Characterization of Uncertainty in Remotely-Sensed Precipitation Estimates

by

Seyed Hamed Alemohammad

B.Sc., Sharif University of Technology (2007)
M.Sc., Sharif University of Technology (2009)

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

February 2015

Author .....
Department of Civil and Environmental Engineering
September 2, 2014

Certified by .....
Dara Entekhabi
Bacardi and Stockholm Water Foundations Professor
Thesis Supervisor

Certified by .....
Dennis B. McLaughlin
H.M. King Bhumibol Professor
Thesis Supervisor

Accepted by .....
Heidi M. Nepf
Chair, Departmental Committee for Graduate Students

Characterization of Uncertainty in Remotely-Sensed Precipitation Estimates

by

Seyed Hamed Alemohammad

Submitted to the Department of Civil and Environmental Engineering
on September 2, 2014, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Satellite-derived retrievals of precipitation have increased in availability and improved in quality over the last decade. There are now several satellites in orbit with instruments capable of precipitation retrieval with various degrees of accuracy, spatial resolution and temporal sampling. These retrievals have the advantage of almost full global coverage when compared to surface gauges and ground-based radars. However, there are uncertainties associated with each of these retrievals. This thesis focuses on developing a new framework for characterizing uncertainties in remotely-sensed precipitation estimates. This characterization is a prerequisite if these estimates are to be used in hydrological models. Precipitation forcing is the primary source of uncertainty in surface hydrological models used for forecasting and data assimilation.

In the first part of the thesis, a new metric of error is applied to evaluate precipitation products from Special Sensor Microwave/Imager (SSM/I) instrument. The SSM/I microwave measurements are used for quantitative precipitation rate retrievals and they are key to the development of precipitation data products with high temporal sampling. Results show marked seasonality and precipitation intensity dependence as well as a lower bias at higher intensities and in geographic locations where precipitation rates are generally higher. Next, a new stochastic method is developed to generate spatially intermittent precipitation replicates. These replicates constitute a prior population that can be updated in a Bayesian framework using observations. Bayesian approach allows us to both merge different measurements and investigate the associated uncertainties.

Finally, a new ensemble-based approach to the characterization of uncertainties (in both magnitude (intensity) and phase (location)) associated with precipitation retrieval from space-born instruments is introduced. Unlike previous studies, this method derives the error likelihood using an archive of historical measurements and provides an ensemble characterization of measurement error. The ensemble replicates are generated using the proposed stochastic method, and they are intermittent in space and time. The replicates are first projected in a low-dimensional subspace using a problem-specific set of attributes. The attributes are derived using a dimensionality-reduction approach that takes advantage of singular value decomposition. A non-parametric importance sampling technique is for-

mulated in terms of the attribute vectors to solve the Bayesian sampling problem. Results indicate that this ensemble estimation approach provides a useful description of precipitation uncertainties with posterior ensemble that is narrower in distribution than its prior.

Thesis Supervisor: Dara Entekhabi
Title: Bacardi and Stockholm Water Foundations Professor

Thesis Supervisor: Dennis B. McLaughlin
Title: H.M. King Bhumibol Professor

Acknowledgments

I am indebted to many people for their support throughout my education years, especially the last five years at MIT. First and foremost, I would like to thank my supervisors Prof. Dara Entekhabi and Prof. Dennis McLaughlin for their invaluable support throughout my Ph.D. studies at MIT. I tremendously benefited from their broad-range knowledge and experience, and I am thankful to them for that. They encouraged me to work on this problem and guided me through different stages of it, at the same time, allowed me to become an independent researcher by exploring different aspects of this field. It has been a real pleasure working with them.

I would like to also thank my thesis committee members Prof. Emmanouil Anagnostou and Dr. William Blackwell for their valuable advice and feedback on my work and all the fruitful discussions we had.

My life at MIT has been a great experience thanks to friends and lab-mates. I would like to thank former and current members of Entekhabi and McLaughlin research groups including: Leila Farhadi, Aldrich Castillo, Sigurður Pétur Magnússon, Alexandra Konings, Kaighin McColl, David Whittleson, Parag Narvekar, Anjuli Jain Figueroa, Reetik Kumar Sahu, and Tiziana Smith; as well as former and current members of Parsons Laboratory: Teresa Yamana, Benjamin Scandella, Amir Pahlavan, Mohamed Siam, Mariam Allam, Mason Stahl, Noriko Endo, Rebecca Gianotti, and Gajan Sivandran. Moreover, thanks to my fellow friends in the Persian Students Association at MIT for their friendship and making my experience here so enjoyable.

There is a special thank to Sheila Frankel, James Long and Vicki Murphy who made everything in Parsons easy and smooth. Their dedication and commitment to enhance the working and social environment in Parsons has made me feel home while being away from my home country for five years. I would also like to thank the staff in the Department of Civil and Environmental Engineering headquarter specially Kris Kipp, Pat Dixon, Jeanette Marchocki and Patricia Glidden who have always been there to help and guide how to do different administrative tasks at MIT.

I would like to take this opportunity to thank my Master's studies supervisor Prof. Reza

Ardakanian and my professor and mentor Prof. Ahmad Abrishamchi for their support during my time at Sharif University of Technology and encouragement and advice to pursue my graduate studies.

I wish to express my deep gratitude to my family who have encouraged me with their love and care for the long time. I am indebted to my parents, Manijeh and Ahmad, for their love and understanding through these years. I am grateful to them for their sacrifice, unconditional support and dedication toward my education. I am also thankful to my parent-in-laws, Zahra and Saeed, for their kind emotional support throughout this endeavor. My brother, Hamid, and my brother-in-law, Ali, have always been a great source of friendship and support that I appreciate it. Love you all!

Finally, but most importantly, I am deeply grateful to my wife, Marzieh, for her love, patience and faith in me. I could have not completed this work without her continuous support. I was very fortunate to get to know her at MIT and share many enjoyable moments with her here. I can only hope to bring her as much joy as she has given me.

In loving memory of my grandmother, Pouran, who taught me to be strong.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Precipitation is a major component of the global water and energy cycle and the principal source of uncertainty in modeling the terrestrial water cycle. Characterizing uncertainties in the estimates of precipitation is necessary before these estimates can be used as input into hydrological and meteorological models. Precipitation is inhomogeneous and evolves quickly in time. Moreover, precipitation fields are non-stationary in space. Hence, it is challenging to quantify precipitation and it's associated uncertainties.

Different instruments are used to monitor precipitation. These include gauges, ground-based radars and space-born instruments. Gauges are sparse point measurements and do not describe the spatial variability of precipitation storms. However, they provide relatively accurate data for validation purposes. Ground-based radars provide spatially continuous measurements of precipitation. However, they are limited to land areas (no radar over the oceans) and do not provide good measurements in mountainous regions due to beam blockage by terrain. On the other hand, satellite-based instruments (both radiometer and radars) have almost full global coverage compared to gauges and ground-based radars [127, 40].

Over the past decade, precipitation retrievals from earth-orbiting satellites have increased in availability and improved in quality. Current satellite retrieval techniques are capable of producing precipitation retrievals on high spatial resolutions; however, high uncertainties are present in such high-resolution products. Satellite observations of precipitation are indirect, noisy, and sometimes limited in space and time [93]. Therefore, it

is necessary to characterize the true precipitation using uncertain observations [135]. The common approach toward this characterization is deterministic. Deterministic techniques quantify errors in precipitation estimates by providing the single best estimate of the true precipitation assuming an error model [81, 5, 3, 102, 120]. However, these approaches suffer form assumptions such as parametric distributions for error and additive/multiplicative error models. On the contrary stochastic and ensemble-based approaches provide an elegant way to express these uncertainties [25, 40, 69, 46]. An ensemble of precipitation replicates that captures the true uncertainty can also be used as random forcing in Land Surface Models (LSM) to derive probability distributions of major hydrological variables.

Markov Chain Monte Carlo (MCMC) and Data Assimilation are the two most popular stochastic approaches for uncertainty quantification. MCMC techniques provide probabilistic characterizations through simulation. There are different algorithms to implement the MCMC technique including Gibbs sampling and Metropolis-Hastings [148]; however, they mostly rely on parametric prior and measurement error models. Data assimilation techniques characterize the state of a natural system using data from different uncertain sources (e.g. models and observations). Each source will be given a relative weight based on its uncertainty, and the estimate will balance information from different sources. This characterization will provide an estimate of the true state of the system together with the uncertainty of that estimate [129, 130]. A new approach toward data assimilation is to provide ensemble or interval estimates rather than point estimates. This characterization has many applications in forecasting too (e.g extreme events).

Bayesian estimation theory provides a convenient framework for merging prior information with measurements from a diverse set of data sources. It simultaneously accounts for both model and parameter uncertainties. Moreover, Bayesian approach allows us to update prior ensembles using measurements and investigate the associated uncertainties [30]. In the recent decade, Bayesian-based approaches have been widely used in estimating hydrologic variables especially for probabilistic hydrologic forecasting [91, 35, 147, 133, 47, 138, 183, 48]. However, all of these studies have been focused on quantifying scalar variables or low-dimensional parameter spaces.

Spatial environmental variables are high-dimensional and it's impractical to implement

Bayesian update in high-dimensional space. Therefore, it is necessary to reduce the problem size by mapping data to a reduced-dimension space. In this thesis, for the first time, I propose a framework for applying Bayesian inference to high-dimensional estimates of precipitation and characterize the uncertainty associated with the true precipitation estimation. In this framework, an ensemble of equally probable precipitation replicates is generated using a stochastic prior replicate generator. Then, the prior replicates are assigned a likelihood weight from historical observations. The resulting replicates with updated weights constitute the posterior distribution that characterizes the true precipitation.

In the rest of this chapter, a review of the instruments and global products of precipitation along with a review of error characterization studies are presented. Finally, the structure of the thesis is explained.

1.1 Monitoring Precipitation using Space-born Instruments

This section provides a brief summary of the current state of precipitation estimation using space-born instruments. A more detailed review of this can be found in [144, 92, 13, 161, 112, 96, 97, 176].

There are three groups of instruments on-board satellites that are capable of precipitation retrieval. The first group are Infrared (IR) and Visible (VIS) bandwidth instruments that are mostly on-board the Geostationary Orbit (GEO) satellites. These two types of instruments allow observation of emissions from cloud top or near cloud top and generally cannot be used to directly estimate precipitation intensities. There are empirical relationships developed to estimate precipitation rate from these observations. Auto-Estimator (AE) and its newer version Hydro-Estimator (HE) algorithms use an empirical relationship derived from 6000 collocated radar-satellite pixels to compute convective rain rates from observations of IR instruments [179]. The HE method is now operational in the US National Weather Service for extreme precipitation events analysis. However, the AE is highly dependent on surface radar data to correctly identify non-rainy cold cloud pixels which makes it impossible to use in regions with no radar coverage. Moreover, HE has a tendency to overestimate the area and magnitude of heavy precipitation for cold-topped

(below 281 K) systems and to underestimate the heavy rain that can fall from warm-topped (above 281 K) systems.

