sources of rainfall measurement and producing a comprehensive rainfall ensemble. We first introduced the Ensemble Kalman filter (EnKF), its derivation and its formulation. The EnKF uses a dynamic model to propagate a state ensemble through time, and it uses the ensemble covariance to incorporate measurement data and update the forecast ensemble. The algorithm is efficient and proven to work well in many hydrological applications. However, when the ensemble size is small relative to the state or measurement dimension, the EnKF can become unstable.

We substituted the matrix inversion algorithm in the EnKF with the stable pseudo-inversion technique. The new EnKF is proven to be more stable than the original one, especially for a small ensemble size. Moreover, we provided the SVD sampling strategy in the Appendix C for minimizing sampling error. Next, we accounted for future measurements by using the Ensemble Kalman smoother (EnKS) algorithm. The EnKS updates the current state with measurements from times after the estimation time using the correlation matrix. The algorithm is easy to implement and requires minimal computational cost if the EnKF has already been performed. Both the EnKF and EnKS algorithms were implemented on synthetic experiments and the both algorithms showed good estimation results. The EnKS algorithm provides superior results when the time of interest and later measurement times are close.

After choosing the appropriate data assimilation technique, we estimated the parameters of the RCR rainfall model by using the state-augmentation technique. This technique treats the parameters as random variables and updates them in real-time along with the state ensemble. The synthetic experiments illustrated that the state-augmentation could provide good estimate of the model parameters when the conditions were right. These conditions are (1) the ensemble size must be relatively large, (2) the measurement data should be sufficient in space and time and be of good quality, and (3) reasonable initial conditions must be given. We were unable to estimate the model parameters in real time because of computation limitations. Therefore, we estimated parameters off-line using different storm events and used the results as guidelines to create time-invariant parameters for the USGP project. Next, we created measurement perturbations from the position error statistics and intensity error statistic given in Chapter 2.

Finally, we implemented the ensemble data assimilation on the USGP project using the EnKF and EnKS with the RCR model. We were able to efficiently merge multiple sources of rainfall measurements and provide a reasonable comprehensive rainfall ensemble. However, the results can only be comparable to NOWRAD measurement at coarser resolutions. This study's approach cannot capture fine scale features and characteristics of rainfall. The result from merging multiple scattered measurements with a simple stochastic based model cannot replace NOWRAD data. We should not expect this data assimilation procedure to match fine scale rainfall data or to compete with complex meteorological modeling. The algorithm, however, may be used for coarser scale analyses. It is fast, simple, and easy to implement. In addition, it can provide an ensemble of a rainfall field that captures correlations in space and time.

Chapter 5

Parameter Estimation for the Multiplicative Cascade Rainfall Model Using the EM-SRE Algorithm

5.1 Introduction

This chapter provides method for estimating the scaling parameters of the multiplicative cascade rainfall model and its associated measurement error variance. These unknown parameters are estimated using the Expectation-Maximization technique on the Scale-Recursive Estimation (EM-SRE) framework. The multiplicative cascade model [54, 94, 97, 98, 112, 114, 116] is well known to provide reasonable spatial characteristics of rainfall. In addition, we can use this algorithm to merge static rainfall data from multiple measurement sources. This chapter emphasizes two major concerns when applying the multiplicative cascade rainfall model: (1) rainfall intermittency and (2) identifiability and uniqueness of the scaling parameters and measurement variances. The tree pruning technique is proposed to solve the rainfall intermittency issue, while the expectation-maximization (EM) algorithm is applied to estimate all unknown parameters in the SRE algorithm.

The organization of this chapter is as follows. In the next section, we introduce the scale-recursive representation and outlines of the SRE framework in a general formulation.

Section 5.3 presents the multiplicative cascade rainfall model in the form that can fit the SRE algorithm. We also point out some special properties and limitations of this rainfall representation. In Section 5.4, the tree pruning technique is proposed for dealing with rainfall intermittency. Next, Section 5.5 discusses the expectation-maximization (EM) algorithm for estimating the SRE parameters. We focus on identifiability and uniqueness of the SRE parameters, as well as test the parameter sensitivities to changes in SRE structure. In Section 5.6, we apply the combined expectation maximization and the scale-recursive estimation (EM-SRE) algorithm to estimate rainfall intensity, scaling parameters and a measurement variance of the NOWRAD rainfall observation. Finally, we discuss the SRE algorithm and conclude the chapter in Section 5.7.

5.2 A Scale-Recursive Estimation Algorithm

5.2.1 The Scale-Recursive Representation

The Scale Recursive Estimation (SRE) algorithm represents the state of interest by a pyramidal-like grid (2-D) or by an inverse tree structure (1-D) [21, 76], as illustrated in Figure 5-1. For convenience, the word "tree" will be used to represent both the 1-D and the 2-D structure. The top node on the tree, called the "root node", embodies total areas of interest. The bottom nodes, referred to as a "leaf nodes", represent a random variables at the finest scale where the finest observations are available. Nodes in the middle scales usually correspond to the coarser observations.

Let $m(s)$ be the level of node s on an M-level tree, whose root node has $m(s) = 0$ and the leaf nodes have $m(s) = M$. Thus, there are a total of $M+1$ scales on an M-level tree. For a node s at a level $m(s)$, we denote its parent at level $m(s)-1$ by $s\gamma$, and its q children at level $m(s)+1$ by $s\alpha_i$; $i=1...q$. The value q is called the branch number and usually is a constant across all scales on the tree. Hence, a number of nodes on each level is equal to

$q^{m(s)}$. Furthermore, let *x(s)* and *z(s)* denote a state vector and a measurement vector at node *s*, respectively.

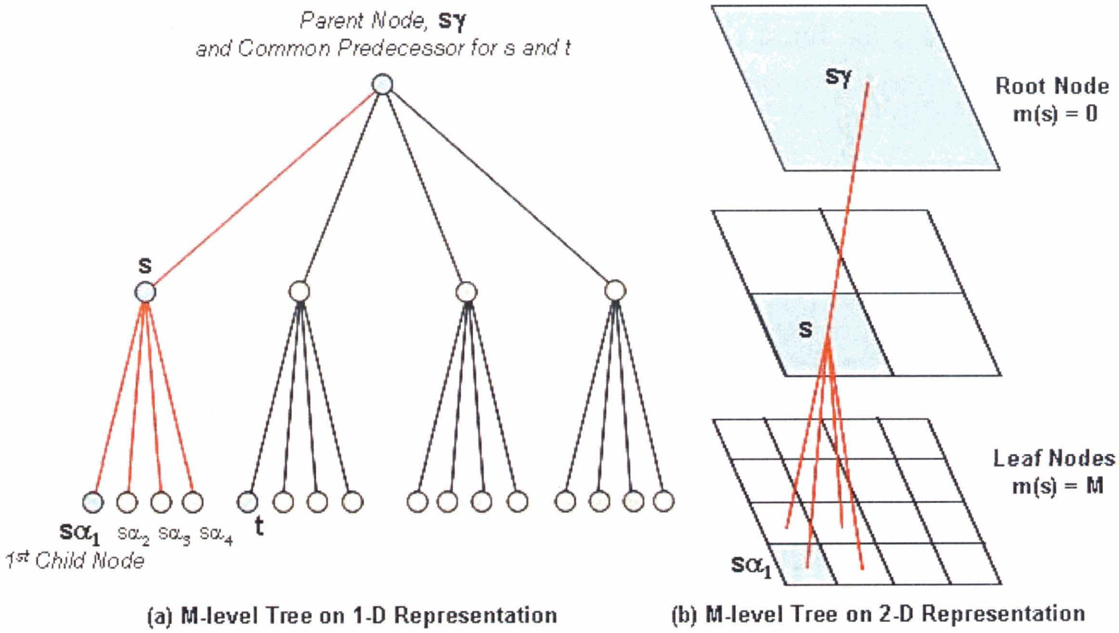

Parent Node, sγ
and Common Predecessor for s and t
s
sα1 sα2 sα3 sα4 t
1st Child Node
(a) M-level Tree on 1-D Representation
sγ
Root Node
m(s) = 0
s
Leaf Nodes
m(s) = M
sα1
(b) M-level Tree on 2-D Representation

Figure 5-1: Examples of scale-recursive structures with (a) a one-dimensional inverse-tree, and (b) a two-dimensional pyramidal-like grid

The scale-recursive dynamic model propagates state information from node *sγ* to *s,* and the measurement equation linearly relates an observation to a state at node *s* via,

$$x(s) = F(s)x(s\gamma) + w(s) \tag{5.1}$$

$$y(s) = H(s)x(s) + v(s) \tag{5.2}$$

where *w(s)* is a process noise, and *v(s)* is a measurement noise. Both *w(s)* and *v(s)* are independent, zero-mean white noise processes with covariance matric*es Q(s)* and *R(s)*, respectively. The term *F(s)x(sγ)* denotes a coarse-to-fine scale prediction with *w(s)* representing higher resolution details. From (5.1), a prior covariance matrix at node *s* and a prior cross-covariance matrix between node *s* and *t* are given by

$$P_0(s) = F(s)P_0(s\gamma)F^T(s) + Q(s) \tag{5.3}$$

$$C_0(s,t) = \phi_0(s, s\wedge t)P_0(s\wedge t)\phi_0^T(t, s\wedge t) \tag{5.4}$$

where node $s{\wedge}t$ is the finest common predecessor of the node s and the node t. ϕ_0 is given by:

$$\phi_0(s_1, s_2) = \begin{cases} I & , s_1 = s_2 \\ F(s_1)\phi_0(s_1\gamma, s_2) & , m(s_1) > m(s_2) \end{cases} \tag{5.5}$$

5.2.2 The Two-Sweep SRE Algorithm

Given the scale-recursive model equation (5.1) and (5.2), the SRE algorithm optimally estimates the state conditioned on all measurements on the tree. The SRE algorithm consists of an upward fine-to-coarse filtering sweep followed by a downward coarse-to-fine smoothing sweep. The upward and downward sweeps are analogous to the temporal Kalman filter algorithm and the Rauch-Tung-Striebel (RTS) smoothing algorithm [43], respectively. However, the SRE algorithm propagates information through scale, instead of time.

The SRE algorithm begins at the leaf nodes, and the information is first propagating up toward the root node. In order to propagate information upward in the recursive fashion, equation (5.1) is modified to be:

$$x(s\gamma) = \tilde{F}(s)x(s) + \tilde{w}(s) \tag{5.6}$$

$$\tilde{F}(s) = P_0(s\gamma)F^T(s)P_0^{-1}(s) \tag{5.7}$$

with,

$$\tilde{Q}(s) \equiv \mathrm{E}\left[\tilde{w}(s)\tilde{w}^T(s)\right] = P_0(s\gamma) - \tilde{F}(s)P_0^{-1}(s)\tilde{F}^T(s) \tag{5.8}$$

where $\tilde{F}(s)$ represents the upward propagation function, and $\tilde{w}(s)$ is the upward process noise. Note that $\tilde{w}(s)$ is a zero mean white noise with a covariance $\tilde{Q}(s)$ and uncorrelated with $x(s)$. In addition, we assume that the prior covariance matrix at node *s*, denoted by $P_0(s)$, is known. Thus, $\tilde{F}(s)$ and $\tilde{w}(s)$ can be specified before applying the SRE algorithm.

At each scale, the upward fine-to-coarse filtering sweep consists of three steps: (1) the update step, (2) the prediction step, and (3) the merging step. We denote the set of all measurements at node *s* and all nodes under *s* as Y_s and the set of all measurements strictly under node *s* as Y_s^+. The update step incorporates the measurement *y(s)* at node *s* to provide an update estimate $\hat{x}(s \mid Y_s)$ and its corresponding update covariance $P(s \mid Y_s)$ from

$$\hat{x}(s \mid Y_s) = \hat{x}(s \mid Y_s^+) + K(s)\left[y(s) - H(s)\hat{x}(s \mid Y_s^+)\right] \tag{5.9}$$

$$P(s \mid Y_s) = \left[I - K(s)H(s)\right]P(s \mid Y_s^+) \tag{5.10}$$

where the Kalman gain matrix *K(s)* is given by,

$$K(s) = P(s \mid Y_s^+)H^T(s)\left[H(s)P(s \mid Y_s^+)H^T(s) + R(s)\right]^{-1} \tag{5.11}$$

Next, the prediction step propagates the update estimate one-scale level upward using (5.6). Suppose that we obtain all updated states $\hat{x}(s\alpha_i \mid Y_{s\alpha_i})$ and covariance matrices $P(s\alpha_i \mid Y_{s\alpha_i})$ at all children nodes below node *s*; and the fine-to-coarse predicted estimates are given by:

$$\hat{x}\left(s \mid Y_{s\alpha_i}\right) = F\left(s\alpha_i\right)\hat{x}\left(s\alpha_i \mid Y_{s\alpha_i}\right) \tag{5.12}$$

$$P\left(s \mid Y_{s\alpha_i}\right) = F\left(s\alpha_i\right)P\left(s\alpha_i \mid Y_{s\alpha_i}\right)F^T\left(s\alpha_i\right) + \tilde{Q}\left(s\alpha_i\right) \tag{5.13}$$

These predicted estimates from all children nodes are merged to obtain one best prediction of the state at node *s* in the merging step,

$$\hat{x}\left(s \mid Y_s^+\right) = P\left(s \mid Y_s^+\right)\sum_{i=1}^{q} P^{-1}\left(s \mid Y_{s\alpha_i}\right)\hat{x}\left(s \mid Y_{s\alpha_i}\right) \tag{5.14}$$

$$P\left(s \mid Y_s^+\right) = \left[(1-q)P_0^{-1}(s) + \sum_{i=1}^{q} P^{-1}\left(s \mid Y_{s\alpha_i}\right)\right]^{-1} \tag{5.15}$$

where the branching number *q* denotes the total number of children nodes under the current node *s*. The update, prediction and merging steps are recursively repeated upward until reaching the root node. Next, the downward sweep begins. It computes smooth estimates and corresponding covariance matrices conditioned on all available measurements on the tree. Let Y_A represent all available measurements on the tree. At the root node, Y_A is equal to Y_s. Thus, the smooth estimate $\hat{x}\left(s \mid Y_A\right)$ and covariance $P\left(s \mid Y_A\right)$ are equal to $\hat{x}\left(s \mid Y_s\right)$ and $P\left(s \mid Y_s\right)$, respectively. The smooth estimates at subsequent levels are given by

$$\hat{x}\left(s \mid Y_A\right) = \hat{x}\left(s \mid Y_s\right) + J(s)\left[\hat{x}\left(s\gamma \mid Y_A\right) - \hat{x}\left(s\gamma \mid Y_s\right)\right] \tag{5.16}$$

$$P\left(s \mid Y_A\right) = P\left(s \mid Y_s\right) + J(s)\left[P\left(s\gamma \mid Y_A\right) - P\left(s\gamma \mid Y_s\right)\right]J^T(s) \tag{5.17}$$

with

$$J(s) = P\left(s \mid Y_s\right)\tilde{F}^T(s)P^{-1}\left(s\gamma \mid Y_s\right) \tag{5.18}$$

Note that $\hat{x}\left(s\gamma \mid Y_s\right)$ and $P\left(s\gamma \mid Y_s\right)$ are the predicted state and covariance before the merging step. They are obtained from (5.12) and (5.13), respectively. The coarse-to-fine smoothing sweep recursively propagates information downward until reaching the finest scale. Consequently, all final estimates are optimally conditioned on all available measurements on the tree.

5.2.3 Characteristics of SRE Framework

There are a few important characteristics of the SRE algorithm that should be emphasized. First, the parameters *F(s)* and *Q(s)* govern the prior correlations among all nodes and implicitly represent the full prior covariance matrix (i.e., covariance matrix with a full state vector from all nodes on the tree). The SRE framework tends to generate a blocky covariance matrix, which can be inconsistent with physical properties of the process. For example, look at the nodes $s\alpha_3$, $s\alpha_4$ and *t* in Figure 5-1 and assuming that *F(s) = 1.* It is easy to verify from equation (5.4) that the cross-covariance $C(s\alpha_3, s\alpha_4) = P(s)$ is greater than the cross-covariance $C(s\alpha_4, t) = P(s\gamma)$, even though the physical distance from $s\alpha_3$ to $s\alpha_4$ and the distance from $s\alpha_4$ to *t* are the same. Nevertheless, their correlations are sufficient for capturing a variety of scale-dependent effects and proven to provide reasonable estimates of the states of interest.