The Geostationary Operational Environmental Satellite (GOES) Precipitation Index (GPI) is another algorithm used for tropical rainfall estimation based on cloud top temperature. This method uses the IR observations of cloud top temperature that is far removed from the physics of the precipitation generation process [88].

There are also instruments in IR and VIS bandwidths that are mounted on Low Earth Orbit (LEO) platforms. These include the Moderate-Resolution Imaging Spectroradiometer (MODIS) [100] on-board Aqua and Terra satellites and the Visible and InfraRed Scanner (VIRS) on-board Tropical Rainfall Measuring Mission (TRMM).

In general, precipitation retrieval methods that use only the IR data work well for the tropics that most of the precipitation are originated from convective cloud systems. However, they are less accurate for mid-latitudes that most of the precipitation are originated from advective or stratiform clouds.

The second group are Passive Microwave (PMW) or radiometers. Microwave frequencies have been widely used for precipitation retrieval on-board LEO platforms. Two distinct processes that are used to retrieve precipitation rate from PMW instruments are scattering of the background emission (from the Earth's surface) by the precipitation ice particles and emission from precipitation droplets. The former decreases the background microwave radiation and the later increases the background microwave radiation [176]. Several PMW instruments are now in orbit including: Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-E), Special Sensor Microwave Imager (SSM/I) and its new version Special Sensor Microwave Imager/Sounder (SSMIS), Advanced Microwave Sounding Unit-B (AMSU-B), TRMM Microwave Imager (TMI), and Global Precipitation Measurement (GPM) Microwave Imager (GMI). Each of these instruments have several channels that are capable of detecting different types of precipitation. GMI is the most recent and advanced PMW instrument that has 13 channels from 10 GHz to 183 GHz.

TRMM was the first space mission that was developed to mainly monitor tropical precipitation. It is a joint mission between National Aeronautics and Space Administration (NASA) and Japan Aerospace Exploration Agency (JAXA). The satellite was launched in

1997 and it has been active until recently. On July 25^{th}, 2014 NASA announced that the satellite ran out of fuel after 17 years of successful mission. As a result, TRMM has begun it's downward drift and it will be shutdown when it descends to an orbit of 335 km from its current operating altitude of 402 km [108].

GPM is the latest joint effort of NASA and JAXA to provide a more accurate estimate of global precipitation with improved temporal and spatial coverage. GPM has a core satellite and a constellation of research and operational satellites. GPM core satellite that was launched on February 27^{th}, 2014 is the newest satellite devoted to monitor global precipitation. Unlike TRMM that was limited to tropical regions, GPM has almost full global coverage [82].

The third group of instruments are Active Microwave instruments or radars. There are only two precipitation radars on-board satellites. TRMM and GPM are the two missions that have precipitation radars. TRMM Precipitation Radar (PR) was the first one that has only one frequency channel (13.8 GHz). GPM Dual-frequency Precipitation Radar (DPR) is the newest one that has two channels (13.6 and 35.5 GHz). The dual channel allows monitoring a wider range of precipitation rates. The radar-based retrievals of precipitation are relatively accurate compared to PMW-based ones. However, radar instruments have a narrow swath and a low temporal revisit. Therefore, their retrievals are usually used as calibration standard for precipitation estimation from the more frequent PMW observations.

Single platform estimates of precipitation have low temporal revisit and usually their time of measurement is unevenly distributed across the diurnal cycle. Therefore, it is necessary to combine different retrievals to obtain a reasonably high spatial and temporal sampling resolution. During the last decade several multi-sensor precipitation products are developed among which [192, 84, 104, 190, 83, 87] are now operational, and [26, 175, 105, 173, 71, 94, 122, 27, 54, 9, 188] are products developed for research purposes.

1.2 Error Characterization of Remotely-Sensed Precipitation Estimates

There are many studies that investigate the uncertainties associated with precipitation retrievals from space-born instruments. These studies can be categorized into two groups; those who investigate and evaluate precipitation products [134, 170, 74, 153, 172, 171, 156, 14, 158, 75, 101, 37, 38, 10, 164] and those who model the errors associated with these retrievals [81, 24, 4, 182, 3, 169, 102, 120]. These studies can be further categorized to those who investigate global products that are based on merging retrievals from different instruments, and those who investigate single instrument products.

Two common approaches among those that have modeled the retrieval errors are assuming a parametric distribution for the precipitation rate or the associated error and considering an additive or multiplicative error model. These two approaches provide great tools to model the errors of precipitation products; however, we need to go beyond these simple and deterministic descriptors.

Figure 1-1 shows examples of spatial error distribution of precipitation from several instruments and platforms. These figures reveal that the spatial error in precipitation products from PMW instruments does not have a pattern or parametric distribution. These errors are not structured and can not be modeled and removed easily from the estimates. Furthermore, evaluation of precipitation rates from different instruments provides the same conclusion [10]. Figure 1-2 also shows the Probability Density Function (PDF) of precipitation rate from different instruments as well as the ground-based NEXRAD-IV radar product (ground validation). This figure shows that none of the instruments have the same distribution as NEXRAD-IV and none of them have a parametric and well-defined distribution that can be modeled.

Stochastic methods of synthesizing precipitation fields from satellite data have been also used to investigate uncertainties in satellite estimates; however, some have been focused on quantifying the effects of limited satellite overpasses [21, 22, 17, 159, 66], some have not considered the spatial correlation of precipitation and errors [32, 49, 177, 18, 109, 8, 68], some are only precipitation simulation rather than conditional (on the measurement)

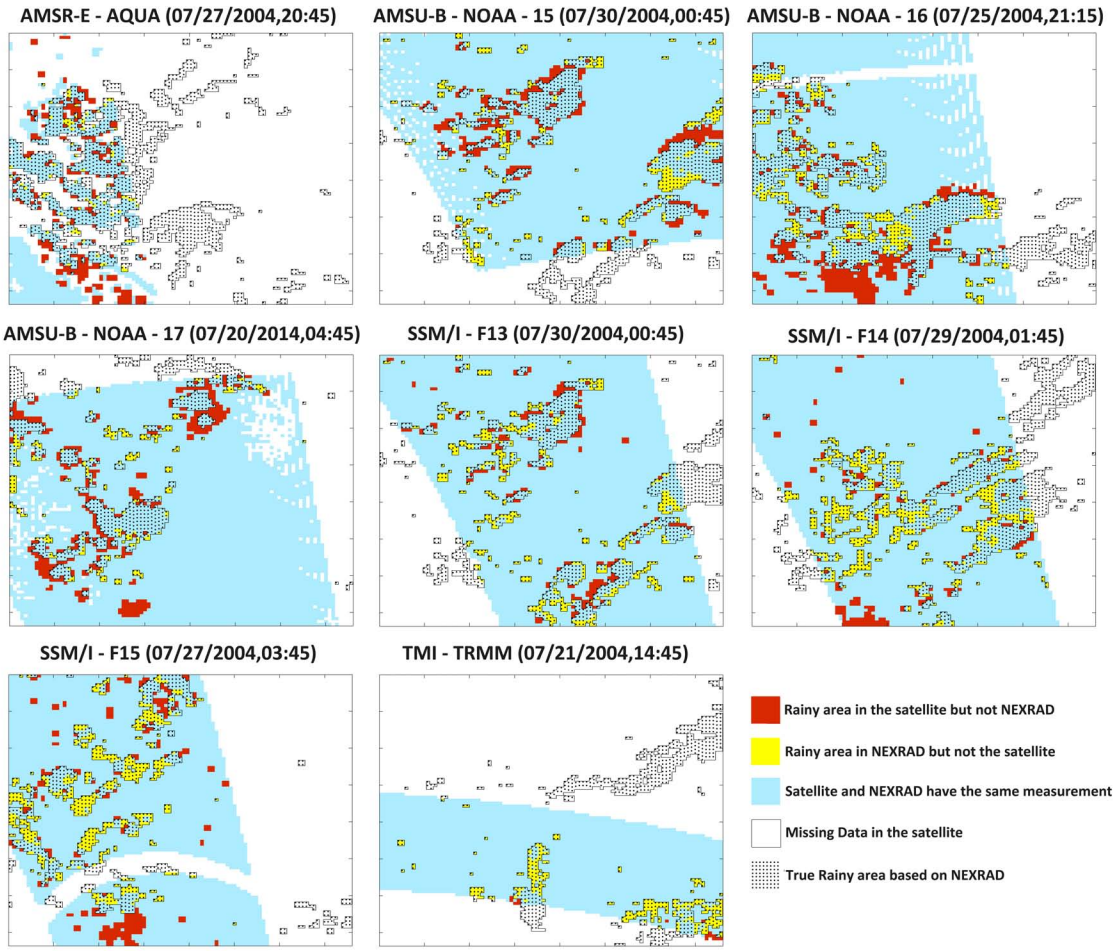


Figure 1-1: Spatial Distribution of Error in Different Instruments/Platforms. Time of measurement is based on UTC.

precipitation estimation [157, 184, 44, 56], and some other use parametric distribution for characterizing the space-time variability of rainfall [73, 139, 40, 167, 141]. (See Section 3.1 for a more detailed explanation of stochastic-based methods.)

These observations reveal that we need to go beyond using simple descriptors for characterizing uncertainties in precipitation estimates. It is necessary to develop a comprehensive and stochastic-based framework to incorporate the uncertainties in the precipitation products using the historical observed error likelihoods.

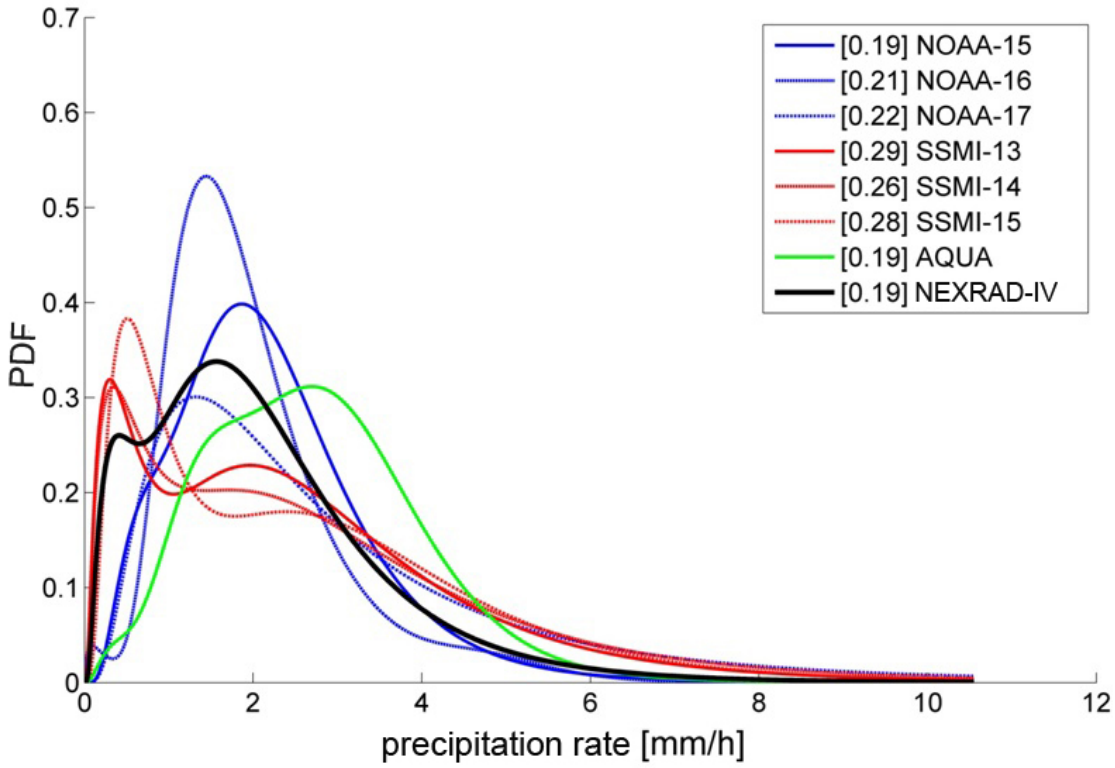


Figure 1-2: Probability Density Function (PDF) of Precipitation Rate based on Measurements from Different Instruments/Platforms. Numbers in the bracket show probability of no precipitation. The data are from Jun. - Aug. 2004 over the Central Part of the Continental US.

1.3 Thesis Structure

The goal of this thesis is to provide an improved characterization of the precipitation rate retrievals from microwave sensors on-board low earth orbit satellites without assuming any parametric distribution. This characterization is based on Bayes' theorem that merges prior information with error statistics from historical observations. To this end, Chapter 2 provides an evaluation of the precipitation products SSM/I instrument. This evaluation together with many others in the literature provide the basis and state of the knowledge about the quality of remotely-sensed precipitation estimates.

Chapter 3 introduces a new stochastic-based method to generate ensemble of precipitation replicates from a given precipitation measurement. This method provides the tool for generating prior replicates to be updated in the Bayesian framework. Chapter 4 presents

the Bayesian framework for characterizing the uncertainties in precipitation estimates. It also introduces the new dimensionality reduction scheme developed for this purpose.

Chapter 5 concludes the findings and provides future directions for research. In addition, the two appendices to this thesis explore some other aspects of the precipitation estimation uncertainties. Appendix A provides a comparison between a newly introduced similarity metric (Jaccard distance, details in section 2.3.2) and the traditional Probability of Detection (POD) and False Alarm Ratio (FAR) metrics. Appendix B presents a new methodology for merging static images of precipitation from two instruments in order to reduce missing data and improve accuracy of precipitation estimates.

Chapter 2

Evaluation of Long-Term SSM/I-based Precipitation Records over Land

Abstract

The record of global precipitation mapping using the Special Sensor Microwave/Imager (SSM/I) instrument measurements now extends over two decades. Similar measurements, albeit with different retrieval algorithms, are to be used in the Global Precipitation Measurement (GPM) mission as part of a constellation to map global precipitation with more frequent data refresh rate. Remotely-sensed precipitation retrievals are prone to both magnitude (precipitation intensity) and phase (position) errors. In this chapter I use the ground-based radar precipitation product from Next Generation Weather Radar Stage-IV product (NEXRAD-IV) to evaluate a new metric of error in the long-term SSM/I-based precipitation records. The new metric quantifies the proximity of two multi-dimensional data sets. Evaluation of the metric across the years shows marked seasonality and precipitation intensity dependence. Drifts and changes in the instrument suite are also evident. Additionally, an estimate of the precipitation retrieval errors conditional on an estimate of background surface soil moisture is presented. The dynamic soil moisture can produce temporal variability in surface emissivity which is a source of error in retrievals. Proper filtering has been applied in the analysis to differentiate between the detection error and retrieval error. The identification of the different types of errors and their dependence on season, inten-

sity, instrument and surface conditions provide guidance to the development of improved retrieval algorithms for use in GPM constellation-based precipitation data products.

2.1 Introduction

From the first launch of the Special Sensor Microwave/Imager (SSM/I) on board the Defense Meteorological Satellite Program (DMSP) platforms in 1987, there has been 25 years of continuous monitoring of the earth atmosphere using this microwave instrument. This includes the launch of the new version of the instrument Special Sensor Microwave Imager Sounder (SSMIS) as well. Figure 2-1 shows the time-line of these two instruments in orbit [28]. This dataset provides a valuable climatological record that can be used for different applications. However, it is necessary to evaluate the products of these instruments at different spatial and temporal scales to have a better and deeper insight to the quality of the records. Furthermore, the results of this evaluation can help guide improved retrieval algorithms.

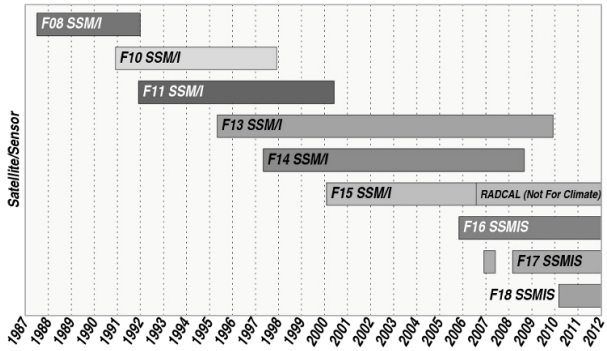


Figure 2-1: Time line of SSM/I and SSMIS instruments in orbit. (*source: [28]*)

There is an increasing need for real-time and near real-time measurements from satellites to be incorporated into meteorological and hydrological hazards decision support systems. A better quantification of errors in space-borne instrument data and associated retrievals is important to guide where efforts need to focus in improving data products algorithms. Currently, there are several multi-platform satellite precipitation products that merge different space-borne estimates of precipitation together and produce global maps of precipitation over land and oceans [84, 87, 83, 190, 104]. All these methods use the bright-

ness temperature measurements from infrared instruments in geostationary orbit as their main input since they provide the space-time continuity of measurements within the instrument viewing disc. Mostly, these approaches calibrate their algorithms using the more accurate but less frequently-refreshed estimates of precipitation from passive microwave (PMW) instruments such as SSM/I in low Earth orbit as well as gauge measurements. However, there has not been a comprehensive spatial and temporal analysis of the quality of the PMW measurements, especially those from SSM/I and its follow-on instruments. It is thus necessary and useful to quantify the uncertainty in the PMW-based retrievals for these merged products.

There is a long history of studies to evaluate the errors associated with PMW-based precipitation estimates. Many of the studies are associated with field campaigns or have limited-duration and limited-coverage domains in their comparisons. Hence, the types of errors such as their dependence on intensity, season, etc. cannot be fully characterized.

During the 1990s, Ferraro and Marks used the ground based radar measurements of rainfall over United States, United Kingdom and Japan to develop a precipitation retrieval algorithm for SSM/I sensor measurements [59]. They categorized the precipitation into several magnitude bins and provided coefficients for nonlinear fits to the instrument data. They found an error of about 10% for scattering algorithm and 20% for emission algorithm. Although the algorithm has gone through revisions, the original version is still used in practice. The SSM/I rain rate product that is used in the present study is also based on this algorithm. Several operational products that produce global maps of precipitation by merging measurements from different sensors take advantage of the SSM/I rain rate based on the Ferraro and Marks algorithm as well (Table 2.1). There has been several studies that use this algorithm [175, 103, 132, 95].

Bell et al., 2001 investigate the monthly-average precipitation estimates based on the SSM/I sensor on-board the DMSP F10 and F11 [23]. They apply spatial averaging to form coarse spatial resolution $2.5^\circ \times 2.5^\circ$ products and evaluate the errors of these products based on independent estimates from surface data and from atmospheric models over the western tropical Pacific (ocean only). Their major finding is that the Root Mean Square Error (RMSE) inferred from the SSM/I measurements is larger than the one based on

Table 2.1: Merged Satellite Estimations of Rainfall that use SSM/I Precipitation Product based on Ferraro and Marks Algorithm

Product Name	Sources	Dates	Temporal Resolution	Pixel Size
CMORPH [89, 87],	PMW & IR	2002 - present	0.5 hr	0.25°
PERSIANN [80, 83],	PMW & IR	2000 - present	3 hr	0.25°
CMAP [191]	PMW& IR	1979 - present	monthly	2.5°
GPCP [2]	PMW & IR	1979- Present	monthly	2.5°

surface data. They also find that the Tropical Rainfall Measurement Mission microwave radiometer-based estimates perform better in the comparisons. Other multi-sensor studies like [51] remove the bias in the SSM/I-based precipitation estimates by calibrating the SSM/I estimate using a more accurate measurement, in this case TRMM precipitation radar (PR). They have also shown that the calibrated estimations have considerably less error. Others like [193] include geographic location and topographic variables (such as surface roughness, slopes facing toward or away from moisture pathways, etc.) to improve the precipitation estimates from SSM/I measurements. [57] uses gauge measurements from the Oklahoma Mesonet network to estimate the errors in instantaneous SSM/I precipitation estimates over land. By incorporating an error model for the Mesonet (as the truth measurement in their study) they estimate that the SSM/I-based precipitation rate over $0.5°$, $1.0°$ and $2.5°$ boxes have an error of 150%, 100% and 70% respectively. This shows the high correlation between the resolution of data and the associated errors in the estimates. Similar results are also evident in the earlier study by [115].

Of the previous studies on PMW-based precipitation retrieval errors, perhaps the most relevant are those by McCollum et al., 2002 [127] and Wolff and Fisher, 2009 [189]. McCollum et al. make an evaluation of the bias in SSM/I-based precipitation retrievals over the contiguous United States by creating a bias-adjusted ground-based radar estimate as the truth. They consider 3-years of data from SSM/I on-board the DMSP F13, F14 and F15. The results show a dependency of the satellite-based data products errors on geographic

location and hence climate. They find that retrievals based on SSM/I on-board F14 and F15 overestimate precipitation over the central part of US by about 45%. However, they have used the NEXRAD Stage-III product which is not a gauges corrected product, and as a result they had to implement a bias adjustment.

Wolff and Fisher, 2009 use four years of data from SSM/I on board DMSP F13, F14 and F15, as well as Advanced Microwave Sounding Unit (AMSU-B on board N15, N16, N17), Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E on board Aqua satellite) and TRMM Microwave Imager (TMI), and evaluate them with ground-based measurements from TRMM ground validation sites. Statistics such as probability density function, correlation coefficient, and performance skill score are reported. In general, they conclude that the PMW-based precipitation estimates tend to overestimate precipitation over land and ocean, and SSM/I and AMSU-B had lower performance skills compared to AMSR-E and TMI.