Second, the smooth estimates obtained from (5.16) and (5.17) are equivalent to those calculated from a one-step temporal Kalman filter with a full state vector (i.e., an augmented state from all nodes.) These estimates are the best linear least-square estimators and are optimal only if all variables in (5.1) and (5.2) are jointly Gaussian. However, the SRE algorithm never explicitly calculates the full covariance matrix. Therefore, the SRE algorithm is much more efficient and more practical for a large system. Moreover, the prior cross-covariance matrix and the posterior-covariance matrix between any two nodes after incorporating all available measurement [76] are given by

$$C(s,t \mid Y_A) = \phi(s, s \wedge t) P(s \wedge t \mid Y_A) \phi^T(t, s \wedge t) \tag{5.19}$$

$$\phi(s_1, s_2) = \begin{cases} I & , s_1 = s_2 \\ J(s_1)\phi(s_1\gamma, s_2) & , m(s_1) > m(s_2) \end{cases} \tag{5.20}$$

5.3 Multiplicative Cascade Rainfall Model

Various studies in past decades support scaling properties of spatial rainfall. These studies include multi-fractal characterizations [75, 79, 85, 117], multiplicative cascade models and clustered point processes [54, 94, 97, 98, 112, 114, 116]. These fractal or multiplicative cascade rainfall models fit nicely with the SRE framework. The model is proven to be useful for rainfall data assimilation purposes [42, 47, 70], as well as rainfall model verification purposes [99, 126].

To employ the multiplicative cascade rainfall model with the SRE framework, we commonly use the normalized rainrate instead of actual rainrate. The multiplicative cascade rainfall model states that the normalized rainrate state at node *s*, denoted by a capital *X(s)*, is log-normally distributed. It evolves from a coarser scale *m(sγ)* to the next finer scale *m(s)* by multiplying with a cascade weight, *W(s)*:

$$X(s) = X(s\gamma) \cdot W(s) \tag{5.21}$$

where *W*'s are mutually independent random variables. Assume that both *W(s)* and *X(s)* are log-normally distributed with mean of 1.0, e.g.

$$X(s) = \exp\{x(s) - \sigma_x^2(s)/2\} \tag{5.22}$$

$$W(s) = \exp\{w(s) - \sigma_w^2(s)/2\} \tag{5.23}$$

By taking the log of (5.21), we obtain the additive form of the dynamic model in equation (5.1) with all *F(s)* equals to 1.0. Similarly, we assume that the rainrate observation *Y(s)* is related to the state *X(s)* by

$$Y(s) = X(s) \cdot V(s) \tag{5.24}$$

where *V(s)* is a log-normal measurement noise and uncorrelated with the state, e.g.

$$V(s) = \exp\left\{v(s) - \sigma_v^2(s)/2\right\} \tag{5.25}$$

By taking the log of (5.24), we arrive at the measurement equation (5.2) with *H(s)* equal to 1.0. We refer to *x(s)* and *y(s)* as a log-rainfall state and a log-rainfall measurement at node *s*, respectively. The log-transformed state-space equations for the multiplicative cascade rainfall model are given by:

$$x(s) = x(s\gamma) + w(s) \tag{5.26}$$

$$y(s) = x(s) + v(s) \tag{5.27}$$

where process noise *w(s)* and measurement noise *v(s)* are mutually independent white noise processes with zero-mean and variances $\sigma_w^2(s)$ and $\sigma_v^2(s)$, respectively. With the state-space equations (5.26) and (5.27), the SRE framework in Sections 5.2.1 and 5.2.2 can be easily applied to the multiplicative cascade rainfall model.

There are some properties of the SRE algorithm that should be noted when being applied to the scale-recursive cascade rainfall model. First, the state and measurement variables in (5.26) and (5.27) are scalar values. Consequently, all covariance matrices *P(s)*'s in the SRE algorithm are also scalar values, and their inverses required for the Kalman gain in equation (5.11) are cheap to calculate.

Second, the scale-homogeneity is normally assumed when using the SRE algorithm with the cascade rainfall model. The scale-homogeneity constraint assumes that the parameters are homogeneous across one scale but may vary from scale to scale (e.g., $P_0(s) = \sigma_{x0}^2(m(s))$, $Q(s) = \sigma_w^2(m(s))$, and $R(s) = \sigma_v^2(m(s))$ for all nodes s on the same scale *m(s)*). Thus, on any M-level tree with k levels of observations, there are *M+k+1* unknown parameters. These parameters are (1) one-*$P_0(0)$* at the root node, (2) M-*Q(s)*'s at all transition levels, and (3) k-*R(s)*'s at all observation levels. These parameters affect the estimation result and their values must be known before applying the SRE algorithm.

Third, when all *F(s)'s* are equal to *1.0*, a prior cross-covariance between any two nodes is simply a variance at their common predecessor (e.g., given by equations (5.4) and (5.5)). If the scale homogeneous assumption is applied, the scale-recursive representation always generates a blocky covariance matrix. This covariance matrix may be inconsistent with true physical properties of rainfall, as mentioned in Section 5.2.3. In addition, a change in the scale-recursive structure (e.g., change in the total number of scales and the branch number) affects the prior covariance matrix, and can influence the estimation of the rainfall state.

Finally, the posterior estimates after incorporating all measurements are no longer scale-homogenous if there is any missing data. From (5.19) and (5.20), it can be seen that the value of *J(s)* at each node varies depending on availability of measurements in its vicinity. Furthermore, the cross-covariance can become even more complex if these posterior rainfall states are dynamically advecting in space. Consequently, we cannot reconstruct the state-space model with *F(s) = 1.0* and with the scale-homogeneity assumption at the next time step. It thus seems that the SRE algorithm with the multiplicative cascade rainfall model is not applicable for the temporal dynamics of rainfall.

5.4 Rainfall Intermittency and the Tree-Pruning Technique

A major problem in applying the multiplicative cascade rainfall model in the SRE framework is rainfall intermittency (i.e., when rainrate is equal to zero). Since the states and the measurements on the tree are the log-transformation of rainfall rate, they are undefined when rainfall intensity is zero. There are a number of proposed techniques to deal with rainfall intermittency on the multiplicative cascade rainfall model. The most prominent and perhaps the simplest method is to set a minimum threshold of rainrate to a small but positive number. This causes the log-transform of the rainrate to be valid, and allows the SRE to be utilized. Nevertheless, it is shown that the estimates strongly depend on the threshold value [42]. From our experience with the threshold technique, the results are generally unstable when the threshold is too low. On the other hand, setting the threshold too high creates strange artifacts and alters the estimation results. By selecting a reasonable threshold, the SRE algorithm can perform well when there are only a few zero rainrate measurements. Unfortunately, a rainfall field usually contains large portions of dry regions. This creates a distribution that concentrates around the threshold value. Consequently, the Gaussian assumption is severely violated, and the estimation accuracy drastically reduces. Alternatively, [42] and [52] proposed a transformation method that raises the measurement to some empirical power. This transformation mimics the Gaussian distribution in a way similar to the logarithmic transformation. This method is less subjective than using the threshold. Nevertheless, when there are many zero measurements, the spatial distribution still concentrates around zero. Again, the Gaussian assumption is severely violated and the estimates can be inaccurate.

Although some methods have been developed to account for intermittency in rainfall estimation, we propose a tree pruning technique that excludes regions with zero rainrate from the SRE algorithm altogether. As the name suggests, the tree pruning technique ignores information on nodes with zero rainrate observations. Beginning at the finest scale, node with the zero-rainfall observation is removed. Then at coarser scales, if all children nodes are removed, the current node is excluded. The pruning procedure continues until we reach the root node. At this stage, we obtain the pruned tree (the solid line in Figure 5-2), consisting of only rainy nodes with finite logarithmic value. This pruned tree is then used to estimate the log-rainfall by the SRE algorithm.

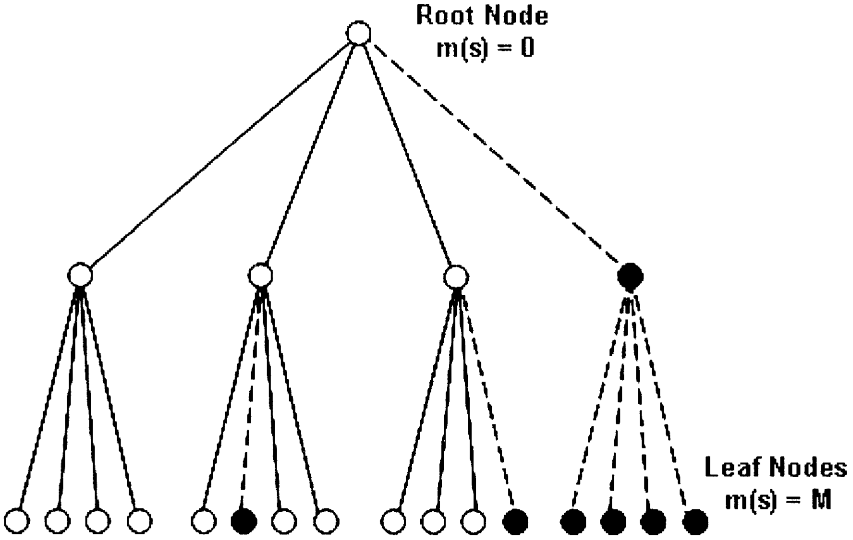


Figure 5-2: A pruned tree with zero-rainfall intensity nodes (black nodes) removed

To make the technique consistent, the tree pruning is based on the measurement at the finest scale available. If any node on the pruned tree has a zero rainfall measurement, we will assume that the measurement at that node is missing. This assumption is based on the fact that the finer scale observations are usually the most accurate in measuring zero rainrate. If we know that any coarser scale measurement is more reliable, we may adjust the method accordingly. With the tree-pruning technique, the SRE algorithm and all its procedures to obtain the smooth estimates remain unchanged. The only exception is that the number of children nodes q in the merging step (e.g., equations (5.14) and (5.15)) has to be adjusted for the new tree structure. Moreover, the Gaussian assumption is valid no matter how large the zero rainfall regions are as long as the underlying distribution of rainy regions is Gaussian. It is important to note that the estimates from the tree pruning technique are sub-optimal estimators in the sense that they do not use all available observations in the update (e.g., they ignore zero rainrate measurements). While the former techniques use all available data including zero-rainrate when updating their states, the concentration around zero or the threshold value skews the distribution and decrease the accuracy of the result. Therefore, the results from the SRE algorithm are more reliable with the tree pruning than the former techniques.

5.5 Parameter Estimation of the SRE Algorithm

There are M+k+1 unknown parameters when the SRE algorithm is applied to the multiplicative cascade rainfall model under the scale-homogeneity assumption. Ultimately, the accuracy of these unknown parameters determines the quality of rainfall estimates from the SRE algorithm. In this section, we introduce the expectation-maximization (EM) algorithm for estimating these unknown parameters. We then focus on the identifiability, uniqueness and sensitivity of model parameters to the change in the scale-recursive structure.

5.5.1 The Expectation-Maximization (EM) Algorithm

The expectation-maximization (EM) algorithm [29] is the iterative algorithm designed to provide the maximum-likelihood (ML) estimates of parameters. This algorithm is suitable for a problem where a direct access to necessary data for estimating the parameters is unavailable, or when some of the data are missing. The EM algorithm is proven to be useful and is commonly used in many applications, including parameter estimation, ARMA modeling, image modeling reconstruction and processing, simultaneous detection and estimation, pattern recognition and neural networking training, etc. Readers should refer to [88] for a list of references, detailed descriptions and applications of the EM algorithm.

The EM algorithm consists of two major steps in each iteration: (1) the expectation step, and (2) the maximization step. The expectation step, or the E-step, computes the conditional expectations of sufficient statistics using the current estimate of the parameters and the posterior states conditioned on all observations. Then the maximization or the M-step re-estimates new model parameters from those sufficient statistics. These two steps are iterated until the parameters converge. The algorithm is known to converge, but possibly to a local instead of the global optimum. However, in many applications, the EM algorithm usually converges to the global optimum solution typically within a few iterations.

5.5.2 The EM-SRE Algorithm for the Multiplicative Cascade Rainfall Model

[65] derived the EM algorithm formula specifically to fit the SRE framework. The combined expectation-maximization with the scale-recursive estimation (EM-SRE) algorithm provides an effective and straightforward framework to estimate both the unknown parameters and the states of interest. Let all unknown parameters of the SRE algorithm be denoted by θ; this represents *P(0)* at the root node, and *F(s), Q(s), H(s),* and *R(s)* for all nodes on the tree. The EM-SRE algorithm estimates the parameters by iteratively maximizing the expected log-likelihood of the observed data from all independent runs $i = 1, \ldots, N$:

$$\hat{\theta}_{EM} = \arg\max_{\theta} \sum_{i=1}^{N} \mathrm{E}\left[\zeta\left(X_A^i, Y_A^i, \theta\right) \mid Y_A^i\right] \tag{5.28}$$

where X_A^i and Y_A^i are the full state vector and the full measurement vector constructed by augmenting all states and observations on the tree for independent run *i*. Then the log-likelihood $\zeta\left(X_i, Y_i, \theta\right)$ is given by

$$\begin{aligned} \zeta\left(X_A^i, Y_A^i, \theta\right) = & -\sum_{s\in\{S_1\}} \left\{\log\left|Q(s)\right| + \left[x(s) - F(s)x(s\gamma)\right]^T Q^{-1}(s)\left[x(s) - F(s)x(s\gamma)\right]\right\} \\ & -\sum_{s\in\{S_2\}} \left\{\log\left|R(s)\right| + \left[y(s) - H(s)x(s)\right]^T R^{-1}(s)\left[x(s) - H(s)x(s)\right]\right\} + const \end{aligned} \tag{5.29}$$

The subset S_1 in equation (5.29) represents the set of all but the root node, while S_2 is the set of all measurement nodes. By maximizing the expected likelihood (5.29) using multivariable regression in the M-step, the new estimates of the SRE parameters are

$$\hat{F}(s) = \llbracket x(s)x^T(s\gamma) \rrbracket \llbracket x(s\gamma)x^T(s\gamma) \rrbracket^{-1} \tag{5.30}$$

$$\hat{Q}(s) = \llbracket x(s)x^T(s) \rrbracket - \hat{A}(s) \llbracket x(s\gamma)x^T(s) \rrbracket \tag{5.31}$$

$$\hat{H}(s) = \llbracket y(s)x^T(s) \rrbracket \llbracket x(s)x^T(s) \rrbracket^{-1} \tag{5.32}$$

$$\hat{R}(s) = \llbracket y(s)y^T(s) \rrbracket - \hat{H}(s) \llbracket x(s)y^T(s) \rrbracket \tag{5.33}$$

$$\hat{P}_0(0) = \llbracket x(0)x^T(0) \rrbracket \tag{5.34}$$

where the operator $\llbracket \bullet \rrbracket$ represents the averages across all independent runs *i* of the expected sufficient statistics over all nodes that share the same statistics, e.g.

$$\llbracket f(x,s) \rrbracket = \frac{1}{N}\sum_{i=1}^{N} \mathrm{E}\left\{f(x,s) \mid Y_A^i\right\} \tag{5.35}$$

These expected quantities in equations (5.30) to (5.34) are calculated in the E-step using posterior estimates and measurements from the SRE algorithm with older parameters from the previous iteration:

$$\mathrm{E}\left[x(s)x^T(s) \mid Y_A^i\right] = P(s \mid Y_A^i) + \hat{x}(s \mid Y_A^i)\hat{x}^T(s \mid Y_A^i) \tag{5.36}$$

$$\mathrm{E}\left[x(s)x^T(s\gamma) \mid Y_A^i\right] = C(s, s\gamma \mid Y_A^i) + \hat{x}(s \mid Y_A^i)\hat{x}^T(s\gamma \mid Y_A^i) \tag{5.37}$$

$$\mathrm{E}\left[y(s)x^T(s) \mid Y_A^i\right] = y(s) \cdot \hat{x}^T(s \mid Y_A^i) \tag{5.38}$$

$$\mathrm{E}\left[y(s)y^T(s) \mid Y_A^i\right] = y(s) \cdot y^T(s) \tag{5.39}$$

where the posterior cross-covariance $C(s, s\gamma \mid Y_A^i)$ can be obtained from (5.19), e.g.

$$C(s, s\gamma \mid Y_A^i) = J(s)P(s\gamma \mid Y_A^i) \tag{5.40}$$

In addition, at the node where the measurement is missing, the expected statistics $\mathrm{E}\left[y(s)x^T(s) \mid Y_A^i\right]$ and $\mathrm{E}\left[y(s)y^T(s) \mid Y_A^i\right]$ in equations (5.38) and (5.38) are calculated from

$$\mathrm{E}\left[z(s)x^T(s) \mid Y_A^i\right] = H(s)\mathrm{E}\left[x(s)x^T(s) \mid Y_A^i\right] \tag{5.41}$$

$$\mathrm{E}\left[y(s)y^T(s) \mid Y_A^i\right] = R(s) + H(s)\mathrm{E}\left[x(s)x^T(s) \mid Y_A^i\right]H^T(s) \tag{5.42}$$

In conclusion, all expected sufficient statistics are calculated from equations (5.36) – (5.42) using the old parameters in the E-step, and the new parameters are updated using equation (5.30) – (5.34) in the M-step. Performing the E and M steps iteratively will eventually will lead to a converged estimation of the parameter set θ; this is commonly referred to as a maximum-likelihood (ML) estimate.

5.5.3 Identifiability and Uniqueness

It was shown that if parameters are scale invariant (e.g. $F(s) = F$, $H(s) = H$, $Q(s) = Q$ for all nodes on the tree) and the measurement noise variance $R(s)$ is given, the EM usually converges to true parameters [52, 65]. From our experiments, the EM-SRE algorithm can also estimate all process noise variances $Q(s)$'s under the scale-homogeneity assumption. The estimated parameters converge regardless of number of measurement scales used as long as all the measurement error variances $R(s)$'s are specified. This implies that the spatial correlation of the observations on the SRE structure contain much more information about the scaling properties and should be able to estimate more unknown parameters than suggested by [52]. This motivates us to explore the maximum limit of the parameters uniquely identifiable by the EM-SRE algorithm.

Using the scale-homogeneity assumption, our multiplicative cascade rainfall model on an M-level tree consists of $M+k+1$ unknown parameters, where k is the total number of measurement scales. These unknown parameters are one-state variance at the root node $P(0)$, M-process noise variances $Q(s)$ at all M transition scales, and k-measurement noise variances $R(s)$ at all k observation scales. To find the maximum number of parameters that can be uniquely identified by the EM-SRE aglroithm, we seek the existing Cramer-Rao Bounds (CRB). Since our underlying distributions are Gaussian according to the multiplicative cascade model assumption, if the CRB exist, the maximum likelihood estimators calculated from the EM-SRE algorithm must be efficient estimators that satisfy the CRB with equality.

Appendix B gives detailed analysis of the CRB for $M+k+1$ unknown parameters of the EM-SRE algorithm with the multiplicative cascade rainfall model. The analysis indicates that the CRB matrix of all $M+k+1$ unknown parameters do not exist. However, the Fisher information, which is an inverse function of the CRB, is rank deficit by only 1 row (or column.) The rank deficiency in the Fisher information is because the rows (or columns) corresponding to the process noise variance at the finest scale $Q(M)$ and the measurement noise at the finest scale $R(M)$ are identical. These two columns always represent *Q(M)* and *R(M)* in the summation form. Therefore, if we treat the summation of process noise variance and measurement noise variance at the finest scale of the M-level tree as one unknown variable, e.g. *Q(M)+R(M)*, the CRB matrix will exist.

The CRB analysis implies that if either the process noise at the finest scale *Q(M)* or the measurement noise variance at the finest scale *R(M)* is specified, the EM-SRE algorithm can estimate up to all the remaining *M+k* parameters. This implication is true regardless of the number of *M* or *k*. Thus, even if we have measurements only at the finest scale, we may be able to estimate all remaining parameters at all other scales assuming that one of the parameters is specified at the finest scale (e.g., *Q(M)* or *R(M)* is specified). Nevertheless, the CRB analysis gives only the maximum numbers parameters that can be estimated. The CRB analysis does not guarantee that all those *M+k* remaining parameters can always be estimated from the EM-SRE algorithm.

We set up a series of synthetic experiments to test whether the EM-SRE algorithm can accurately estimate *M+k* parameters of an M-level tree when *R(M)* is specified according to the CRB analysis. One thousand independent replicates are generated on an 8-level tree with a branching number of 2 x 2 in two-dimensional space using the state-space equations (5.27) and (5.28). The parameters *Q(s)*'s and *R(s)*'s are specified in Table 5.1. The EM-SRE algorithm is employed for two different cases. In the first case, we use the measurement data at all 8 levels to estimate one $P_0(0)$ at the root node, all 8 *Q(s)*'s at all scale and 7 *R(s)*'s at all but the finest scale. The second case uses only the measurements at the finest scale to estimate one $P_0(0)$ at the root node and all 8 *Q(s)*'s at all levels. In both cases, the measurement error variance at the finest scale *R(M)* is given.

Table 5.1: Root-node error variance *P(0) ≡Q(0)*, process error variance *Q(s)*, and measurement error variance *R(s)* used to generate synthetic experiment

Scale m(s)	Dimension	'True' Parameters on an 8-level tree	
		Q(s)	***R(s)***
0	1 x 1	0.10	0.10
1	2 x 2	0.30	0.20
2	4 x 4	0.50	0.30
3	8 x 8	0.70	0.40
4	16 x 16	0.90	0.50
5	32 x 32	0.70	0.40
6	64 x 64	0.50	0.30
7	128 x 128	0.30	0.20
8	256 x 256	0.10	0.10

Ensemble means of the scale-recursive parameters and their ensemble standard deviations that are found using the Monte Carlo technique with 1000 independent replicates are shown in Figure 5-3. For each individual replicate, the parameters are estimated from only one run. In other words, "*i*" in equation (5.35) is equal to 1 and the expectation is an average over all nodes that share the same parameter. Consequently, parameters at the root node and coarser scales are prone to more error because of inadequate sampling average in the expectation values, as indicated by a larger standard deviation. Nevertheless, all parameters estimated from the EM-SRE algorithms converge to the true values, even when the process noise variances *Q(s)*'s and measurement noise variances *R(s)*'s vary with scale. In addition, there are insignificant differences between using all measurements at all scales (case 1) and using only measurements at the finest scale (case 2). Thus, when the measurement noise variance at the finest scale *R(M)* is given, the EM-SRE algorithm can asymptotically estimate all the remaining *M+k* parameters on the M-level tree regardless of the number of measurement scale *k*. In addition, the accuracy of each parameter greatly depends on how many nodes share the same parameters.

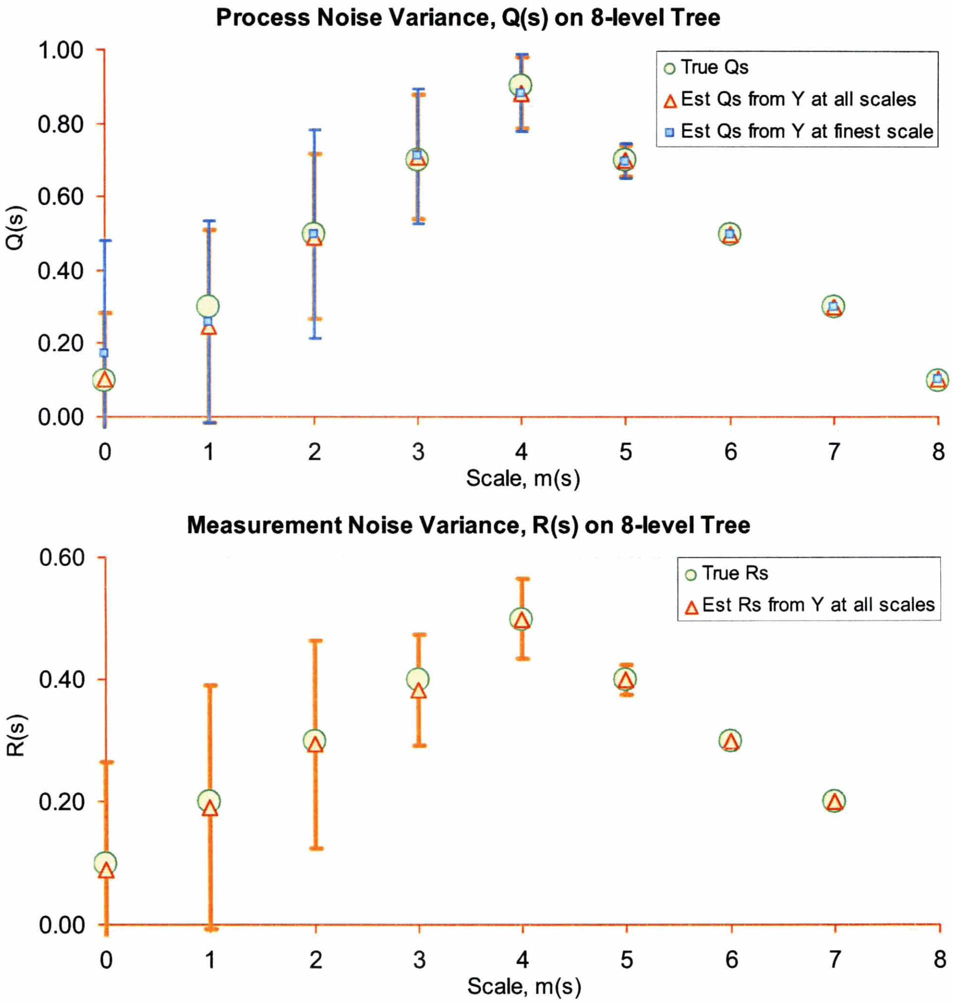


Figure 5-3: (a) Process noise variances and (b) measurement noise variance estimated from the EM-SRE algorithm from: case A – observations from all scales, and case B – observations only from the finest scale

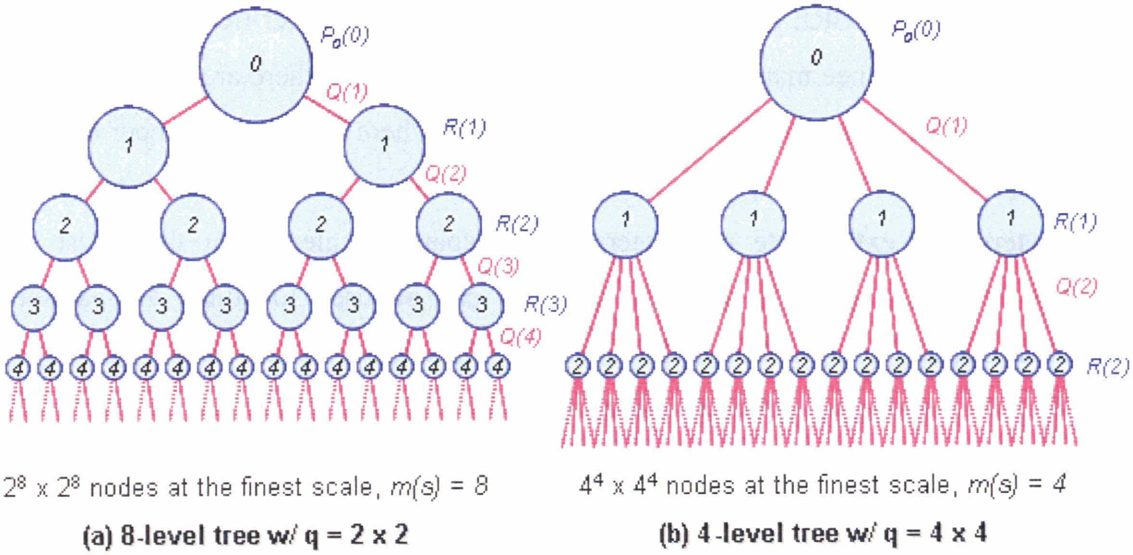
Further experiments showed that the EM-SRE algorithm can estimate all the remaining parameters as well if the process noise variance at the finest scale *Q(M)* is given instead of the measurement noise variance at the finest scale *R(M)*. However, when neither *Q(M)* or *R(M)* is given (e.g., estimating all *M+k+1* unknown parameters on the tree), the EM-SRE algorithm can correctly estimate all process noise variances and measurement noise variances at all scales except for those at the finest scale. At the finest scale, however, the summation of process noise and measurement noise variances *Q(M)+R(M)* converges to the true summation. These results agree well with the CRB analysis mentioned earlier.

It may be possible to estimate all $M+k+1$ unknown parameters using the EM-SRE algorithm if there is a trend on the $Q(s)$ values from scale to scale. The following procedure is proposed. First, we use the EM-SRE algorithm to estimate all $M+k+1$ unknown parameters; thus, we should obtain M+k-1 parameters at all but the finest scale. At the finest scale, we are only interested in the summation $Q(M)+R(M)$ from the EM algorithm and ignore individual values of $Q(M)$ and $R(M)$. Next, we extrapolate for from $Q(s)$'s at coarser scales to find $Q(M)$ at the finest scale. Finally, we subtract the extrapolated $Q(M)$ from the summation $Q(M)+R(M)$ to obtain R(M). Note that we can work with the state variance $P_0(M)$ as well because it is the cumulative sum of $Q(s)$ from the root node down to the current node s. This may become useful if there is a stronger trend in $P_0(s)$ than in $Q(s)$. In this case, the extrapolated $P_0(M)$ is subtracted out from the summation $P_0(M)+R(M)$.

5.5.4 Parameter Sensitivity to Changes in the Tree Structure

The correlations among all nodes on the tree depend on the tree structure as given in equations (5.4) and (5.5). When the SRE algorithm is applied, only one tree structure is usually used. The EM-SRE algorithm will provide the scaling parameters that best fit the data to that particular tree structure. These parameters best describe the correlations among all nodes under that tree constraint. However, the best tree structure to fit the data and give the best correlations is unknown. Unfortunately, it is difficult to anticipate differences in estimated parameters calculated from the EM-SRE algorithm when the tree structure changes. Thus, it is beneficial to investigate the sensitivity of the parameters obtained from the EM-SRE algorithm when the tree structure changes.

To perform the sensitivity analysis, we use the same synthetic observation generated from the previous experiment, but only at every other scale (e.g., $m(s) = 0, 2, 4, 6, 8$). We assume that these observations are from a 4-level tree, as shown in Figure 5-4. Then, we employ the EM-SRE algorithm to find all the scaling parameters of the 4-level tree given that measurement error variance at the finest scale $R(M)$ is known. Finally, we compare scaling parameters of the 4-level tree with those of the original 8-level tree given in the previous section. If the parameters calculated from the EM-SRE algorithm are not sensitive to changes in the tree structure, we would expect to see the same state variances $P_0(s)$ and measurement variances $R(s)$. Moreover, the process noise variance $Q(s)$ of the 4-level tree should be twice the values from the 8-level tree. Nevertheless, we would expect some differences because it is clear that the correlation among all nodes has changed and the EM-SRE algorithm should compensate for these changes.



2^8 x 2^8 nodes at the finest scale, $m(s)$ = 8

(a) 8-level tree w/ q = 2 x 2

4^4 x 4^4 nodes at the finest scale, $m(s)$ = 4

(b) 4-level tree w/ q = 4 x 4

Figure 5-4: One-dimensional diagram of (a) the 8-level tree with branching number of 2 x 2, and (b) the 4-level tree structure with branching number of 4 x 4

The result of the sensitivity analysis is shown in Figure 5-5. The mean of the state variances $P_0(s)$'s and measurement noise variances $R(s)$'s from 1000 independent runs are plotted against the original 8-level tree scales. The result shows that the state variances $P_0(s)$ obtained from both the 4-level tree and the 8-level tree are the very similar. This implies that the state noise variance $P_0(s)$ and the process noise variance $Q(s)$ are insensitive to changes in the tree structures. However, the measurement noise variances $R(s)$ estimated from the 4-level tree are higher than those of the original 8-level tree. This means that the EM-SRE algorithm believes that the error in the cross covariance structure is due to the measurement error and compensates by inflating the measurement noise variance.