In a recent study ([180]) , Villa et al. applied a new quality control (QC) scheme to the antenna temperatures of SSM/I and SSMIS sensors and evaluated monthly hydrological products such as rain rate, liquid water path and total precipitable water. They used the Ferraro and Marks algorithm for rain rate retrieval and concluded that the quality controlled antenna temperatures produce a more accurate product on the monthly scale compared to the non-QC version. They also used a histogram matching technique to modify SSMIS temperatures and match them with the SSM/I reference across all the seven channels. The results showed good agreement between the two sensors' products for all the analyzed variables.

Here, I will use longer years of data products based on SSM/I instrument measurements to assess the errors in estimation of precipitation intensity over land. I also explore a larger spatial domain and introduce a new multi-scale metric for evaluating rainfall products. In contrast to [127], I will use NEXRAD-IV radar product which is a gauge corrected rain rate product, and I also use up to 6 years of data for this analysis. The statistics that are presented show the seasonality, and intensity dependence of the errors in more detail. Moreover, I take advantage of their results showing poor quality of radar products over mountainous regions, and excluded those areas from this analysis.

This study specifically aims to quantify the retrieval errors in the precipitation product based on the suite of SSM/I instruments using the Ferraro and Marks algorithm. For ground validation, I use ground-based radar precipitation estimates. Specifically, I use the Next Generation Weather Radar (NEXRAD) Stage IV product that is a merged re-analysis with gauge precipitation station observations. The NEXRAD-IV product provides realistic precipitation estimates that are less prone to error than radar-only estimates (Lin and Mitchell, 2005). In order to have a more comprehensive evaluation, the error statistics of the SSM/I-based precipitation products using instruments on-board DMSP F13, F14 and F15 (hereafter referred to as SSM/I-13, SSM/I-14 and SSM/I-15 respectively) are presented. In order to distinguish between detection error and retrieval error, I have incorporated snow and freeze/thaw state measurements into the comparisons. The samples have been monitored for snow and freeze/thaw status, and those with high snow percentage on the ground and/or frozen condition have been removed from the analysis. The details of this monitoring is provided in Section 2.2.3.

Results show the significant and similar seasonal patterns in errors. The SSM/I-15 statistics reveal the effect of the interference from the radar calibration suite (RADCAL) that was activated on DMSP F15 on August 14, 2006. The study area is over part of the contiguous US where Stage IV data are available, and hence the focus is on land. Since many of the meteorological and hydrological hazards applications of GPM data products with frequent refresh-rate are over land, understanding the error structure and where the algorithms need improvement is important.

Section 2.2 presents the data sets used and the study region along with the monitoring approach for removing detection errors. Section 2.3 includes the statistical measures and the results, while Section 2.4 presents discussions and conclusions.

2.2 Datasets and Monitoring

2.2.1 Study Area

The study region is a large portion of the contiguous US (CONUS) contained within the region bounded by $31°37'30''$N and $47°37'30''$N latitudes and $104°37'30''$W and $80°37'30''$W longitudes. The region is selected based on the results from [127] to exclude mountainous regions in which radar measurements are prone to errors. The spatial scale of comparisons is $0.25°$ (i.e. $0°15'00''$) for both NEXRAD- and SSM/I-based precipitation products. The NEXRAD data are coarse-grained to this resolution since their true resolution is considerably higher. The SSM/I-based products are posted at $0.25°$ even though the instrument measurements inputs are at similar or coarser resolutions. The match between the two products is not complete because of the swath limitations of the satellite instrument measurements. Missing pixel information also contribute to the mismatch. Therefore, I divide the region into six sub-regions each of which cover $8° \times 8°$. Figure 2-2 shows the study area and the six sub-regions. I also exclude the pixels that fall over part of Great Lakes in the North-Eastern part of the region and over the ocean in the South-Eastern part of the region because the Ferraro and Marks algorithm is only for precipitation retrieval over land. Some pixels close to these regions are also removed to mitigate water body contamination.

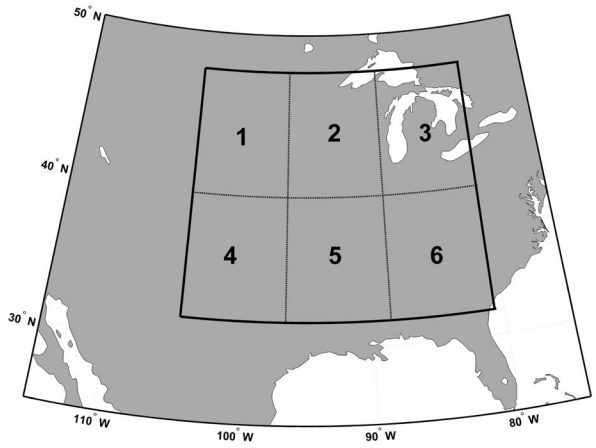


Figure 2-2: The study area shown by the bold rectangular over the contiguous US. The dashed lines show the six sub-regions defined for sampling

Table 2.2: Time domain and number of samples for each of the satellite platforms

Sensor	Start Date	End Date	Number of Samples
SSM/I-13	Jul. 2003	Nov. 2009	1525
SSM/I-14	Jul. 2003	Aug. 2008	1358
SSM/I-15	Jul. 2003	Dec. 2010	2696

2.2.2 Data

The precipitation estimates based on SSM/I instrument measurements are available through the Microwave Surface and Precipitation Products System (MSPPS) Orbital Data from National Oceanic and Atmospheric Association (NOAA). These data products are available on orbital grids and I map them into a $0.25^\circ \times 0.25^\circ$ grid using a nearest neighbor sampling that does not affect the marginal distribution of precipitation intensities. Nearest neighbor sampling is a method that assigns the value of the nearest SSM/I orbital pixel to the center of each grid box. The time domain for each of the satellites is different as listed in Table 2.2.

The retrieval algorithm of MSPSS product is based on [58]. This algorithm uses a scattering approach for rain rate retrieval over land. The Scattering Index is defined based on the earlier works by Grody, 1991 [70]. The SI over land is defined as:

$$SI_L = (451.9 - 0.44 \times TB_{19V} - 1.775 \times TB_{22V} + 0.00575 \times TB_{22V}^2) - TB_{85V} \quad (2.1)$$

In which SI_L is the scattering index over land, TB is the brightness temperature (K) and the subscripts denote the frequency of the vertical polarization channel used. This algorithm has been calibrated with ground based radar measurements of rainfall and the following equation is used to retrieve rain rate:

$$RR = 0.00513 SI_L^{1.9468} \quad (2.2)$$

RR is the rain rate in mm/hr. A minimum threshold of 10 K is set for SI based on earlier studies in the literature that results in a 0.45 mm/hr minimum for rain rate retrieval. Moreover, as the RR increases exponentially with SI based on Equation 2.2, a maximum rain rate of 35 mm/hr is also set for this algorithm, and values higher than that are set to

35 mm/hr. This algorithm is also known as National Oceanic and Atmospheric Administration National Environmental Satellite, Data, and Information Service (NOAA/NESDIS) algorithm for SSM/I.

The ground validation measurements are from the Next Generation Weather Radar (NEXRAD) data set obtained from the National Weather Service's ground-based WSR-88D radar network [64]. The merged radar and gauge estimates produced by the 12 River Forecast Centers (RFCs) in CONUS is mosaicked on a 4 km grid at the National Centers for Environmental Prediction (NCEP). The result is a national product (NEXRAD-IV), and is available through National Center for Atmospheric Research (NCAR) Earth Observing Laboratory (EOL) [116]. I coarse-grain these measurements spatially to a $0.25^\circ \times 0.25^\circ$ grid to be consistent with the SSM/I-based products. The temporal resolution of data are hourly. The NEXRAD-IV data are available from January 2002 to present and cover the study period for each of the satellites considered here.

For generating samples, I match coincident measurements of NEXRAD-IV with each of the satellite products separately over the six sub-regions defined earlier. This means a sample is generated if the satellite overpass completely covers one of the sub-regions without any missing pixels. This condition is necessary for the application of the metrics defined in Section 2.3. The result is different number of samples for each satellite product in Table 2.2. In Section 2.3.4, the comparisons are made over the whole study region using climatological indices, but in the rest of the study the pairs of samples that are generated for the sub-regions are compared.

The freeze/thaw (FT) measurements are obtained from National Snow & Ice Data Center (NSIDC) as part of the MEaSUREs Global Record of Daily Landscape Freeze/Thaw Status product version 2 [98, 99]. This product is a global daily record of FT status derived using microwave observations from SSM/I and Scanning Multichannel Microwave Radiometer (SMMR). The data are provided on a Climate Modeling Grid (CMG) at 25 km grids. I retrieved the AM and PM FT status separately using this dataset, and excluded any pixel that has a frozen status at the time of measurement.

Snow data are from Moderate Resolution Imaging Spectroradiometer (MODIS) instrument on board Terra satellite. This dataset contains daily snow cover data in a 0.05° Climate

Modeling Grid [76]. I coarse-grain these data to a $0.25^\circ \times 0.25^\circ$ spatial grid to match other measurements in this study. These measurements are used as a monitoring tool to exclude pixels that have a high percentage of snow cover from this analysis.

2.2.3 Monitoring Algorithm

In order to filter out the detection errors from the analysis, I used FT and snow measurements. The logic behind this monitoring is that frozen ground or a ground covered with snow have a high brightness temperature and the instrument might detect these pixels as rainy [60]. Therefore, I filtered out all the pixels that have a frozen condition at the time of measurement. As mentioned in the previous section, the FT data provides AM and PM status separately. I used the AM status for any measurement taken before noon and PM status for measurements taken after noon.

The snow cover data has been used in the sampling procedure that was described in Section 2.2.1. If more than 80% of the pixels in a sample is covered by snow (either partially or fully), the sample is rejected. On the other hand, the samples that are collected have a snow tag for each of their pixels. This tag is used during the analysis, and any pixel that is partially or fully covered by snow is not included in the calculations of metrics. For example, when calculating the probability of detection, I assume that the pixel covered by snow does not exist.

In Section 2.3.2, I present a comparison between two groups of samples, first one only includes the samples with no snow cover and in thaw condition, and the second one contains only the samples with snow cover or frozen condition. This comparison shows that there is a considerable difference between the two groups, and it is necessary to remove the samples that have detection errors to have a better understanding of the retrieval errors.