We recommended using the maximum number of levels or the tallest tree possible when using the EM-SRE algorithm to estimate the scaling parameters. For example, if there are 256 x 256 grids at the finest scale, it may be more conservative to construct an 8-level tree with a branching number of 2 x 2 instead of 2-level tree with branching number of 16 x 16. Although a taller tree requires more parameters to be estimated, additional cost when using the EM-SRE algorithm is insignificant. With a taller tree, the blocky effect in the cross covariance matrix is less prominent. In addition, there are more correlations among all nodes on the tree, which is beneficial when there are a large number of missing observations. Furthermore, with more parameters it is easier to draw a trend in the parameters and extrapolate parameters at the coarser scales or at the finest scale. Extrapolated parameters at the coarser scales can be used to adjust the parameters obtained directly from the EM-SRE algorithm. Since there are less nodes, parameters at these coarser scales are normally inaccurate. In addition, we can use the extrapolated state variance or process noise variance at the finest scale to identify measurement noise variance at the finest scale as described earlier.

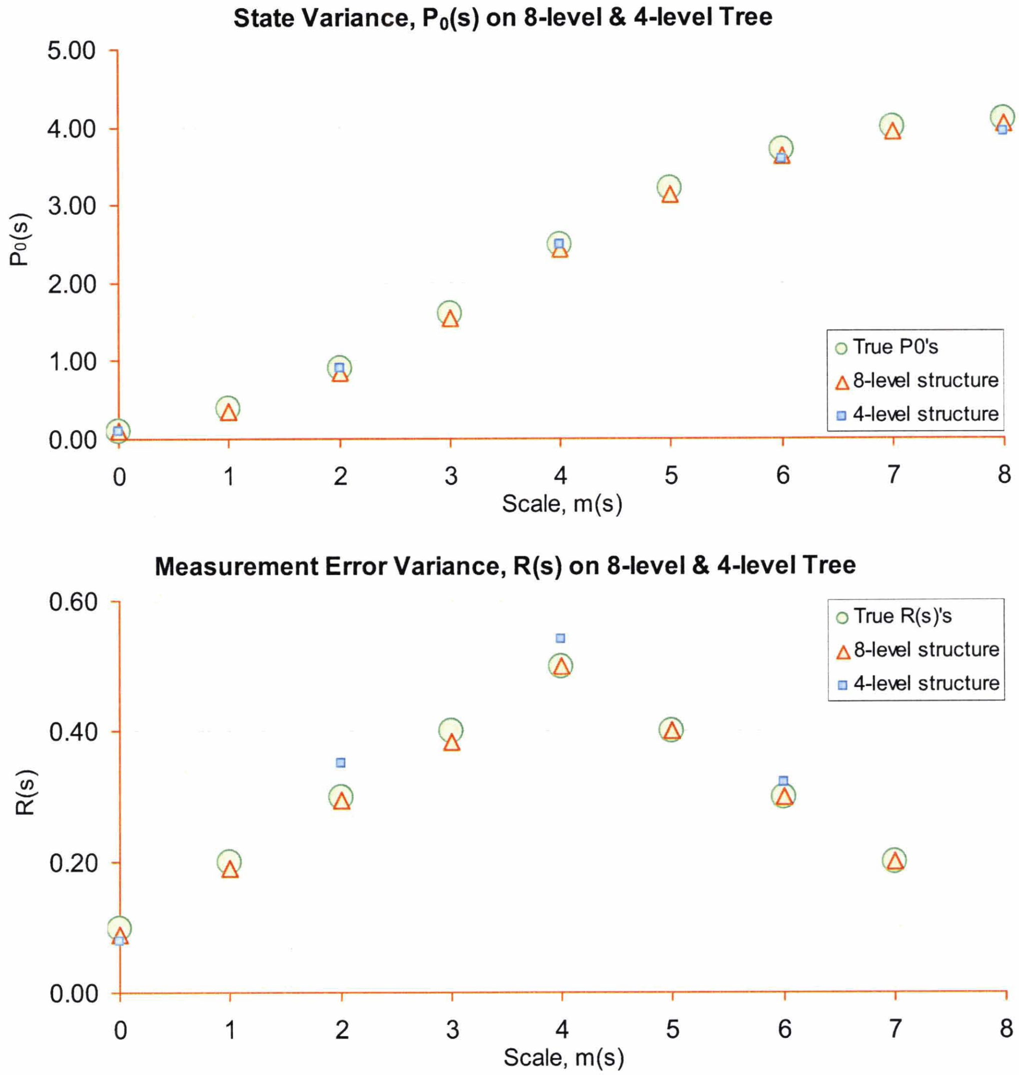


Figure 5-5: (a) State variances and (b) measurement noise variances estimated from the EM-SRE algorithm from the 8-level and 4-level tree structure

5.6 The EM-SRE Algorithm with NOWRAD Measurements

The purpose of this section is to demonstrate that the tree-pruning technique with the EM-SRE algorithm can be used to estimate the scaling parameters and measurement error statistics from real rainfall observation data. We will not focus on multiple scales data assimilation in this section since it has been mentioned in many studies including [47, 52]. In addition, since NOWRAD observations are very accurate and comprehensive in space, having measurements at coarser resolutions is unlikely to improve significantly the accuracy of the rainfall estimates.

Rainfall observations used in this experiment are 15-minute cumulative NOWRAD data provided by the Atmospheric and Environmental Research Incorporation (AER). Rainfall intensity data in Figure 5-6a shows a severe convective storm at 20:30 UTC on Aug 19, 2004 over the latitudes 31.8 °N-33 °N and the longitudes 97.2°W-96 °W, approximately above Oklahoma. The normal probability plot of the log-rainrate observation data after applying the tree-pruning technique is also shown in Figure 5-6b to illustrate that the Gaussian assumption is acceptable.

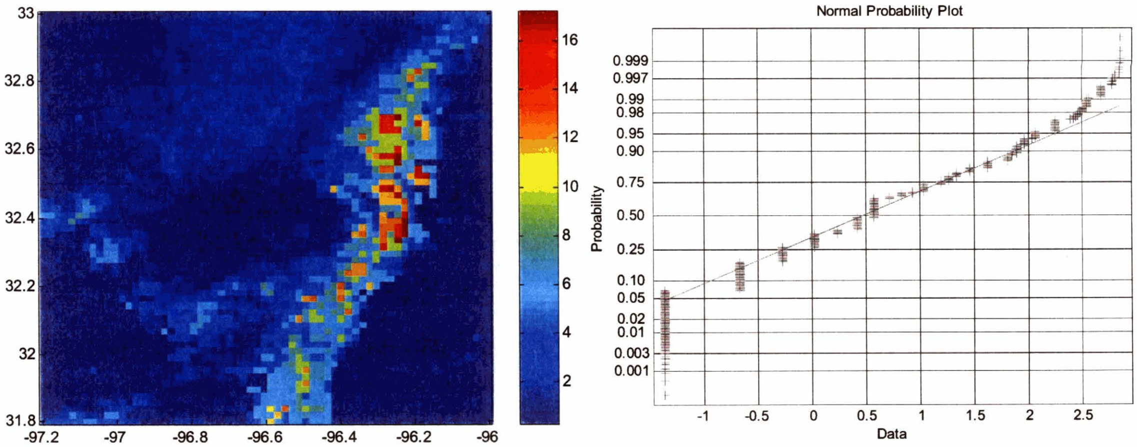


Figure 5-6: (a) 15-minute NOWRAD rainfall observations and (b) the normal probability plot after pruning out zero rainrate of a convective storm on Aug. 19, 2004 at 20:30:00 GMT over the longitude of 97.2° W - 96° W and the latitude of 31.8° N - 33° N

The data are available over a 128 x 128 pixel^2 grid. We use a 6-level tree with the branching number of 2 x 2, and a 3-level tree with the branching number of 4 x 4 to fit this NOWRAD data. Since there is only one measurement scale and the measurement noise variance is unknown, the EM-SRE algorithm can uniquely estimate state variance $P_0(s)$ at all but the finest scale. In addition, it provides the summation of state variance and measurement noise variance at the finest scale $P_0(M)+R(M)$. Figure 5-7 shows the semi-log plot of estimated state variances $P_0(s)$*'s* obtained from the EM-SRE algorithm using the 6-level and the 5-level tree. The state variances are plotted against the length scale ratio, defined as the length of the cell at scale *m(s)* over the length of the root node. This plot is commonly used in multifractal studies and usually establishes a straight-line relationship. However, our result gives a slightly non-linear concave trend in $P_0(s)$.

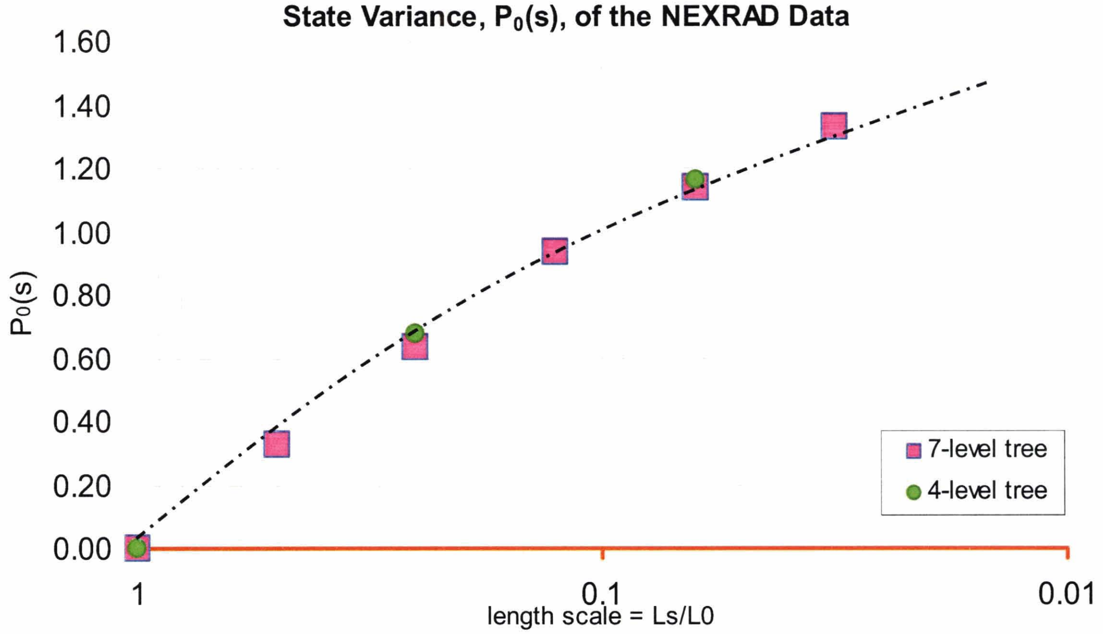


Figure 5-7: State variances $P_x(m_s)$ estimated from the EM algorithm using the NOWRAD data using 7-level tree (pink) and 4-level tree (green) on Aug. 19, 2004 at 20:30:00 GMT over the longitudes of 97.2° W - 96° W and the latitudes of 31.8° N - 33° N; the dashed line is for extrapolating the $P_0(M)$ at the finest scale

To estimate the parameters at the finest scale, we extrapolate to find the state variance at the finest scale $P_0(M)$. The measurement noise variance at the finest scale $R(M)$ can be estimated from subtracting extrapolated $P_0(M)I$ from the summation $P_0(M)+R(M)$ calculated with the EM-SRE algorithm. Table 5-2 summarizes the estimated parameters at the finest scale using the EM-SRE algorithm and the extrapolation technique. The table shows that the NOWRAD observations have log-error variance around 0.03-0.04 units. The multiplicative noise variance, e.g. variance of $V(s)$ in equation (5.24), is given by

$$\sigma_V^2(s) = e^{R(s)} - 1 \tag{5.43}$$

Note that this multiplicative noise variance is unitless with a mean of 1.0. Our experiment indicates that the multiplicative noise variance of the NEXRAD data is around 0.03-0.04, which is considered very accurate. It also implies that the scale-recursive structure fits well with real rainfall field.

Table 5.2: The estimated measurement noise variance obtained from the EM algorithm summation parameters $P_x(M)+R(M)$ and extrapolated Px(M) in Figure 5-7

Tree Structure Used	**Estimates from EM-SRE Algorithm**			**Extrapolate**	**Estimated**
	$P_0(M)$	$R(M)$	$P_0(M)+R(M)$	$P_0(M)$	$R(M)$
7-level tree	1.51	0.01	1.51	1.48	0.03
4-level tree	1.50	0.01	1.52	1.48	0.04

5.7 Conclusions

This chapter presented the combined expectation-maximization with the scale-recursive estimation (EM-SRE) algorithm for estimating rainfall states and scaling parameters conditioned on rainfall observations at various scales. The EM-SRE algorithm can be used to merge static rainfall data given at various measurement scales as well as to estimate unknown parameters including the log-rainfall measurement error variance. This algorithm is simple, yet effective and can be applied to a very large problem. It calculates the best linear least-square estimates of the rainfall state given all observations, and it is optimal if rainfall exhibits a log-normal relationship. In addition, the EM-SRE algorithm iteratively estimates maximum-likelihood scaling parameters that best describe the observations on the tree structures. However, the scale recursive structure results in a blocky effect of the covariance structure. This artifact, however, does not cause significant error in estimating the state and can be minimized by using the tallest tree structure possible. We also focused on the rainfall intermittency problem and proposed the tree-pruning technique to avoid the logarithmic transform of zero rainfall measurement.

With the EM-SRE algorithm, we could estimate all unknown scaling parameters on the tree structure if we specified either process noise variance or the measurement error variance at the finest scale. If there was a trend in process noise variances $Q(s)$'s or state variances $P(s)$'s, we could extrapolate to obtain the values at the finest scale. By subtracting the extrapolated parameters from the summation $Q(M)+R(M)$ or $P_0(M)+R(M)$ obtained from the EM-SRE algorithm, we could estimate the measurement noise variance at the finest scale as well. Hence, it is possible to uniquely identify all M+k+1 scale-homogenous parameters using the EM-SRE algorithm with extrapolation. Moreover, the state variance parameters are not very sensitive to the tree structure, but the measurement noise variances are. This is likely because the EM-SRE algorithm compensates for any error in the cross-correlation by inflating the measurement error.

Chapter 6

Conclusion and Future Research

6.1 Conclusions and Contributions

In Chapter 2, we proposed using the United States Great Plains (USGP) case study to illustrate concepts and implementations of the rainfall data assimilation. In the USGP project, there are six atmospheric forcing and rainfall measurement sources. The atmospheric forcing is provided by infrared cloud-top temperature images from the Geostationary Operational Environmental Satellite (GOES). This measurement was used in the rainfall model to provide information about cloud location and cloud depth. The remaining five rainfall measurement sources are the NOWRAD precipitation product, the Automated Surface Observing Station (ASOS), the Tropical Rainfall Measuring Mission (TRMM), the Special Sensor Microwave Imager (SSM/I), and the Advanced Microwave Sounding Unit-B (AMSU-B). Over the USGP region, accurate rainfall measurements from NOWRAD data is available at very fine frequency and resolution, and would thus dominate the data assimilation results. Consequently, NOWRAD data was excluded from the data assimilation scheme, and was instead used for validation purposes.

The measurement error of each rainfall measurement source was estimated, using NOWRAD data as truth. We divided the measurement error into two categories: (1) position error and (2) intensity error. First, the position error and its statistics were obtained by utilizing the multi-resolution alignment (MRA) algorithm, which lines up rainfall measurements with corresponding NOWRAD data. The average offset distances from multiple storm events were collected and used to provide position error statistics. Second, the intensity error and its statistics were calculated from the residual between NOWRAD measurements and the aligned measurements. With these resulting statistical errors, we could generate a realistic ensemble of perturbed measurements, which are required in the data assimilation algorithm.

To provide a comprehensive rainfall ensemble for the USGP project, we needed a rainfall model that can efficiently describe rainfall characteristics in space and time. This rainfall model was required to propagate local rainfall information from scatter measurement data to a specific place and time. The recursive clustered rainfall (RCR) model was introduced in Chapter 3 to meet these requirements. This spatiotemporal stochastic model was adapted from the cluster-point process rainfall model. It describes the rainfall process in two-dimensions using six parameters and provides a recursive form that can propagate rainfall through time at the any desired spatial or temporal resolution. However, the RCR model by itself cannot account for rainfall intermittency.

To deal with intermittency, the RCR model incorporates cloud-top temperature from GOES. By choosing a reasonable temperature threshold, we can force the model to generate rainfall in realistic locations (e.g., inside deep cloud areas.) Furthermore, two consecutive GOES images can provide the velocity field by utilizing the multi-resolution alignment algorithm. This velocity field is needed for the RCR model to advect existing rainfall. The RCR model, forced with GOES data, is fast, efficient and reliable. It can generate a reasonable rainfall ensemble over a large area without simulating the complex physical dynamics needed for many metrological models. In addition, the RCR model requires only one forcing variable (e.g., GOES cloud-top temperature), which is usually available globally. Chapter 3 ended with the implementation of the RCR model with GOES forcing to generate large-scale rainfall over the USGP region.

After developing an effective rainfall model in Chapter 3, we proposed in Chapter 4 the Ensemble Kalman filter (EnKF) as the appropriate data assimilation algorithm for merging the various sources of rainfall measurements and producing a comprehensive rainfall ensemble. The EnKF uses the state ensemble to calculate the necessary statistics (e.g., ensemble mean and covariance) to sequentially update the ensemble with new measurement information. We also produced a more stable pseudo-inversion technique for the EnKF. This technique can be extremely useful when the ensemble size is much smaller than the state dimension, as is the case in our USGP rainfall data assimilation problem. The EnKF is relatively fast, effective and is proven to work well in many hydrological data assimilation applications. It is a very powerful tool for merging measurement data in real-time applications.

For a reanalysis problem in which the state of interest usually occurred prior to some of the observation times, the EnKF cannot account for any measurements after the estimation time. In this case, the Ensemble Kalman Smoother (EnKS), an extension of the EnKF, is more useful. It accounts for measurements after the estimation time by using the statistics provided by the forward EnKF algorithm. The EnKS is convenient, fast and requires only a minimal amount of extra calculation beyond the EnKF.

We then estimated the model parameters by using the state-augmentation technique. This technique treats the parameters as additional states and updates them using the same methodology applied for updating actual states. Although the state-augmentation technique can estimate parameters in real time, it requires very large ensemble sizes in order to produce reliable parameter results. Because it is impractical to use a large ensemble in our rainfall data assimilation problem, we decided to estimate the parameters offline using many storm events over a smaller region and time interval. We were able to estimate the parameters by constraining them and providing reasonable initial conditions. The time-invariant parameters used in the USGP project were selected from these parameters estimation results. The three main elements in the data assimilation framework were then complete: (1) the RCR model with GOES forcing input with known parameters, (2) rainfall measurements with error statistics, and (3) the analysis algorithm (the EnKF and the EnKS).

Finally, we implemented the data assimilation technique to calculate a comprehensive rainfall ensemble for the USGP project. It was shown that we could merge multiple sources of rainfall measurements and could produce a comprehensive rainfall ensemble using the RCR model and the EnKS algorithm. However, the algorithm cannot provide as fine scale features and characteristics of rainfall in as NOWRAD measurements. It is evident that we cannot replace the NOWRAD data using a simple stochastic rainfall model and available measurements, because they are too scattered in space or time. Nevertheless, given the limited availability and often poor quality of the raw measurements, the data assimilation proposed can provide satisfying coarse scale results over a large region. The algorithm can generate multiple realizations of rainfall, which are essential in many ensemble-based land-surface models. It is fast, efficient and simple relative to many meteorological models currently used.

Chapter 5 presented the Expectation-Maximization and the Scale-Recursive Estimation (EM-SRE) algorithms to estimate scaling parameters for the multiplicative cascade rainfall model. This rainfall model is known for its ability to provide spatial characterization of rainfall process. The EM-SRE algorithm can be used to merge static rainfall measurements from multiple sources. It is an efficient algorithm and can perform well even with missing measurement data. There are two main issues when applying the multiplicative rainfall model. First, the multiplicative cascade model works with the logarithmic transformation of rainfall intensity, and thus cannot cope with rainfall intermittency (i.e., when rainfall intensity is equal to zero). To solve this problem, the tree pruning technique was proposed. The technique excludes zero-measurement nodes from the tree and only uses the remaining nodes to perform the EM-SRE algorithm. Second, there are scaling parameters and measurement uncertainty parameters that must be specified. The EM-SRE algorithm iteratively estimates these parameters along with the state. When the parameters are scale-homogeneous (i.e., the parameter is constant over a single scale but can vary between different scales) and one of the parameters at the finest scale is given, the EM-SRE algorithm can estimate all remaining unknown parameters. Lastly, the combination EM-SRE algorithm was applied to the NOWRAD measurement data to estimate the scaling properties and measurement error variance.

6.2 Suggestions for Future Research

There are three major aspects in our rainfall data assimilation algorithm that can be improved to provide more accurate and reliable ensemble results. These are (1) the rainfall model, (2) the rainfall observation, and (3) the data assimilation technique. In this section, potential improvements in each of these areas are discussed.

6.2.1 Rainfall Models

In the ensemble data assimilation, the dynamic model is used to propagate the state ensemble forward between update times. This propagated ensemble is used to provide the background covariance for an update; thus, the accuracy of the analysis ensemble is closely related to the accuracy of the model used. An advantage of using the ensemble technique is that it does not require an analytical form of the processes or it statistics. Consequently, we are free to modify the model to make forecasts more accurate.

One of the major problems with our RCR model is rainfall intermittency. In Chapter 4, we incorporated GOES cloud-top temperature forcing variable to ensure that between update time steps (1) new rain cells were generated within thick cloud regions, and (2) no rainfall occurs under thin cloud regions. The RCR model with GOES forcing performs well in the presence of small but deep convective thunderstorms. However, for a large storm system with greater deep cloud coverage or for a frontal system, rainfall fields generated from the RCR model are too scattered. It might be possible to modify the usage of GOES or include other atmospheric forcing variables to constrain the location of rain cells. For instance, one might try to use the cloud-top temperature gradient or the Convective Available Potential Energy (CAPE) to locate rain cells instead. It is important, however, that data for a forcing variable should be easy to obtain for the entire region.

Another important component in the rainfall model is its parameters. In this thesis, the RCR model parameters were estimated off-line using the state-augmentation technique. The technique can provide useful estimates of the parameters, but only with serious constraints. Apparently, these estimated parameters are highly sensitive to the constraint and initial condition given. In addition, we assumed that these parameters are time-invariant. The estimation would be more accurate if we can quantify these parameters more objectively through a more robust method. In addition, the accuracy would improved if we could update these parameters in real-time along with the rainfall state.

6.2.2 Rainfall Measurements

The quality of rainfall measurements has a significant impact on the accuracy of the data assimilation results. In this thesis, the stochastic RCR model can only provide rough physical description of the rainfall process. A forecast ensemble generated from the model is usually very scattering due to the uncertainty of its position. As a result, the data assimilation scheme will likely to be uncertain about the forecast and have more confidence in the measurement. Other than incorporating more sources of rainfall measurement, most improvements to the measurement component of the data assimilation are beyond our control (e.g., obtaining finer spatial resolution of the raw data, more frequent revisit times). However, we could improve the measurement model and error statistics.

There are three suggestions that should be considered. First, we ignored the position error of the rain gauge data and used the data to represent the average rainfall over an entire pixel. There may be other ways to introduce this point measurement to update the rainfall state as well as incorporate position error. Second, when the position error perturbation was added to the replicate, the entire measurement field was shifted in one direction with a constant displacement. With this simplification, we ignored spatial correlation, and the possible rotation or twisting of the measurement images. It may be more accurate to generate measurement position perturbations that account for the displacement field in more complex manner. Third, the statistics of the measurement error were estimated from a simple linear regression analysis with an arbitrary weight function. It would be beneficial to use a more robust method or include the measurement uncertainty from the raw data source itself.

6.2.3 Rainfall Data Assimilation Techniques

Among the three factors considered to improve the rainfall data assimilation result, the assimilation technique is perhaps the most significant and extendable subject. The current project utilized the ensemble Kalman filter and the ensemble Kalman smoother. These algorithms ignore the higher moment statistics when updating the state ensemble with measurements. This should not be a serious issue for our results because the state is directly observable. However, when the state is non-linearly related to the measurement, as was the case with parameter estimation using state-augmentation, the Kalman filter-based algorithm may not provide reliable results [6, 120].

Another important consideration is the separation of rainfall position and intensity updates [71, 92, 107]. The current data assimilation method properly deals with amplitude error, but it can perform poorly when there are significant position errors either in the forecast, or in the measurement. By disaggregating the position error problem from the intensity error problem, the data assimilation framework should be improved. Finally, we should be able to speed up the data assimilation algorithm by localizing the calculations in the (vector) subspace spanned by the ensemble perturbations [64, 96].

Appendix A

Multi-Resolution Alignment Algorithm

The multi-resolution alignment (MRA) algorithm [106] is a position adjustment technique [103-107]. The MRA algorithm is similar to the feature calibration alignment (FDA) algorithm [49, 56, 92]. It iteratively solves for the position error problem by minimizing a penalty function based on a gradient and a divergence term. The MRA algorithm is practical for data without well-defined features. The algorithm is also more robust than the correlation based approaches where the displacement is given by the maximum correlation between two patches of images within a searching distance [31, 63]. In addition, the MRA algorithm uses local constraints for relating displacements and represents the displacements as smooth flow fields. This can be useful when working on a large region where characteristics and features vary in space.

The MRA algorithm consists of solving a nonlinear quadratic estimation problem. Solutions to this problem are obtained by regularizing the ill-posed inverse problem. Let $\mathbf{X(r)}$ and $\mathbf{Y(r)}$ be two random vectors defined over a Euclidian grid Ω where $\mathbf{r}^T = \{r_i = (x_i, y_i)^T, i \in \Omega\}$ represent a position vector. Moreover, let $\mathbf{q}^T = \{q_i = (\Delta x_i, \Delta y_i)^T$ $, i \in \Omega\}$ be a displacement vector and **X(r-q)** to represent a displacement of **X** by **q**. Suppose that a random vector Y is linearly related to X via,

$$\mathbf{Y(r) = HX(r - q) + V} \tag{A.1}$$

where **H** is a transformation matrix and **V** is a Gaussian random noise with zero mean and covariance matrix **R**. We assume that all components in (A.1) are Gaussian and write the likelihood function P(**X**,**Y**|**q**) as

$$P(\mathbf{X},\mathbf{Y}\,|\,\mathbf{q}) = \frac{1}{(2\pi)^{n/2}\,|\mathbf{R}|^{1/2}} \exp\left\{-\tfrac{1}{2}\left(\mathbf{Y}-\mathbf{X}(\mathbf{r}-\mathbf{q})\right)^{T}\mathbf{R}^{-1}\left(\mathbf{Y}-\mathbf{X}(\mathbf{r}-\mathbf{q})\right)\right\} \tag{A.2}$$

By using Bayes' rule, we can obtain the probability $P(\mathbf{q}\,|\,\mathbf{X},\mathbf{Y})$ by

$$P(\mathbf{q}\,|\,\mathbf{X},\mathbf{Y}) \propto P(\mathbf{X},\mathbf{Y}\,|\,\mathbf{q})P(\mathbf{q}) \tag{A.3}$$

Assume that the displacement prior density P(q) is given by

$$P(\mathbf{q}) \propto e^{-L(\mathbf{q})} \tag{A.4}$$

where *L(q)* is the energy function. We propose to construct *L(q)* from a smooth flow fields commonly used in the fluid flows. The smoothness assumption leads to Tikhonov type formulation [124]. Particularly, the penalty function *L(q)* is composed of a gradient and a divergence penalty term and expressed in a quadratic form as,

$$L(\mathbf{q}) = \frac{w_1}{2}\sum_{j\in\Omega} tr\left\{\left[\nabla q_j\right]\left[\nabla q_j\right]^{T}\right\} + \frac{w_2}{2}\sum_{j\in\Omega}\left[\nabla\cdot q_j\right]^{2} \tag{A.5}$$

Equation (A.5) represents a weak constraint weighted by the corresponding weights w_1 and w_2. From these sets of equation, we obtain the solution to the displacement field by minimizing the log-likelihood function ξ (i.e., log of equation (A.3)) as,

$$\frac{\partial\xi}{\partial\mathbf{q}} = \nabla\mathbf{X}\big|_{\mathbf{r}\text{-}\mathbf{q}}\,\mathbf{H}^{T}\mathbf{R}^{-1}\left(\mathbf{H}\mathbf{X}(\mathbf{r}-\mathbf{q})-\mathbf{Y}\right)+\frac{\partial L}{\partial\mathbf{q}} \tag{A.6}$$

Using the regularization constraints from (A.5), the solution at node *i* becomes:

$$w_1\nabla^2 q_i + w_2\nabla(\nabla \cdot q_i) + \left[\nabla\mathbf{X}\big|_{\mathbf{r}-\mathbf{q}}\mathbf{H}^T\mathbf{R}^{-1}\left(\mathbf{HX}(\mathbf{r}-\mathbf{q})-\mathbf{Y}\right)\right]_i = 0 \tag{A.7}$$

Equation (A.7) is the field alignment formulation. It is non-linear and is solved iteratively as a Poisson equation. In each iteration, **q** is computed by holding the forcing term constant. The estimate of displacement is then used to deform a copy of the original random variable **X** for the next iteration. The process is repeated until a small displacement residual is obtained, the misfit with **Y** does not improve, or an iteration limit is reaches. Upon convergence, we have an aligned image $\mathbf{X}(\hat{\mathbf{p}})$ and displacement field $\hat{\mathbf{q}} = \sum_{k=1}^{n}\mathbf{q}^{(k)}$ from each displacement $\mathbf{q}^{(k)}$ at iteration k = 1…n.

The convergence of the solution is linearly dependent on the expected displacement between the two fields. It is possible to speed up the computation time by perform the multi-resolution alignment. The multi-resolution alignment begins at the coarser scale by coarsen the resolution of the random field **X** and **Y** and obtain the coarse-scale displacement. At the coarser resolution, the alignment will converge faster because the displacement will be small relative to the coarser resolution. We then rescale the displacement field to the finer-scale, use it to deform the finer resolution image **X**, and solve for another displacement field at the finer resolution. By repeating the process until the resolution of interest is reached. Note that when iteratively solving for (A.7), the unit of the displacement field $\hat{\mathbf{q}}$ is equal to the resolution of the underlying field **X** and **Y**. Therefore, it is essential to rescale the displacement field when we utilize it at different resolution or when perform the multi-resolution approach.

Appendix B

Cramer Rao Bound Analysis for the Scale-Recursive Estimation (SRE) Algorithm

The Cramer Rao bound (CRB) [27] of an unbiased parameter estimator vector $\hat{\theta}(\bullet)$ is given by the inversion of the Fisher information matrix $\mathbf{I}_Y(\theta)$

$$CRB \equiv \mathbf{I}_Y^{-1}(\theta) \tag{B.1}$$

Element *(i,j)* of the Fisher Information matrix is defined by

$$[\mathbf{I}_Y(\theta)]_{ij} = -\mathrm{E}\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j}\zeta_Y(Y;\theta)\right] \tag{B.2}$$

$$= \mathrm{E}\left[\left(\frac{\partial}{\partial\theta_i}\zeta_Y(Y;\theta)\right)^2\right] \tag{B.3}$$

where $\zeta_Y(Y;\theta)$ (i.e., the log-likelihood matrix) is a function of an measurement vector Y given the parameter vector θ. In a special case where Y is jointly Gaussian distributed, the analytical form of the log-likelihood function can be written as:

$$\zeta_Y(Y;\theta) \propto \ln|P_Y| + (Y-\bar{Y})^T P_Y^{-1}(Y-\bar{Y}) \tag{B.4}$$

Consider the state-space equation of the multiplicative cascade rainfall model given in (5.26)-(5.27) and let *X* be a full state vector consisting of all scalar states *x(s)* on the tree. Under the scale-homogeneity assumption, the prior state variance $P_0(s) = P_0(m_s)$, the process noise variance $Q(s) = Q(m_s)$, and the measurement variance $R(s) = R(m_s)$ for all node s at scale $m(s) = m_s$. The full prior covariance matrix $\mathbf{P}_0$ is given by

$$\mathbf{P}_0 = \begin{bmatrix} P_0(0) & \cdots & C_0(0,s) & \cdots & C_0(0,t) & \cdots & C_0(0,N) \\ & \ddots & \vdots & & \vdots & & \vdots \\ & & P_0(s) & \cdots & C_0(s,t) & \cdots & C_0(s,N) \\ & & & \ddots & \vdots & & \vdots \\ & & & & P_0(t) & \cdots & C_0(t,N) \\ & & & & & \ddots & \vdots \\ & & & & & & P_0(N) \end{bmatrix} \tag{B.5}$$

The scalar value $P_0(s)$ represents the prior state variance at node *s*. It is calculated from a cumulative summation of the root node variance and all process noise variance up to the node scale m_s as given by equation (3.3). When all *F(s) = 1.0* and the scale-homogeneity is enforced, the prior state variance can be written as

$$P_0(s) = P_0(m_s) = P_0(0) + \sum_{m_i=1}^{m_s} Q(m_i) \tag{B.6}$$

$C_0(s,t)$ in (B.5) denotes the prior cross-covariance between node *s* and node *t* as given in equation (5.4) and (5.5). Similarly, when all *F(s) = 1.0* and the scale-homogeneity is applied, $C_0(s,t)$ equals the prior state variance at their finest common predecessor node *s^t*.

$$C_0(s,t) = P_0(s \wedge t) = P_0(m_{s \wedge t}) \tag{B.7}$$

The full prior cross covariance matrix $\mathbf{P}_0$ in (B.5) is a positive definite symmetric matrix with dimension of NxN, where N is the total number of nodes on the tree and *m(s)* = *1…M* is the level index of the M-level tree. The diagonal element of $\mathbf{P}_0$ is the state variance at each node on the tree and the off-diagonal element is a cross-covariance between any two nodes on the tree corresponding to the row and column of matrix $\mathbf{P}_0$. Since the finest common predecessor node *s^t* cannot be on the finest scale, any off-diagonal term of $\mathbf{P}_0$ will never contain the process noise variance at the finest scale *Q(M)* where *M* is the finest scale of an M-level tree.