2.3 Statistical Comparisons

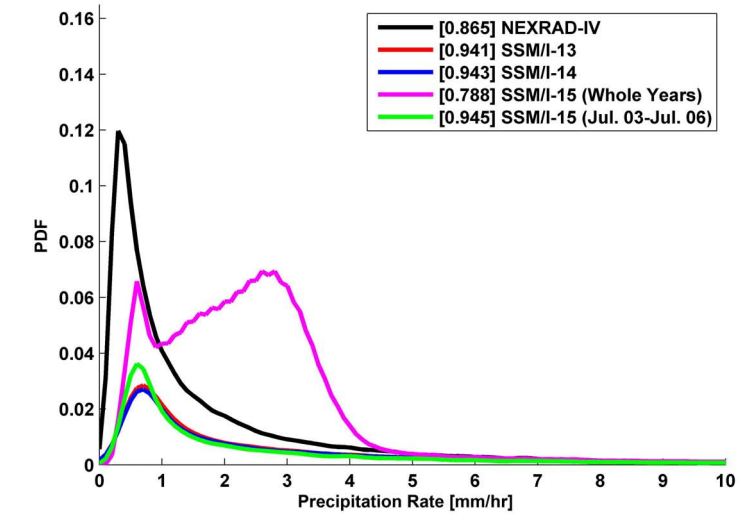
In this section I consider several statistics to compare the SSM/I-based retrievals of precipitation with those based on NEXRAD-IV. These include the Jaccard distance at different rainfall thresholds, the Probability of Detection (POD) and the False Alarm Ratio (FAR),

Probability Density Function (PDF) as well as climatological indices of rain rate. Moreover, in section 2.3 (2.3.5) I investigate the correspondence between precipitation rate retrieval error and surface soil moisture. These are selected strategically in order to differentiate different types of errors.

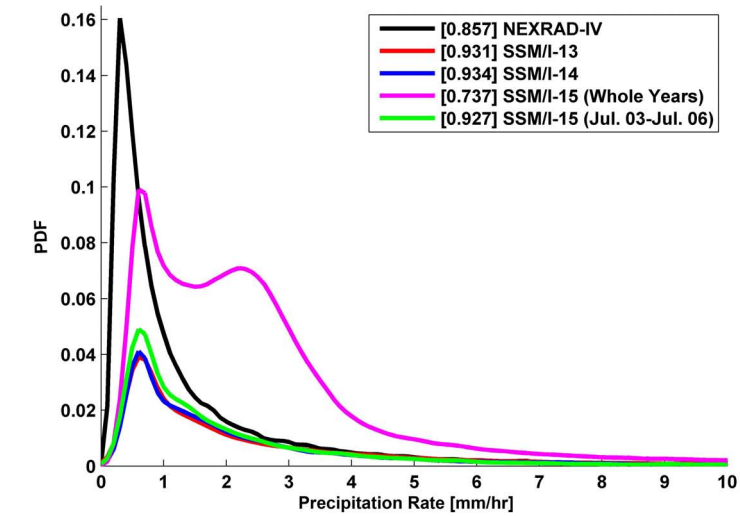
2.3.1 Probability Density Function

Marginal PDFs give basic insight into systematic errors in the magnitude. Figure 2-3 shows the PDFs of precipitation estimates from different datasets (SSM/I-13, SSM/I-14, SSM/I-15 and NEXRAD-IV) for both winter (Oct. - Mar.) and summer (Apr. - Sep.) months. The first result from these two comparisons is that the PDF of precipitation estimates from SSM/I-15 from all the years of data is very different than the other data products. However, removing the data after Jul. 2006 the new PDF is more similar to the other two SSM/I-based PDFs as well as the one from NEXRAD-IV. The reason for selecting this cut-off date is the activation of the radar calibration suite on-board DMSP F15 during August 2006 that has degraded the quality of measurements due to interference. This issue is investigated in more details in the next section.

In general, the three SSM/I-based PDFs have the same shape and magnitude, but compared to the NEXRAD-IV's PDF they have a peak at a slightly higher precipitation rate, and the probabilities are in general lower. The reason is that the probability of no precipitation is higher for the SSM/I-based precipitation products (bracketed numbers in Figure 2-3). This can be partially due to the resolution differences in the coarse-resolution satellite data products and the high-resolution ground-based radar product. Smaller precipitation events may be below the detection threshold of instruments with coarse resolution. Such resolution differences complicate merging PMW measurements into either gap-filling or calibrating other sensor measurements. Moreover, there is a slight difference between the peak of the PDF for SSM/I-13 and 14 and the peak for SSM/I-15. Further analysis showed that sampling SSM/I-13 and 14 for the same time period as SSM/I-15 does not change the qualitative differences between the PDFs. Different probability of no rain (shown in brackets in Figure 2-3) might be a contributing factor that itself might be caused by the



a) Probability Density Function of Precipitation Rate from Different Datasets (Summer)



b) Probability Density Function of Precipitation Rate from Different Datasets (Winter)

Figure 2-3: Probability Density Function (PDF) of precipitation rate from different data products during (a) summer, (b) winter months. The numbers in brackets show the probability of no precipitation

difference in the time of measurement between SSM/-13 and 14 compared to SSM/I-15 (see Figure 2-7).

There is not a major difference between the PDFs from summer months and winter months other than the probability of no precipitation. The slightly larger differences during the summer season are consistent but not necessarily a confirmation of the resolutions-difference interpretation. Note that the NEXRAD-IV measurements used to derive the

PDF in this figure are concurrent with SSM/I-14 measurements. Results not shown here indicate that there are no noticeable differences between the PDF of precipitation from NEXRAD-IV if it is matched with data based on SSM/I-13 or SSM/I-15.

The PDF-based comparisons provide insights into the general distribution of magnitude errors. Because they are marginal distributions they cannot characterize phase (location) errors. Nor can the metric provide concise insights into the dependence on intensity. Other evaluation metrics that are capable of characterizing the latter attributes are Probability of Detection (POD) and False Alarm Ratio (FAR). These two metrics are traditionally used and each is sensitive to a different type of error (but not the other). In the next section, I introduce a new metric that is related to both POD and FAR. This compact metric allows parsimonious evaluation of the data products.

2.3.2 Jaccard Distance

A metric used in this section to quantify proximity between two data sets is called the Jaccard Distance. Jaccard distance is used to evaluate dissimilarity of binary data [163]. Let **x** and **y** be two objects that consists of *n* binary attributes. By comparing these two objects, four quantities can be defined:

f_{00} = number of attributes that **x** is 0 and **y** is 0;

f_{10} = number of attributes that **x** is 1 and **y** is 0;

f_{01} = number of attributes that **x** is 0 and **y** is 1;

f_{11} = number of attributes that **x** is 1 and **y** is 1;

Based on the definitions of a 2×2 contingency table, f_{11} is the number of *true positives (hits)*, f_{10} and f_{01} are the number of *misses* and *false alarms*, and f_{00} is the number of *true negatives*. Since Jaccard is a symmetric distance, there is no difference between f_{10} and f_{01}, and they can be used interchangeably. The Jaccard distance is defined as:

$$J = 1 - \frac{f_{11}}{f_{01} + f_{10} + f_{11}} \tag{2.3}$$

Jaccard distance will have a value between zero and one; a smaller value of Jaccard

Table 2.3: Results of comparing Jaccard Distance between non-snow covered and thaw samples (group 1) with snow covered and frozen samples (group 2)

Sensor	No. of Samples in Group 1	No. of Samples in Group 2	Mean and Standard Dev. of Jaccard in Group1	Mean and Standard Dev. of Jaccard in Group 2
SSM/I-13	277	19	0.8217±0.1713	0.9536±0.0733
SSM/I-14	255	16	0.8414±0.1555	0.9539±0.0781
SSM/I-15	163	16	0.8441±0.1639	0.9524±0.0851

distance shows that the two images are more similar.

Appendix A relates the Jaccard distance to FAR and POD. Both the FAR and POD also depend on various combinations of f_{01} and f_{10} errors. But whereas neither POD and FAR contain both f_{01} and f_{10} errors, neither is a complete measure of all error types in itself. The Jaccard distance conveniently combines the two error types and the information captured by POD and FAR into a single and parsimonious metric. The Appendix A contains the relationship between the metrics and Figure A-1 shows that the Jaccard distance is an equitable combination of both FAR and POD. The Appendix A also includes simple case examples where an error type is missed by either FAR or POD but has impacts on the magnitude of the Jaccard distance.

First, the result of comparison between two group of samples as mentioned in Section 2.2.3 is presented. These samples are only taken from region 1 in Figure 2-2 and are only from months of October and November. Group 1 has only samples that are free of snow and have a thaw condition at all times; group 2 has only samples that have snow cover on the ground or a frozen condition at the time of measurement. Table 2.3 presents the results for each of the 3 platforms along with the number of samples. This shows that there is a considerable difference between the mean of the two groups, and group 2 has a higher mean. This proves the necessity for filtering out the snow covered and frozen samples to provide a better analysis of the retrieval errors.

I apply the metric to compare coincident SSM/I and NEXRAD-IV precipitation intensity data products. In order to gain more concise insights, I have evaluated the Jaccard

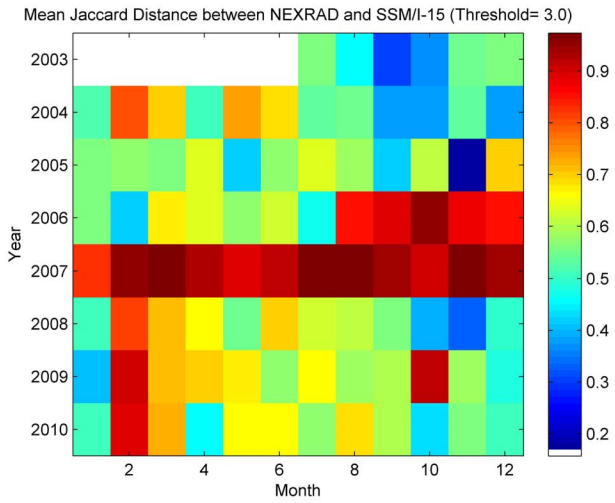
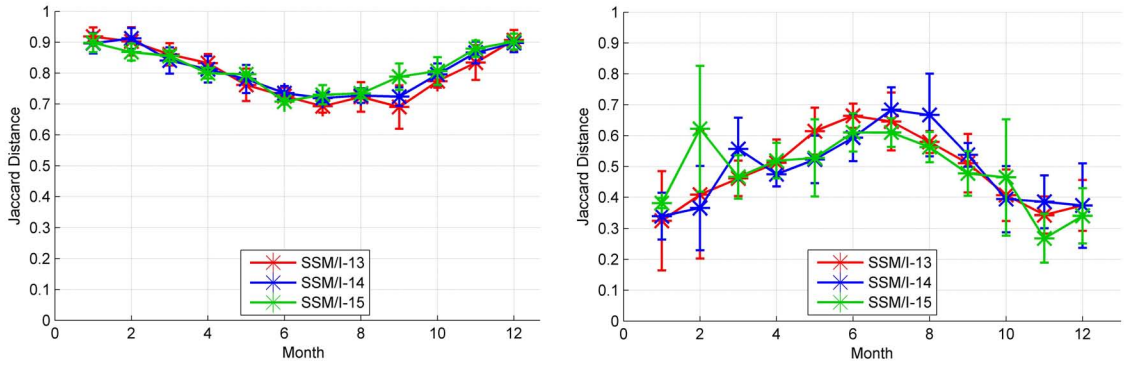


Figure 2-4: Mean of Jaccard distance between pairs of NEXRAD-IV and SSM/I-15 precipitation data products in each month at 3 mm/hr threshold (white colors show months with no data)

distance at different thresholds of precipitation intensity (0, 0.5, 1.0, 1.5, 3.0, 5.0, 7.0, 10.0 mm/hr) in different seasons. Figure 2-4 shows the mean of the Jaccard distance in each month for threshold of 3 mm/hr from SSM/I-15 measurements. The general pattern in this figure is also present for other thresholds, albeit at different magnitudes (see analysis and figures discussed further below). Figure 2-4 shows that there is greater dissimilarity (larger Jaccard distance and closer to unity) for the period between August 2006 and December 2007. This is more distinguished at lower magnitude thresholds. This shows the effect of the radar calibration suite (RADCAL) that was activated on DMSP F15 during August 2006 that produced interference with the 85 GHz channel on the unit. This issue is described in the technical report by Remote Sensing Systems (RSS) Group [79]. This pattern disappears after 2007 and the only reason that can be found for that is the drift in the orbit of the DMSP F15 in early 2008. It has been reported that the effect of the interference from the RADCAL suit is dependent on the thermal environment that changes with the drift of the satellite [78]. However, based on the observations of [78] the interferences exist with an intermittent pattern after 2007. Therefore, in the remainder of this chapter, all the statistics that are presented for SSM/I-15 dataset will exclude the measurements after August 2006 unless otherwise indicated. Comparing the results of Jaccard Distance for the pairs



(a) Threshold = 0.5 mm/hr

(b) Threshold = 5 mm/hr

Figure 2-5: Mean and one standard deviation of Jaccard distance between NEXRAD-IV and different satellite platform precipitation data products in each month

of NEXRAD-IV and SSM/I-13 and SSM/I-14, no systematic anomalous patterns can be detected.

More detailed examination of the Jaccard distance between the NEXRAD-IV and SSM/I estimates reveal that in low thresholds during the summer the SSM/I-based precipitation estimates have better performance (low Jaccard values). During winter the SSM/I land precipitation retrieval algorithm appears to have larger errors. This pattern is reversed at higher thresholds. Figure 2-5 shows the monthly mean of Jaccard Distance at two different thresholds for all the three instruments. There are clearly large differences between panels (a) and (b) of this figure. The seasonality error magnitudes are convex up and down depending on the precipitation rate magnitude threshold.

Figure 2-6 also illustrates the mean and one standard deviation of Jaccard distance over all the samples at different thresholds. There is a decreasing trend of the Jaccard distance with respect to the threshold. The retrieval algorithms have greater challenges with low precipitation rates than higher precipitation rates. As it was mentioned in Section 2.2.2, the Ferraro and Marks algorithm also has a 0.45 mm/hr minimum limit that is also contributing to this problem. Depending on the application, this difference in errors may be compensated by the application itself (e.g. flood forecasts).

The last evaluation using Jaccard distance is presented to show the effect of diurnal cycle on the quality of measurements. Figure 2-7 shows the box plot of Jaccard distance

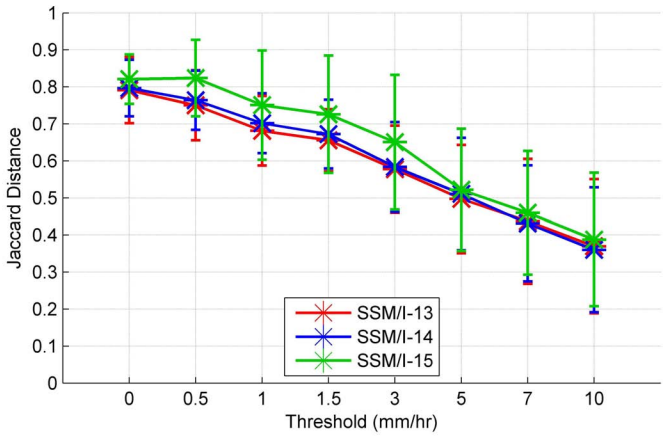


Figure 2-6: Mean and one standard deviation of Jaccard distance between NEXRAD-IV and different satellite platforms precipitation data products at different thresholds

at threshold 0 mm/hr for each of the sensors as a function of time of the day (local time). These plots are derived using samples that are only from regions 2 and 5 in Figure 2-2, as they have a more similar longitude compared to the whole study region. This figure shows that there is not a general pattern for quality of the measurements with respect to the time of the measurement. It should be noted that the reason for the box plot of SSM/I-14 to be spread over several hours compared to the SSM/I-13 is the drift of the satellite during the year 2003-2008. This is evident to some extent in the SSM/I-15 plot as well. Evaluation of the Jaccard distance at higher thresholds revealed similar trends for all the platforms.

2.3.3 Probability of Detection and False Alarm Ratio

The Jaccard distance is a metric of proximity in two binary fields. It is related to both POD and FAR. POD and FAR each capture a separate and distinct type of error. Some applications may be sensitive to one but not the other. For the sake of completeness, I also report on the dependencies of errors captured by POD and FAR as a function of season.

The POD metric shows the quality of capturing precipitation in the estimates. The FAR captures possible over-estimation of precipitating areas. Ebert, 2007 provides a comprehensive review of the statistics [53]. Here, I use POD and FAR to evaluate the SSM/I precipitation product with respect to NEXRAD-IV product

Figure 2-8 show the mean and one standard deviation of POD and FAR for all the three

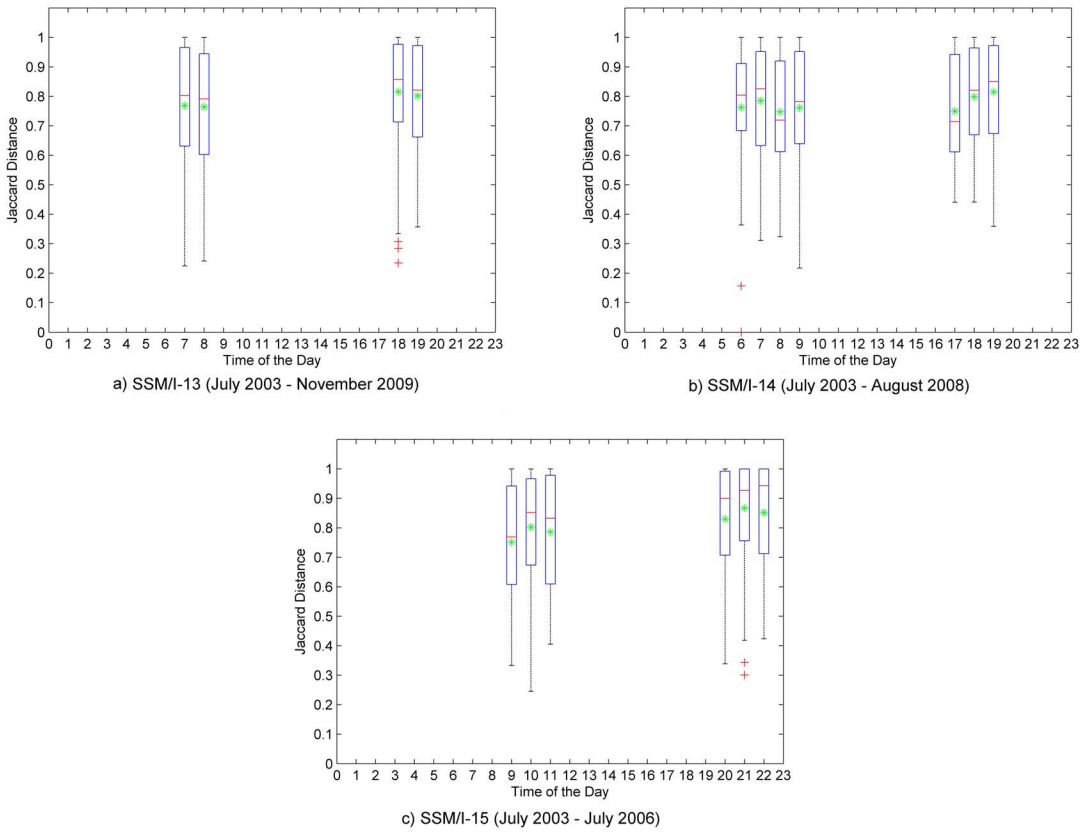
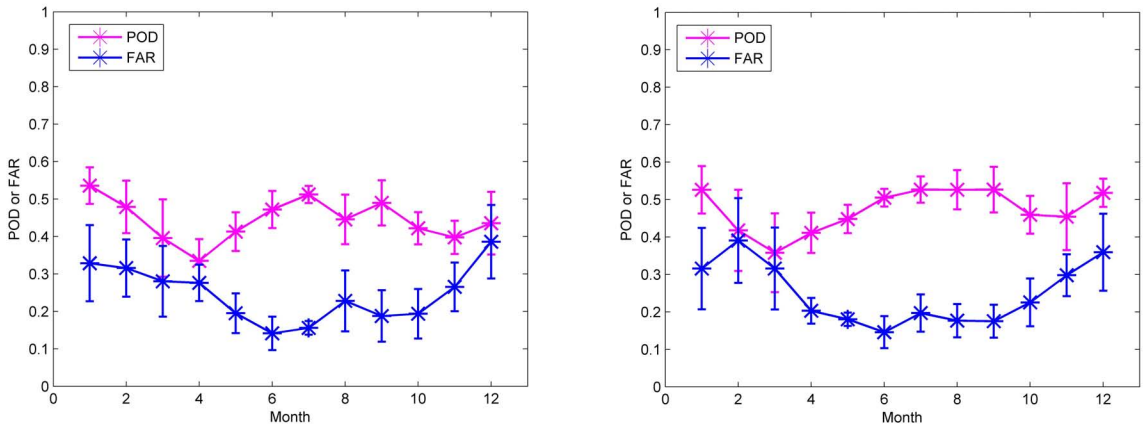


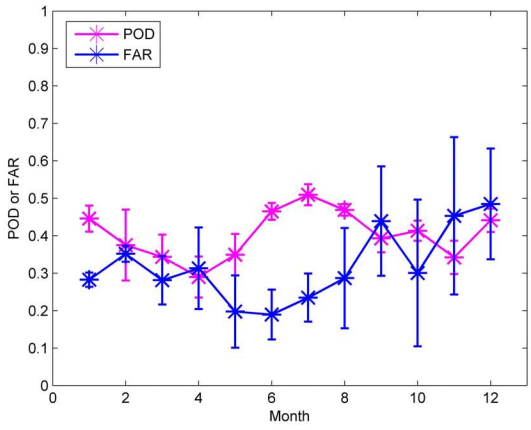
Figure 2-7: Boxplot of Jaccard distance as a function of the time of the day (local time) for (a) SSM/I-13 (b) SSM/I-14 and (c) SSM/I-15. The central red line indicates the median, the edges of the box are 25th and 75th percentiles, the whiskers show the most extreme values and outliers are plotted as red crosses. The mean is also plotted as green star.

platforms across different months. For SSM/I-15 I have only included data from July 2003 until July 2006. The general trend for all the platforms is that during summer months POD increases and FAR decreases, i.e. better estimates are obtained during summer time. This pattern is consistent with the Jaccard distance at lower thresholds; however, at higher thresholds the Jaccard distance had a different pattern that provides another perspective. The pattern revealed by Jaccard distance at different thresholds (Figure 2-5) indicates that moderate to high precipitation rates are estimated more accurately in this algorithm. The fact that this algorithm does not have a classification scheme to distinguish between different precipitation regimes (e.g. stratiform vs. convective) can be a contributing factor to the different performance patterns at different thresholds.