From (3.27), the full measurement vector Y is related to the full state vector X by

$$Y = \mathbf{H}X + V \tag{B.8}$$

where the vector V is an uncorrelated white measurement noise with zero-mean and the covariance matrix **R**. The full measurement error covariance **R** is a diagonal matrix consists of scalars measurement noise variances $R(s)$'s at all measurement nodes. Using (B.6), the full measurement covariance matrix $\mathbf{P}_Y$ is

$$\mathbf{P}_Y = \mathbf{H}\mathbf{P}_0\mathbf{H}^T + \mathbf{R} \tag{B.9}$$

When **H** is a selective matrix consisting of only 1 and 0, $\mathbf{H}\mathbf{P}_0\mathbf{H}^T$ is equivalent to selecting elements in $\mathbf{P}_0$ that corresponds to locations of measurement Y. Since **R** is diagonal, it will only be added to the main diagonal of $\mathbf{H}\mathbf{P}_0\mathbf{H}^T$ matrix.

Again, considering the M-level tree with the scale-homogeneity assumption, unknown parameters set θ consists of M+k+1 unknown parameters. These parameters are one root node state variance $P_0(0)$, M process noise variances *Q(m)*'s at all transition scale *m* = *1…M*, and *k* measurement noise variances *R(m)*'s at all measurement scale $k \leq M+1$. It is easy to verify from (B.7) that the off-diagonal elements of $\mathbf{P}_Y$ do not contain measurement noise variance R(s). The off-diagonal terms have the following form:

$$\mathbf{P}_Y(s,t)\,|_{s\neq t} = P\left(m_{s^\wedge t}\right) = P(0) + \sum_{m_i=1}^{m_{s^\wedge t}} Q(m_i) \tag{B.10}$$

On the other hand, the diagonal elements of $\mathbf{P}_Y$ must contain measurement noise variance and have the following form:

$$\mathbf{P}_Y(s) = P_0\left(m_s\right) + R\left(m_s\right) = \left[P_0(0) + \sum_{m_i=1}^{m_s} Q(m_i)\right] + R\left(m_s\right) \tag{B.11}$$

Since *s^t* cannot be on the finest scale, of diagonal term of $\mathbf{P}_Y$ will not contain *Q(M)*. In addition, the process noise variance at the finest scale *Q(M)* only appears in the main diagonal elements of $\mathbf{P}_Y$ at the finest scale and it always appear in the summation term with the measurement noise variance *R(M)*, e.g. $P_0(0)+\ldots+[Q(M)+R(M)]$. Thus, element in the full measurement covariance matrix $\mathbf{P}_Y$ will consist of *M+k* unique terms with *Q(M)+R(M)* always appears as the summation. If we consider these terms as new *M+k* variables, the log-likelihood function in (B.4) will be a function of these *M+k* variables. Consequently, the Fisher Information matrix having the dimension *(M+k+1)* square is calculated from the 2$^{\text{nd}}$ derivative of the log-likelihood function with respect to a pair of these *M+k+1* unknown parameters.

Since the process noise variance and measurement noise variance at the finest scale always appear as a summation *Q(M)+R(M),* their derivatives of the expected log-likelihood function with respect to either of these two variables are the same, e.g.

$$\frac{\partial}{\partial Q(M)}\zeta\left\{P_0(0),\ldots,Q(M)+R(M)\right\} = \frac{\partial\zeta}{\partial\left[Q(M)+R(M)\right]}\cdot\frac{\partial\left[Q(M)+R(M)\right]}{\partial Q(M)} \tag{B.12}$$

$$= \frac{\partial \zeta}{\partial [Q(M)+R(M)]} \cdot (1) \tag{B.13}$$

$$= \frac{\partial \zeta}{\partial [Q(M)+R(M)]} \cdot \frac{\partial [Q(M)+R(M)]}{\partial R(M)} \tag{B.14}$$

$$= \frac{\partial}{\partial R(M)} \zeta \left\{ P_0(0), \ldots, Q(M)+R(M) \right\} \tag{B.15}$$

where ζ representing the log-likelihood function given by (B.4). It is a function of $P_0(0)$, $Q(1)$, $R(1)$, ..., and $Q(M)+R(M)$. The equity in (B.14) is a result from using the chain rule.

It is evident that if the likelihood-function is given by (5.26) and (5.27) with $F(s) = 1.0$, the partial derivative of the likelihood-function with respect to the $Q(M)$ and that with respect to $R(M)$ are identical. Consequently, the last two rows and last two columns of the Fisher Information matrix given in (B.3) will have the same value regardless of the number of state scales M or the measurement scales k. When all $M+k+1$ unknown parameters are to be estimated, the fisher information is always rank deficit and the Cramer Rao Bound will not exist.

However, in the case when either $Q(M)$ or $R(M)$ is know or the summation $Q(M)+R(M)$ is treated as one unknown parameter, the Fisher information usually have a full rank. As a result, the Cramer Rao Bound will exist. In addition, since all states and measurements are assumed to be jointly Gaussian in the scale-recursive algorithm, the maximum-likelihood estimators calculated form the EM-SRE algorithm are the efficient estimators. Thus, given either $Q(M)$ or $R(M)$, it is possible to obtain up to $M+k$ optimal parameters from the EM-SRE algorithm.

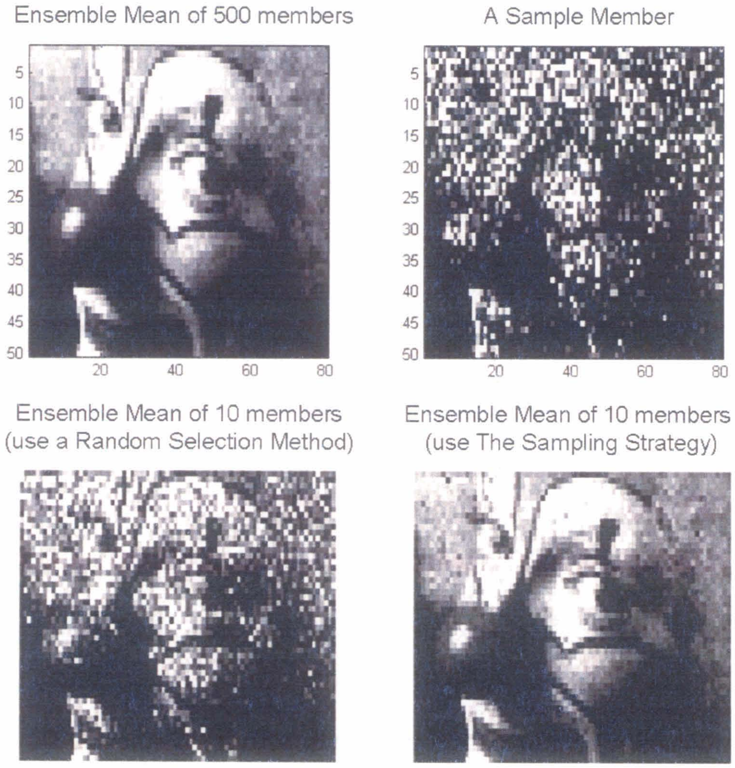
Appendix C

Sampling Strategies

Sampling scheme is one of the most important topic when using the Monte Carlo based technique. The accuracy of the data assimilation problem based on the ensemble method depends on how well we sample the distribution. Effective sampling strategies become even more important when we have limited resources or constraints that allow us to only use small ensemble size. In addition, the sampling strategy is useful for selecting a small number of samples from a population and making sure that these samples can represent the population well. The sampling strategy presented in this chapter uses the concept of eigenvalue spectrum to retain subset of members that span the largest space [39, 100].

Direct application of the sampling scheme to be presented in this section is to improve rainfall estimate from any ensemble approaches. The main idea is to generate large number of ensemble forecasts from the dynamic model, and only select a subset of members that contain the most useful information to be used in the analysis scheme. This scheme is extremely useful when the forecast is less computationally demanding than the analysis scheme. Generally, since the forecast step can be done in parallel, the critical moment that requires the most computation resource should be at the analysis stage where all ensemble members are combined to approximate the distribution and necessary statistics. In stead of using all random samples, a smart sampling strategy will greatly speed up the calculation with very insignificant lost of accuracy.

In additional to direct benefit to the data assimilation framework, an effective sampling strategy is useful when we would like to select a subset of our results to be used in other applications. For example, suppose our rainfall assimilation algorithm can generate up to 500 replicates for each time step but we only need 25 samples rainfall as forcing in a more complicated hydraulic model. How can we effectively sample 25 members that best describe the process? Figure C-1 illustrates a sample benefit of using the sampling scheme to select a small subset of noisy images corrupted by multiplicative log-normal noise. The top left image show the mean from 500 corrupted images with a sample member presented on the top right. The lower left picture shows the mean image from 10 sampling member using the random selection method, while the lower right image show the mean from the sampling strategy to be present in this section. Note that the random selection method is done by calculating the perturbation from the mean, randomly selecting 10 perturbations, and then adding back the mean value.


Ensemble Mean of 500 members

A Sample Member

Ensemble Mean of 10 members (use a Random Selection Method)

Ensemble Mean of 10 members (use The Sampling Strategy)

Figure C-1: Benefit of using the sampling strategy based to a random selection method; the mean and covariance are nicely preserved using the SVD sampling strategy

The derivation of sampling scheme start by defining an error covariance matrix $\mathbf{P} \in \Re^{n \times n}$ with eigenvalue decomposition, $\mathbf{Z\Lambda Z}^T = \mathbf{P}$. To have a sample with maximum rank and best possibly represents the error covariance matrix for a given ensemble size $N << n$, we should sampled from the first N dominant eigenvectors and associated eigenvalues of **P**. In other words, we want to generate a sample matrix $\mathbf{A} \in \Re^{n \times N}$ such that *rank(***A***)* $= N$ and the condition number defined as the ratio between the singular values, $\kappa(\mathbf{A}) = \sigma_1(\mathbf{A})/\sigma_N(\mathbf{A})$, is minimal.

Approximate the error covariance matrix with its ensemble covariance,

$$\begin{aligned}\mathbf{P} \approx \mathbf{P}_e &= \frac{1}{N-1}\tilde{\mathbf{A}}\tilde{\mathbf{A}}^T \\ &= \frac{1}{N-1}\left(\mathbf{U\Sigma V}^T\right)\left(\mathbf{U\Sigma V}^T\right)^T \\ &= \frac{1}{N-1}\mathbf{U\Sigma}^2\mathbf{U}^T\end{aligned}$$

where $\tilde{\mathbf{A}}$ is matrix that holds all ensemble perturbations from the mean, $\mathbf{U} \in \Re^{n \times N}$, $\mathbf{\Sigma} \in \Re^{N \times N}$, and $\mathbf{V} \in \Re^{N \times N}$ are reduce-sized singular value decomposition matrices of $\tilde{\mathbf{A}}$. When the ensemble size N approaches infinity, the n singular vectors in **U** will converge towards the n eigenvectors in **Z** and $\mathbf{\Sigma}^2/(N-1)$ will converge toward the eigenvalues, **Λ.** For a fix ensemble size N, we can improve the rank conditioning of the ensemble by ensuring that the first N singular vectors in **U** are similar to the first N eigenvectors in **Z**.

The implementations of the sampling scheme go as follows. Suppose that we have a large ensemble states $\mathbf{A}_\beta \in \Re^{n \times \beta N}$ and we would like a sub-sample $\mathbf{A} \in \Re^{n \times N}$, i.e. select N members from βN members.

1) Obtain the ensemble perturbations $\tilde{\mathbf{A}}_\beta \in \Re^{n x \beta N}$ by subtracting each column with the ensemble mean $\bar{\mathbf{A}}_\beta \in \Re^{n \times 1}$
2) Compute the reduce-sized SVD: $\mathbf{U}_\beta \mathbf{\Sigma}_\beta \mathbf{V}_\beta^T = \tilde{\mathbf{A}}_\beta$ with $\mathbf{U}_\beta \in \Re^{n \times \beta N}, \mathbf{\Sigma}_\beta \in \Re^{\beta N \times \beta N}$, and $\mathbf{V}_\beta \in \Re^{\beta N \times \beta N}$
3) Store the first *N x N* quadrant of $\mathbf{\Sigma}_\beta$ to a matrix $\mathbf{\Sigma} \in \Re^{N \times N}$
4) Store the first *N* singular vector of $\mathbf{U}_\beta$ to a matrix $\mathbf{U} \in \Re^{n \times N}$
5) Create a random orthogonal matrix $\mathbf{V} \in \Re^{N \times N}$ from right singular vectors of an *N* x *N* random matrix.
6) Obtain the sample perturbations $\tilde{\mathbf{A}} \in \Re^{n \times N}$ from $\tilde{\mathbf{A}} = \frac{1}{\sqrt{\beta}} \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$
7) Obtain sample ensemble matrix $\mathbf{A}$ by adding each column of the sample perturbation matrix $\tilde{\mathbf{A}}$ with the ensemble mean $\bar{\mathbf{A}}_\beta \in \Re^{n \times 1}$ calculated in step 1).

It is important that the ensemble perturbation $\tilde{\mathbf{A}}_\beta$ is used, not the original ensemble matrix $\mathbf{A}_\beta$, because the singular value will only converge to the square root of eigenvalue for a zero mean matrix. Applying the singular value decomposition directly to non-zero mean ensemble matrix will get singular matrix that dominated by the mean value and may cause numerical instability. Secondly, the rescaling factor $1/\sqrt{\beta}$ in step (6) is included to ensure that the variance of the sample will be consistent. As the total size *βN* approaches infinity, the singular vectors and the square of singular value will converge toward the eigenvectors and eigenvalues, respectively. Using the SVD instead of explicit eigenvalue decomposition reduces a lot of computation cost, especially when the dimension is large.

The benefit of using this sampling strategy can be evaluated by the ratio of the largest to the smallest singular value. Figure C-2 illustrates the benefit of sampling 50 members from 100, 150, 200, 250, and 300 total ensemble members, respectively. For an increasing in size of the original ensemble, there are clearly an improvement in the ratio between the first and the 50th singular value. The improvement is, however, greater for less correlated ensemble as in the example of independent measurement noise (top plot) versus more-correlated new rainfall cell noise using the recursive rainfall model (bottom plot) proposed in chapter 4. In addition, the improvement decreases as the number of total increases.

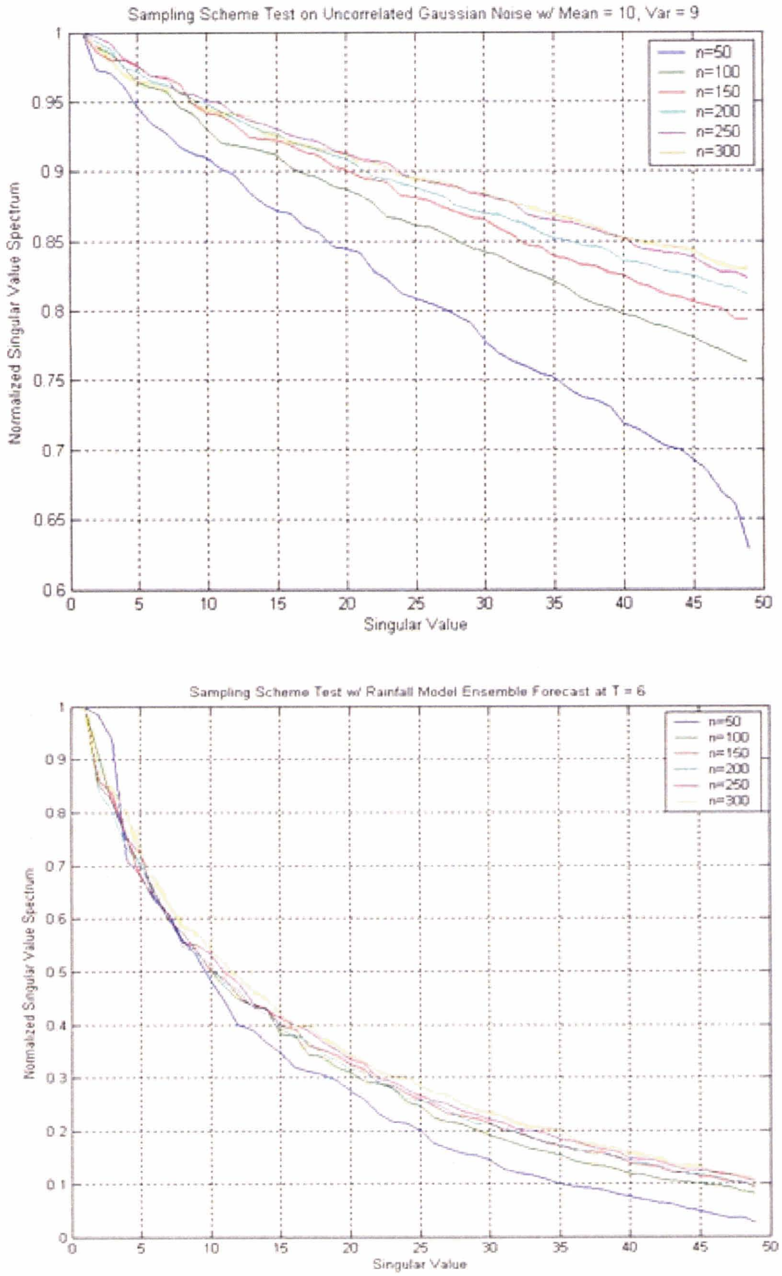


Figure C-2: The first 50 singular vectors of a selective matrix sampling 50 members from 50, 100, 150, 200, 250, and 300 total members

We also apply this sampling strategy to select rainfall replicates generated from the recursive rainfall models given chapter 4. The sample ensemble from this technique has the mean and covariance that are more consistent to the original ensemble than the sample ensemble selected from a random selection method. Figure C-3 shows the benefit of the SVD sampling scheme in comparison to the standard random selection. In the left column we plot the mean from all ensemble members (top) consist of 500 members, the mean from a random selection method (middle), and the mean from the SVD sampling strategy (bottom) both having 10 members. The 2nd and the 3rd column show the first 5 random sampling members from the random selection method and the SVD sampling strategy, respectively. In this plot, the mean is better conserved from the SVD sampling strategy.

The SVD sampling strategy presented in this chapter is a fast and easy to implement technique to optimally select sample member from large population. The algorithms better conserved the mean and covariance of the sampling member better than the random selection method. Thus, we highly recommend to use this sampling strategy every time there need to sample members from full ensemble obtained from the rainfall data assimilation scheme. In addition, if the condition allow, it would be beneficial to dynamically propagate much more ensemble members and use the sampling strategy to select partial of ensemble in the analysis stage so that it is within a computation limit.

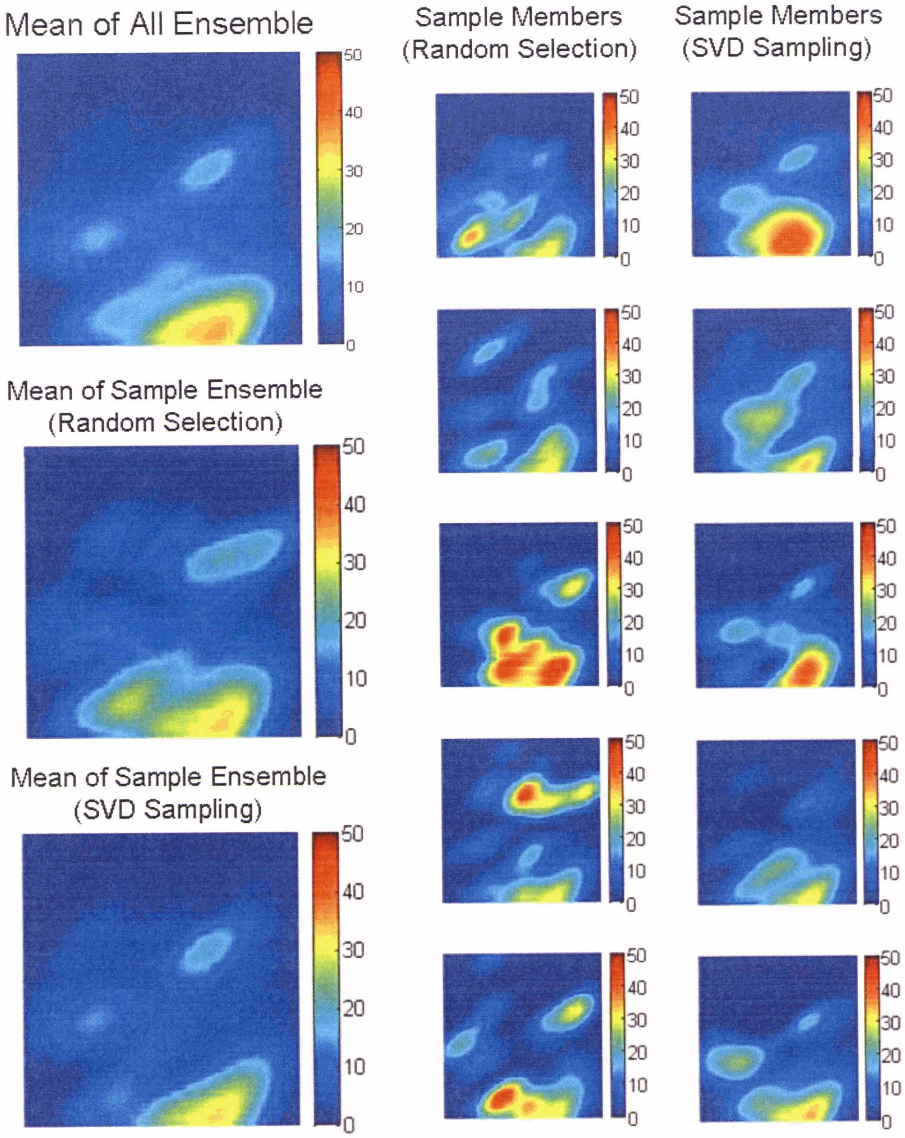


Figure C-3: Comparison between the random selection method and the SVD sampling strategy to select 10 rainfall samples from 500 rainfall members generated by the recursive rainfall model

Bibliography

1. Adler, R.F. and A.J. Negri, *A satellite infrared technique to estimate tropical convective and stratiform rainfall.* Journal of Applied Meteorology, 1988. **27**(1): p. 30-51.
2. Allan, R.J., J. Lindesay, and D.E. Parker, *El niño, southern oscillation & climatic variability*. 1996, Collingwood, Vic., Australia: CSIRO. ix, 405 p.
3. Allen, M., *Do it yourself climate prediction.* Nature, 1999. **401**: p. 642.
4. Annan, J.D., et al., *Parameter estimation in an atmospheric gcm using the ensemble kalman filter.* Nonlinear Processes in Geophysics, 2005. **12**: p. 363-371.
5. Arkin, P.A. and B.N. Meisner, *The relationship between large-scale convective rainfall and cold cloud over the western hemisphere during 1982-84.* Monthly Weather Review, 1987. **115**(1): p. 51-74.
6. Arulampalam, M.S., et al., *A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking.* Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on], 2002. **50**(2): p. 174-188.
7. Austin, P.M. and R.A. Houze, Jr., *Analysis of the structure of precipitation patterns in new england.* Journal of Applied Meteorology, 1972. **11**(6): p. 926-935.
8. Battan, L.J., *Radar observation of the atmosphere.* Rev. ed. 1973, Chicago: University of Chicago Press. x, 324 p.
9. Bennett, A.F., *Inverse methods in physical oceanography.* Cambridge monographs on mechanics and applied mathematics. 1992, Cambridge ; New York: Cambridge University Press. xvi, 346 p.
10. Bennett, A.F., B.S. Chua, and L.M. Leslie, *Generalized inversion of a global numerical weather prediction model.* Meteorology and Atmospheric Physics, 1996. **60**(1 - 3): p. 165-178.
11. Black, T.L., *The new nmc mesoscale eta model: Description and forecast examples.* Weather and Forecasting, 1994. **9**(2): p. 265-278.
12. Black, T.L., D. Deaven, and G. DiMego, *The step-mountain eta coordinate model: 80 km 'early' version and objective verifications.* NWS/NOAA Tech. Procedures Bull., 1993. **412**: p. 31.
13. Boyle, D.P., H.V. Gupta, and S. Sorooshian, *Toward improved calibration of hydrologic models: Combining the strengths of manual and automatic methods.* Water Resources Research, 2000. **36**(12): p. 3663-3674.
14. Buizza, R. and T.N. Palmer, *Impact of ensemble size on ensemble prediction.* Monthly Weather Review, 1998. **126**(9): p. 2503-2518.

15. Buizza, R., et al., *Impact of model resolution and ensemble size on the performance of an ensemble prediction system.* Quarterly Journal of the Royal Meteorological Society, 1998. **124**(550): p. 1935-1960.

16. Burgers, G., P. Jan van Leeuwen, and G. Evensen, *Analysis scheme in the ensemble kalman filter.* Monthly Weather Review, 1998. **126**(6): p. 1719-1724.

17. Chandler, R.E. and H.S. Wheater, *Climate change detection using generalized linear models for rainfall - a case study from the west of ireland. I, preliminary analysis and modelling of rainfall occurrence.* 1998, Department of Statistical Science: University College, London.

18. ---, *Climate change detection using generalized linear models for rainfall - a case study from the west of ireland. Ii, modelling of rainfall amounts on wet days.* 1998, Department of Statistical Science: University College, London.

19. Chen, F. and J. Dudhia, *Coupling an advanced land surface-hydrology model with the penn state-ncar mm5 modeling system. Part i: Model implementation and sensitivity.* Monthly Weather Review, 2001. **129**(4): p. 569-585.

20. ---, *Coupling an advanced land surface-hydrology model with the penn state-ncar mm5 modeling system. Part ii: Preliminary model validation.* Monthly Weather Review, 2001. **129**(4): p. 587-604.

21. Chou, K.C., A.S. Willsky, and A. Benveniste, *Multiscale recursive estimation, data fusion, and regularization.* IEEE Transactions on Automatic Control, 1994. **39**(3): p. 464-478.

22. Colle, B.A., K.J. Westrick, and C.F. Mass, *Evaluation of mm5 and eta-10 precipitation forecasts over the pacific northwest during the cool season.* Weather and Forecasting, 1999. **14**(2): p. 137-154.

23. Courtier, P., J.N. Thepaut, and A. Hollingsworth, *A strategy for operational implementation of 4d-var, using an incremental approach.* Quarterly Journal of the Royal Meteorological Society, 1994. **120**(519): p. 1367-1387.

24. Courtier, P., *Dual formulation of four-dimensional variational assimilation.* Quarterly Journal of the Royal Meteorological Society, 1997. **123**(544): p. 2249-2261.

25. Cowpertwait, P., *Further developments of the neyman-scott clustered point process for modeling rainfall* Water Resource Research, 1991. **27**(7): p. 1431-1438.

26. Cox, D.R. and V. Isham, *A simple spatial-temporal model of rainfall.* Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 1988. **415**(1849): p. 317-328.

27. Cramér, H., *Mathematical methods of statistics.* Princeton mathematical series ; 9. 1946, Princeton,: Princeton University Press. xvi, 575 p.

28. Davis, C., et al., *Development and application of an operational, relocatable, meso-gramma scale weather analysis and forecasting system.* Tellus A, 1999. **51A**: p. 710-727.

29. Dempster, A.P., N.M. Laird, and D.B. Rubin, *Maximum likelihood from incomplete data via the em algorithm.* Journal of Royal Statistical Society, Series B, 1977. **39**(1): p. 1-38.

30. Deque, M., *Ensemble size for numerical seasonal forecasts.* Tellus A, 1997. **49**(1): p. 74-86.

31. Dong, L. and A.L. Boyer, *A portal image alignment and patient setup verification procedure using moments and correlation techniques.* Phys. Med. Biol., 1996. **41**: p. 697-723.

32. Duan, Q., *Calibration of watershed models.* Water science and application ; 6. 2003, Washington, D.C.: American Geophysical Union. vi, 345 p.

33. Dunne, S.C., *Hydrologic data assimilation of multi-resolution microwave radiometer and radar measurements using ensemble smoothing*, in *Civil and Environmental Engineering.* 2006, Massachusetts Institute of Technology: Cambridge. p. 208.

34. Errico, R.M., *What is an adjoint model?* Bulletin of the American Meteorological Society, 1997. **78**(11): p. 2577-2591.

35. Evensen, G., *Using the extended kalman filter with a multilayer quasi-geostrophic ocean model.* Journal of Geophysical Research, 1992. **97**(C11): p. 17905-17924.

36. ---, *Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics.* Journal of Geophysical Research, 1994. **99**(C5): p. 10143-10162.

37. Evensen, G. and P.J. van Leeuwen, *An ensemble kalman smoother for nonlinear dynamics.* Monthly Weather Review, 2000. **128**(6): p. 1852-1867.

38. Evensen, G., *The ensemble kalman filter: Theoretical formulation and practical implementation.* Ocean Dynamics, 2003. **53**(4): p. 343-367.

39. ---, *Sampling strategies and square root analysis schemes for the enkf.* Ocean Dynamics, 2004. **54**(6): p. 539-560.

40. ---, *The combined parameter and state estimation problem.* Submitted to Computational Geosciences, 2005.

41. Foufoula-Georgiou, E. and P. Guttorp, *Assessment of a class of neyman-scott models for temporal rainfall.* Journal of Geophysical Research, 1987. **92**(D8): p. 9679-9682.

42. Foufoula-Georgiou, E., V. Venugopal, and R. Gupta, *Merging of multisensor precipitation estimates: A nonparametric approach based on expectation-maximization and scale recursive estimation.* Journal of Geophysical Research, 2004(Submitted).

43. Gelb, A., *Applied optimal estimation.* 1974, Cambridge, Mass.,: M.I.T. Press. 374 p.

44. Germann, U. and I. Zawadzki, *Scale-dependence of the predictability of precipitation from continental radar images. Part i: Description of the methodology.* Monthly Weather Review, 2002. **130**(12): p. 2859-2873.

45. ---, *Scale dependence of the predictability of precipitation from continental radar images. Part ii: Probability forecasts.* Journal of Applied Meteorology, 2004. **43**(1): p. 74-89.

46. Goodman, B., et al., *A non-linear algorithm for estimating three-hourly rain rates over amazonia from goes/vissr observations.* Remote Sensing Reviews, 1994. **10**: p. 169-177.

47. Gorenburg, I.P., D. McLaughlin, and D. Entekhabi, *Scale-recursive assimilation of precipitation data.* Advances in Water Resources, 2001. **24**(9-10): p. 941-953.

48. Gourley, J.J., et al., *An exploratory multisensor technique for quantitative estimation of stratiform rainfall.* Journal of Hydrometeorology, 2002. **3**(2): p. 166-180.