(a) SSM/I-13

(b) SSM/I-14



(c) SSM/I-15

Figure 2-8: Mean and one standard deviation of POD and FAR for SSM/I precipitation data products over different months

2.3.4 Geographical Distribution of Errors

Whereas PDF, Jaccard distance, POD and FAR metrics allow stratification of errors by season, intensity magnitude and instrument, they do not allow investigation of possible systematic geographical errors that may be associated with topography, climate regime and other site-specific contributions to errors in retrieval. In this section, the results of climatological comparison of SSM/I and NEXRAD-IV is presented by mapping mean values. For this comparison, concurrent precipitation estimates based on SSM/I and NEXRAD-IV measurements over the study region are averaged, and the mean of the precipitation rates are presented in Figure 2-9. I have only selected the concurrent measurements for each platform; therefore, the two climatological averages are statistically comparable.

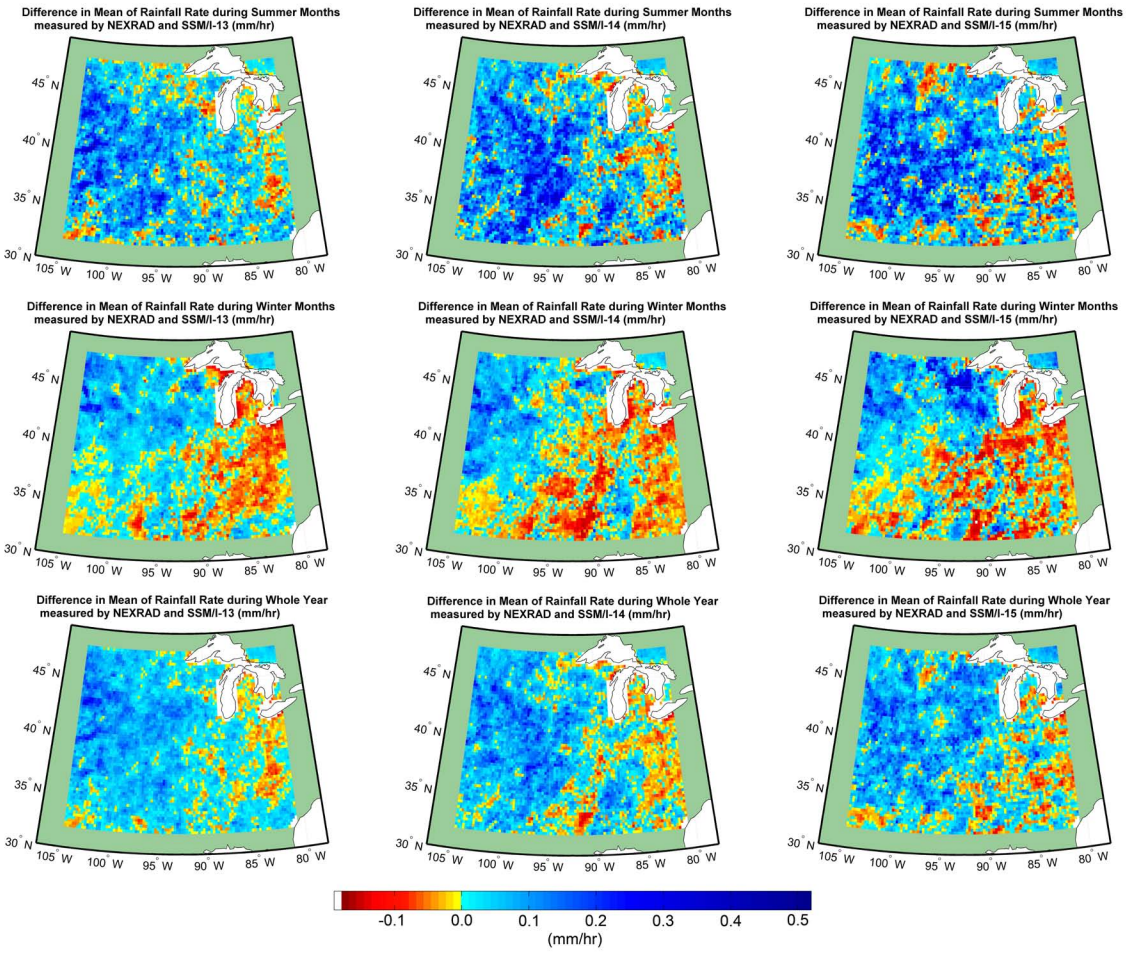


Figure 2-9: Difference in mean of SSM/I and NEXRAD-IV precipitation data products. Left column is for SSM/I-13, middle columns is for SSM/I-14 and right columns is for SSM/I-15. First row shows Summer months (April to September - inclusive), second row shows Winter months (October to March - inclusive), and third row shows the Whole-year

Rows in Figure 2-9 from top to bottom correspond to summer differences, winter differences and whole-year differences in the SSM/I instrument-based precipitation products and NEXRAD-IV. Summer is defined as April to September (inclusive) and winter is defined as October to March (inclusive). For SSM/I-15 only the period not affected by RADCAL is used (July 2003 to July 2006).

A useful backdrop for interpreting the difference maps is the precipitation climatology itself - sampled according to the availability of satellite-based estimates. Figure 2-10 shows the precipitation climatology evident in the NEXRAD-IV data product. In deriving this climatology, I have sampled NEXRAD-IV at the times when SSM/I-14 swath and overpass

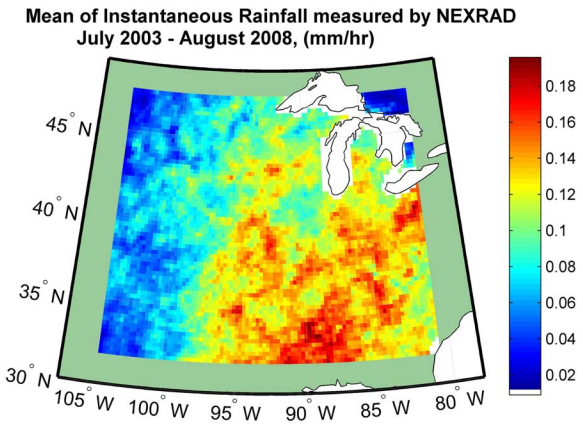


Figure 2-10: Mean of precipitation rate based on NEXRAD-IV sampled concurrent to SSM/I measurements over the region

coverage is available. This conditional sampling has small and almost imperceptible effects on the true climatology when all NEXRAD-IV data are used (not shown). Furthermore, the choice of instrument other than SSM/I-14 does not change the conclusions. The issues are more closely tied with the algorithms themselves.

Figure 2-10 shows a strong a gradient of mean precipitation over the region from north-west to south-east. Superposed on the difference maps in Figure 2-9, a systematic error with geographical structure becomes evident. There are three major differences between the NEXRAD-IV estimates and the SSM/I precipitation data products:

- SSM/I-based precipitation products mostly overestimate the precipitation rate; therefore, the difference in their mean statistic is mostly positive. There are regions that SSM/I underestimates the precipitation rate climatologically, but their area is limited in extent when compared to those that are overestimated.

- All the three platforms and their instrument and data products have similar spatial patterns for summer, winter and whole-year. For example during summer SSM/I-13, SSM/I-14 and SSM/I-15 show similar differences in mean precipitation over the study region. During winter all the three SSM/I products overestimate over the western and north-western parts of the study region. This is a dominant feature in all three and it affects the whole-year difference plots. The dominant differences are localized in the western and north-western part of the region.

- In the southeastern part of the region, all the three platforms have small negative difference values. This region has the highest precipitation rates climatologically. The SSM/I-based precipitation estimates have better quality over the south-eastern part of the study region that generally has higher precipitation rate magnitudes. This is consistent with the findings from Jaccard distance at different thresholds (Figure 2-5) that this algorithm has better performance on moderate to high precipitation rates.

These conclusions are consistent with those evident in the probability density functions (Figure 2-3).

2.3.5 Error Dependence on Surface Soil Moisture

One of the factors affecting the quality of precipitation retrieval from microwave measurements is variable land surface emissivity that is partially due to dynamic surface soil moisture [61]. In-situ or satellite estimates of surface soil moisture are limited by their spatial and/or temporal coverage. Therefore, I used the Antecedent Precipitation Index (API) as an index of soil moisture content to investigate correspondence between the precipitation estimation error and estimated soil moisture. API is defined as:

$$API_t = K \times API_{t-1} + P_t \tag{2.4}$$

In which API_t is the API on day t, API_{t-1} is the API on the preceding day $t-1$, P_t is the precipitation on day t, and K is a constant decay factor. K is dimensionless and the other three parameters in this equation have units of length (in this case *mm*). The value of K controls the decay or loss rate of surface soil moisture. Based on literature values of the application of this filter to precipitation in order to produce estimates of surface soil moisture, I use a typical value of 0.93 for K. Daily API are estimated using the 24hr accumulations of NEXRAD measurements.

For each measurement pixel I calculate the error of the SSM/I estimation with respect to NEXRAD. I segment the errors conditional on values of API over that pixel from the day before. The API of day of observation is affected by the current storm; however, the API from the day before only indicates the soil moisture content as a result of preceding storms