49. Grassotti, C., H. Iskenderian, and R.N. Hoffman, *Fusion of surface radar and satellite rainfall data using feature calibration and alignment.* Journal of Applied Meteorology, 1999. **38**(6): p. 677-695.

50. Grassotti, C., et al., *Multiple-timescale intercomparison of two radar products and rain gauge observations over the arkansas-red river basin.* Weather and Forecasting, 2003. **18**(6): p. 1207-1229.

51. Grell, G., J. Dudhia, and D.R. Stauffer, *A description of the fifth generation penn state/ncar mesoscale model (mm5).* 1993, NCAR Tech Note, NCAR/TN-398+IA.

52. Gupta, R., *Parametric and non-parametric approaches for validation and blending of multi-sensor precipitation estimates.* 2004, University of Minnesota, 2004. p. xix, 201 leaves.

53. Gupta, V.K. and E. Waymire, *On taylor's hypothesis and dissipation in rainfall.* Journal of Geophysical Research, 1987. **92**(D8): p. 9657-9660.

54. Gupta, V.K. and E.C. Waymire, *A statistical analysis of mesoscale rainfall as a random cascade.* Journal of Applied Meteorology, 1993. **32**(2): p. 251-267.

55. Hoffman, R.N., et al., *Distortion representation of forecast errors.* Monthly Weather Review, 1995. **123**: p. 2758-2770.

56. Hoffman, R.N. and C. Grassotti, *A technique for assimilating ssm/i observations of marine atmospheric storms.* Journal of Applied Meteorology, 1996. **35**: p. 1177-1188.

57. Hoffman, R.N., C. Grassotti, and H. Iskenderian, *Distortion fusion of ground-based radar and satellite-based rainfall data*, in *Final Report NAS13-97010.* 1997, NASA, Stennis Space Center.

58. Holt, F.C. and S.R. Olson, *Goes products and services catalog. 4th ed.* 1999, U.S. Department of Commerce, NOAA/NESDIS.

59. Houze, R.A., *Struccture of atmospheric precipitation systems: A global survey.* Radio Science, 1981. **16**: p. 671-689.

60. Houze, R.A.J. and P.V. Hobbs, *Organization and structure of precipitating cloud systems.* Advances in Geophysics, 1982. **24**: p. 225-315.

61. Huffman, G.J., et al., *Global precipitation estimates based on a technique for combining satellite-based estimates, rain gauge analysis, and nwp model precipitation information.* Journal of Climate, 1995. **8**(5): p. 1284-1295.

62. Joyce, R.J., et al., *Cmorph: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution.* Journal of Hydrometeorology, 2004. **5**(3): p. 487-503.

63. Junck, L., et al., *Correlation methods for the centering, rotation, and alignment of functional brain images.* Journal of Nuclear Medicine, 1990. **31**(7): p. 1220-1226.

64. Kalnay, E., *Atmospheric modeling, data assimilation, and predictability.* 2003, Cambridge, U.K. ; New York: Cambridge University Press. xxii, 341 p., 4 p. of plates.

65. Kannan, A., et al., *Ml parameter estimation of a multiscale stochastic process usingthe em algorithm.* IEEE Transactions on Signal Processing, 2000. **48**(6): p. 1836-1840.

66. Kepert, J.D., *On ensemble representation of the observation-error covariance in the ensemble kalman filter.* Ocean Dynamics, 2004. **54**(6): p. 561-569.

67. Khaliq, M.N. and C. Cunnane, *Modelling point rainfall occurrences with the modified bartlett-lewis rectangular pulses model.* Journal of Hydrology, 1996. **180**(1): p. 109-138.

68. Kivman, G.A., *Sequential parameter estimation for stochastic systems.* Nonlinear Processes in Geophysics, 2003. **10**: p. 253-259.

69. Knutti, R., et al., *Constraints on radiative focring and future climate change from observations and climate model ensembles.* Nature, 2002. **416**: p. 719-723.

70. Kumar, P., *A multiple scale state-space model for characterizing subgrid scalevariability of near-surface soil moisture.* IEEE Transactions on Geoscience and Remote Sensing, 1999. **37**(1): p. 182-197.

71. Lawson, W.G. and J.A. Hansen, *Alignment error models and ensemble-based data assimilation.* Monthly Weather Review, 2005. **133**: p. 1687-1709.

72. Le Cam, L. *A stochastic description of precipitation.* in *Proceeding Fourth Berkeley Symposium on Mathematical Statistics and Probability.* 1961. Berkley, CA: Office of Ordinance Research, U.S. Ary.

73. Li, J., et al., *Sensitivity of north american monsoon rainfall to multisource sea surface temperatures in mm5.* Monthly Weather Review, 2005. **133**(10): p. 2922-2939.

74. Lorenc, A.C., *The potential of the ensemble kalman filter for nwp - a comparison with 4d-var.* Quarterly Journal of the Royal Meteorological Society, 2003. **129**(595): p. 3183-3203.

75. Lovejoy, S. and D. Schertzer, *Multifractals, universality classes and satellite and radar measurements of cloud and rain fields.* Journal of Geophysical Research (D) Atmospheres JGRDE3 Vol. 95, 1990: p. No. 3.

76. Luettgen, M.R. and A.S. Willsky, *Multiscale smoothing error models.* IEEE Transactions on Automatic Control, 1995. **40**(1): p. 173-175.

77. Mackenzie, D., *Ensemble kalman filters bring weather models up to date.* SIAM News, 2003. **36**(8): p. 1-4.

78. Maddox, R.A., et al., *Weather radar coverage over the contiguous united states.* Weather and Forecasting, 2002. **17**(4): p. 927-934.

79. Marsan, D., D. Schertzer, and S. Lovejoy, *Casual space-time multifractal processes: Predictability and forecasting of rainfall fields.* Journal of Geophysical Research, 1996. **101**(D21): p. 26333-26346.

80. Martin, D.W., et al., *Estimates of daily rainfall over the amazon basin.* Journal of Geophysical Research, 1990. **95**(D10): p. 17043-17050.

81. Mason, J., *Numerical weather prediction.* Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 1986. **407**(1832): p. 51-60.

82. Mass, C.F. and Y.-H. Kuo, *Regional real-time numerical weather prediction: Current status and future potential.* Bulletin of the American Meteorological Society, 1998. **79**(2): p. 253-263.

83. McLaughlin, D., et al. *A distributed filtering approach to real-time rainfall forecasting.* in *Preprints, Eighth Conf. on Hydrometeorology.* 1990. Kananaskis Park, Alberta, Canada,: American Meteorological Society.

84. ---, *Computational issues for large-scale land surface data assimilation problems.* Submitted to Journal of Hydrometeorology, 2005.

85. Menabde, M. and M. Sivapalan, *Modeling of rainfall time series and extremes using bounded random cascades and levy-stable distributions.* Water Resources Research, 2000. **36**(11): p. 3293-3300.

86. Miller, R.N., M. Ghil, and F. Gauthiez, *Advanced data assimilation in strongly nonlinear dynamical systems.* Journal of the Atmospheric Sciences, 1994. **51**(8): p. 1037-1056.

87. Mitchell, H.L., P.L. Houtekamer, and G. Pellerin, *Ensemble size, balance, and model-error representation in an ensemble kalman filter{ast}.* Monthly Weather Review, 2002. **130**(11): p. 2791-2808.

88. Moon, T.K., *The expectation-maximization algorithm.* IEEE Signal Processing Magazine, 1996. **13**(6): p. 47-60.

89. Moradkhani, H., et al., *Dual state-parameter estimation of hydrological models using ensemble kalman filter.* Advances in Water Resources, 2005. **28**(2): p. 135-147.

90. Morales, C.A. and E.N. Anagnostou, *Extending the capabilities of high-frequency rainfall estimation from geostationary-based satellite infrared via a network of long-range lightning observations.* Journal of Hydrometeorology, 2003. **4**(2): p. 141-159.

91. Navon, I.M., *Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography.* Dynamics of Atmospheres and Oceans, 1998. **27**(1): p. 55-79.

92. Nehrkorn, T., et al., *Feature calibration and alignment to represent model forecast errors: Empirical regularization.* Quarterly Journal of the Royal Meteorological Society, 2003. **129**(587): p. 195-218.

93. Northrop, P., *A clustered spatial-temporal model of rainfall.* Proceedings of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences, 1998. **454**(1975): p. 1875-1888.

94. Onof, C., et al., *Rainfall modelling using poisson-cluster processes: A review of developments.* Stochastic Environmental Research and Risk Assessment (SERRA), 2000. **14**(6): p. 384-411.

95. Orlandi, A., et al., *Rainfall assimilation in rams by means of the kuo parameterisation inversion: Method and preliminary results.* Journal of Hydrology (Amsterdam), 2004. **288**(1-2): p. 20-35.

96. Ott, E., et al., *A local ensemble kalman filter for atmospheric data assimilation.* Tellus A, 2004. **56**(5): p. 415.

97. Over, T.M. and V.K. Gupta, *Statistical analysis of mesoscale rainfall: Dependence of a random cascade generator on large-scale forcing.* Journal of Applied Meteorology, 1994. **33**(12): p. 1526-1542.

98. ---, *A space-time theory of mesoscale rainfall using random cascades.* Journal of Geophysical Research, 1996. **101**(D21): p. 26319-26332.

99. Perica, S. and E. Foufoula-Georgiou, *Model for multiscale disaggregation of spatial rainfall based on coupling meteorological and scaling descriptions.* Journal of Geophysical Research, 1996. **101**(D21): p. 26347-26362.

100. Pham, D.T., *Stochastic methods for sequential data assimilation in strongly nonlinear systems.* Monthly Weather Review, 2001. **129**(5): p. 1194-1207.

101. Phelan, M.J. and C.R. Goodall, *An assessment of a generalized waymire-gupta-rodriguez-iture model for garp atlantic tropical experiment rainfall.* Journal of Geophysical Research, 1990. **95**(D6): p. 7603-7615.

102. Pielke, R.A., et al., *A comprehensive meteorological modeling system rams.* Meteorology and Atmospheric Physics, 1992. **49**(1 - 4): p. 69-91.

103. Ravela, S., *Shaping receptive fields for affine invariance.* Proceedings of Computer Vision and Pattern Recognition, 2004. **2**(8): p. 725-730.

104. Ravela, S., A. Torralba, and W.T. Freeman, *An ensemble prior of image structure for cross-modal inference.* Proc. 10th International Conference of Computer Vision, 2005. **1**: p. 871-876.

105. Ravela, S. *Amplitude-position formulation of data assimilation.* in *LNCS 3993(III).* 2006: Springer-Verlag Berlin Heidelberg.

106. Ravela, S. and V. Chatdarong, *Rainfall advection using velocimetry by multiresolution viscous alignment*, in *ARXIV physics-0604146.* 2006.

107. Ravela, S., K. Emanuel, and D. McLaughlin, *Data assimilation by field alignment.* Physica(D), to appear, 2006.

108. Reichle, R.H., D.B. McLaughlin, and D. Entekhabi, *Hydrologic data assimilation with the ensemble kalman filter.* Monthly Weather Review, 2002. **130**(1): p. 103-114.

109. Reichle, R.H., et al., *Extended versus ensemble kalman filtering for land data assimilation.* Journal of Hydrometeorology, 2002. **3**(6): p. 728-740.

110. Reichle, R.H. and R.D. Koster, *Assessing the impact of horizontal error correlations in background fields on soil moisture estimation.* Journal of Hydrometeorology, 2003. **4**(6): p. 1229-1242.

111. Rickenbach, T.M., *Cloud-top evolution of tropical oceanic squall lines from radar reflectivity and infrared satellite data.* Monthly Weather Review, 1999. **127**(12): p. 2951-2976.

112. Rodriguez-Iturbe, I. and P.S. Eagleson, *Mathematical models of rainstorm events in space and time.* Water Resources Research WRERAQ Vol. 23, 1987: p. No. 1, p 181-190.

113. Rodriguez-Iturbe, I., D.R. Cox, and V. Isham, *A point process model for rainfall: Further developments.* Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 1988. **417**(1853): p. 283-298.

114. Rodrigueziturbe, I., D.R. Cox, and P.S. Eagleson, *Spatial modeling of total storm rainfall.* Proceedings of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences, 1986. **403**(1824): p. 27-50.

115. Rogers, E., et al., *Changes to the operational "Early" Eta analysis{sol}forecast system at the national centers for environmental prediction.* Weather and Forecasting, 1996. **11**(3): p. 391-413.

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

117. ---, *Fractals, raindrops and resolution dependence of rain measurements.* Journal of Applied Meteorology JAMOAX, 1990. **29**: p. No. 11, p 1167-1170.

118. Simpson, J., R.F. Adler, and G.R. North, *Proposed tropical rainfall measuring mission (trmm) satellite.* Bulletin of the American Meteorological Society BAMIAT Vol 69, 1988: p. No. 3, p 278-295.

119. Stern, R.D., *A model fitting analysis of daily rainfall data.* Journal of Royal Statistical Society, 1984. **147**(1): p. 1-34.

120. Storvik, G., *Particle filters for state-space models with the presence ofunknown static parameters.* IEEE Transactions on Signal Processing, 2002. **50**(2): p. 281-289.

121. Surussavadee, C. and D.H. Staelin, *Comparison of amsu millimeter-wave satellite observations, mm5/tbscat predicted radiances, and electromagnetic models for hydrometeors.* 2006, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139. p. 23.

122. Talagrand, O. and P. Courtier, *Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory.* Quarterly Journal of the Royal Meteorological Society, 1987. **113**(478): p. 1311-1328.

123. ---, *Variational assimilation of meteorological observations with the adjoint vorticity equation. Ii: Numerical results.* Quarterly Journal of the Royal Meteorological Society, 1987. **113**(478): p. 1326-1347.

124. Tikhonov, A.N. and V.I.F.A.F. Arsenin, *Solutions of ill-posed problems.* Scripta series in mathematics. 1977, Washington, New York: Winston. 258 p.

125. Todd, M.C., et al., *A combined satellite infrared and passive microwave technique for estimation of small-scale rainfall.* Journal of Atmospheric and Oceanic Technology, 2001. **18**(5): p. 742-755.

126. Tustison, B., E. Foufoula-Georgiou, and D. Harris, *Scale-recursive estimation for multisensor quantitative precipitation forecast verification: A preliminary assessment.* Journal of Geophysical Research, 2003. **108**(D8).

127. Velden, C., et al., *Recent innovations in deriving tropospheric winds from meteorological satellites.* Bulletin of the American Meteorological Society, 2005. **86**(2): p. 205-223.

128. Velden, C.S., et al., *Upper-tropospheric winds derived from geostationary satellite water vapor observations.* Bulletin of the American Meteorological Society, 1997. **78**(2): p. 173-195.

129. Verhoest, N., P.A. Troch, and F.P. De Troch, *On the applicability of bartlett-lewis rectangular pulses models in the modeling of design storms at a point.* Journal of Hydrology, 1997. **202**(1): p. 108-120.

130. Vicente, G.A., R.A. Scofield, and W.P. Menzel, *The operational goes infrared rainfall estimation technique.* Bulletin of the American Meteorological Society, 1998. **79**(9): p. 1883-1898.

131. Waymire, E., V.K. Gupta, and I. Rodriguez-Iturbe, *Spectral theory of rainfall intensity at the meso-beta scale.* Water Resources Research Vol. 20, 1984: p. No. 10, p 1453-1465.

132. West, M. and J. Harrison, *Bayesian forecasting and dynamic models.* 2nd ed. Springer series in statistics. 1997, New York: Springer. xiv, 680 p.

133. Whiton, R.C., et al., *History of operational use of weather radar by u.S. Weather services. Part ii: Development of operational doppler weather radars.* Weather and Forecasting, 1998. **13**(2): p. 244-252.

134. ---, *History of operational use of weather radar by u.S. Weather services. Part i: The pre-nexrad era.* Weather and Forecasting, 1998. **13**(2): p. 219-243.

135. Zeng, X., *The relationship among precipitation, cloud-top temperature, and precipitable water over the tropics.* Journal of Climate, 1999. **12**(8): p. 2503-2514.

136. Zhou, Y., *Multi-sensor large scale land surface data assimilation using ensemble approaches,* in *Civil and Environmental Engineering.* 2006, Massachusetts Institute of Technology: Cambridge, MA. p. 234.