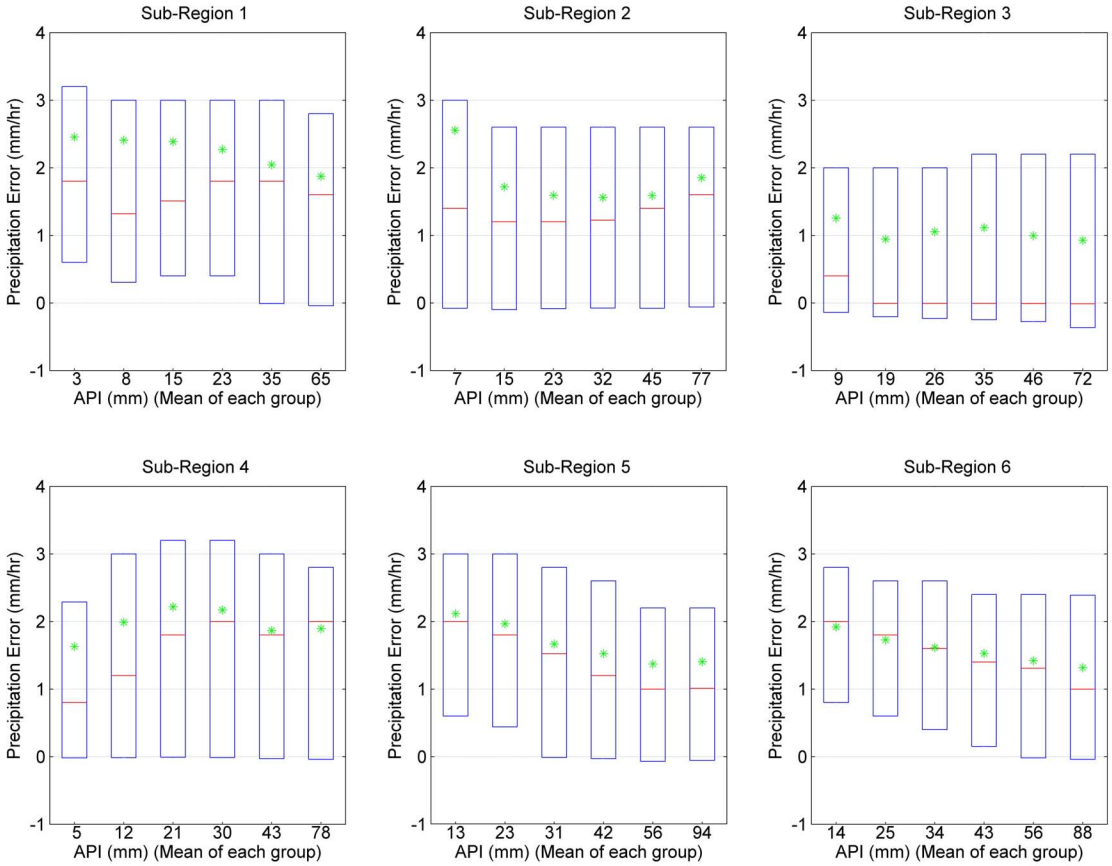


Figure 2-11: Boxplot of SSM/I-based precipitation rate error as a function of API. The central red line indicates the median, the edges of the box are 25th and 75th percentiles, and the mean is plotted as a green star.

and at the time of measurement. Use of same-day API did not appreciably change the conclusions. In order to enhance the sample size, the measurement errors from SSM/I-13, 14 and 15 are combined.

Figure 2-11 shows box plots of error in SSM/I precipitation estimation versus API values. This figure includes six plots, one for each of the sub-regions. In each sub-region the numbers of data for each boxes are equal. Number of data in each box of different sub-regions are: sub-region 1: 23,635; sub-region 2: 27,355; sub-region 3: 16,653; sub-region 4: 12,856; sub-region 5: 23,632; sub-region 6: 30,457. Similar to the other box plot in Figure 2-7, the edges of the boxes are 25th and 75th percentiles, and the central red line indicates the median. The mean is also plotted as green star. No consistent trend is

detected in sub-regions 1 through 4; however, in sub-regions 5 and 6 the mean and median of the error decreases with increase in API. The results of a statistical hypothesis test (t-test) shows that in 86% of the cases the mean of each of the error box plots is different than the others in the same sub-region at 5% significance level.

This analysis shows that the error in SSM/I precipitation estimations are dependent on dynamical changes of the surface conditions. Land surface emissivity changes that are partially driven by dynamics of surface soil moisture can affect the precipitation rate retrieval performances. A possible future path is to use surface soil moisture retrievals from the Soil Moisture Active Passive (SMAP) mission to estimate the soil moisture contribution to changing surface emissivity [61].

2.4 Summary

In this chapter, I investigate and diagnose the errors in the precipitation products based on the Special Sensor Microwave/Imager (SSM/I) instrument on-board the DMSP F13, F14 and F15 platforms. The gauge-corrected ground-based radar product NEXRAD-IV is used to diagnose the dependencies of the errors on precipitation intensity, season, platform and other factors. The stratification of the errors according to these attributes is made possible by using extended duration and large domain comparisons data sets. Data products from three DMSP platforms and the extensive NEXRAD record are used. The stratification identifies priorities for ways of improving the passive microwave-based precipitation retrieval algorithms over land. These algorithms are currently used to calibrate global precipitation products that use infrared brightness temperatures from instruments on geostationary platforms as their major inputs.

I show that current SSM/I-based precipitation products are positively biased in magnitude. The bias is less at higher intensities and in geographic locations where precipitation rates are generally higher. Several marginal distribution and magnitude and phase error statistics are also used in this study to evaluate SSM/I-based precipitation products. The Jaccard distance is introduced as a compact and concise metric to evaluate precipitation products. It is shown that the metric captures errors that have distinct seasonal and

magnitude-dependent characteristics. Together with its constituent POD and FAR metrics, I show that the SSM/I-based precipitation products generally overestimate precipitating areas during the winter months by estimating low magnitude precipitation rates outside of the precipitation clusters. This contributes to the noticeable dependence of the Jaccard distance metric on intensity magnitude. The SSM/I-based products provide better estimates during summer time when compared to the winter season.

Month by month comparisons of the metric across platforms and instruments reveals a major anomaly starting in August 2006 and ending in December 2007. During this period the products based on SSM/I-15 are worse compared to the rest of the study period. This is consistent with the timing of the activation of the radar calibration (RADCAL) suite on board F15 that interferes with 85 GHz channel.

The effects of dynamic surface soil moisture on the quality of precipitation retrieval shows strong connection between the estimation error and changing soil moisture. More accurate land surface emissivity estimates can possibly improve precipitation rate retrievals over land. The future SMAP mission provides accurate fine spatial scale measurements of soil moisture that can be combined with precipitation retrieval algorithms to improve accuracy of precipitation estimations.

There are inherent challenges associated with retrieval of precipitation over land based on the measurements in the set of channels afforded by SSM/I and its heritage instruments. The sources of emissions and scattering cannot all be characterized by the finite measurements. Simple distributional corrections such as probability matching work well on the training data sets, but may not generally extend to times and regions where both ground validation and retrievals are not used in the matching. I suggest that the paradigm for generating and using data products needs to be revisited in the case of such a challenging and important retrieval problem. The error patterns in SSM/I-based precipitation products are not structured in simple ways, and they cannot be easily modeled and removed. Perhaps, we should go beyond deterministic products generation and incorporate uncertainty in the description of the products. There is an emerging trend towards stochastic and ensemble-based approaches to merging and interpreting data products. If we generate a population of equally probable prior replicates based on the SSM/I measurements and update the prior

probabilities using the historical error likelihood; then, it is possible to generate a posterior population of precipitation replicates that are ranked based on the probability of being a characterization of the true situation. Each of these replicates will be consistent with the original measurement while having some added noise. This study aimed at characterizing the types and dependency attributes of retrieval errors that is a necessary first-step toward this goal.

Chapter 3

Generating Spatially Distributed Precipitation Replicates

Abstract

Current satellite observations allow retrieval of precipitation fields but their estimates of location, magnitude and phase of precipitation remains highly uncertain. Therefore, quantitative representation of precipitation retrieval uncertainties is required for widespread adoption of remotely-sensed precipitation products. An elegant way to express these uncertainties is to generate a realistic ensemble of precipitation replicates that differ while sharing basic structural similarities. Available methods for simulating precipitation include physically-based and stochastic models. Physical models can be very expensive to use in an ensemble context and most stochastic models have difficulty reproducing realistic, i.e. conditioned on observations, spatially discontinuous (or patchy) precipitation fields. In this chapter, I present a new method for generating realistic spatially discontinuous precipitation replicates. The method is based on an orientation and scale-dependent spatial filter that describes precipitation images in terms of a small number of random coefficients. Different sets of coefficient values give different spatially distributed precipitation replicates. The statistics of the random coefficients are derived from a training image that defines the common spatial structure of the ensemble members. The implementation discussed here generates static images (snapshots), although it can be extended to generate dynamic rain

storms if the training image is time-dependent. This precipitation generation algorithm is illustrated with an example using ground-based radar data and geostationary satellite cloud-top temperature observations. Results from this example show that the proposed method is able to generate realistic replicates that reflect the general structure of the training image while also conveying sufficient variability to properly represent uncertainty.

3.1 Introduction

Satellite monitoring of precipitation has the advantage of almost global areal coverage compared with surface-based measurements such as gauge and radar measurements. Different remote observations of precipitation (passive microwave, radar and infrared) have distinctive advantages as well as shortcomings. Numerous studies have been carried out to assess the uncertainties in the observations of satellites used to observe precipitation; however, most of them have been focused on characterizing retrieval or rain rate magnitude errors ([165, 62, 140], among others). Some have tried to decompose these errors (in single-satellite retrievals or multi-satellite products) by adopting assumptions such as parametric distribution for error or a multiplicative/additive error model ([80, 81, 172, 181, 150, 3, 156, 169], among others). Development of improved quantitative representations of retrieval uncertainty is a prerequisite to the widespread adoption of satellite precipitation products [25]. Therefore, there is a need to investigate the uncertainties associated with the satellite precipitation retrievals in a more comprehensive framework that includes the spatial geometry of the rainy area as well as the statistical distribution of precipitation intensity. An elegant way to express these uncertainties is to generate a realistic ensemble of precipitation replicates that differ while sharing basic structural similarities.

Ensemble approaches approximate the full probability distribution of interest with a set of equally likely realizations from the true distribution. They are useful for propagating the uncertainty in hydrological and meteorological models, especially those used in data assimilation [15, 128]. They are also helpful in probabilistic flood and drought forecasting and risk analysis. Moreover, an ensemble of spatially distributed precipitation events that captures the correct precipitation uncertainty can be used as a random forcing in a spa-

tially distributed land surface model to derive probability distributions for all major surface hydrological variables. Precipitation forcing is the principal source of ensemble spread in now-popular ensemble data assimilation systems used in hydrology. Without adequate characterization of precipitation uncertainty, the ensemble tends to collapse and new observations get rejected in the analysis due to an over-confident model forecast.

There are two main approaches for generating precipitation ensembles: physically-based models and stochastic models. The first approach is to generate precipitation ensembles using the physically-based numerical weather forecasting models by perturbing the initial or boundary conditions. Despite continual increases in numerical model resolution and significant improvements in the forecasting of many meteorological parameters, the results of these models still suffer from deficiencies in the Bulk Microphysical Parametrization (BMP) schemes, insufficient knowledge of precipitation processes, and inadequate terrain representation among other. These deficiencies have become more apparent as models' resolution have increased [42, 155, 162, 19, 160, 117]. At the same time, running these models for hundreds of different initial or boundary conditions to generate an ensemble of precipitation events is computationally inefficient and impractical.

An alternate approach is to use stochastic models that simulate the observed spatio-temporal patterns of precipitation. Examples include multifractal models and scaling laws [119, 121, 131, 142], multiplicative cascade models [73, 72], wavelet models [106] and geostatistical models [143, 167, 187, 110, 154]. These models either have unrealistic mathematical assumptions about stationarity and homogeneity of the spatial precipitation field or dependence on first-order and second-order statistics (such as variogram) which makes it hard to represent the intermittent characteristics of precipitation fields in the ensemble replicates. Wojcik et al., 2009 present a method for generating spatially and temporally intermittent precipitation replicates which are conditioned on geostationary satellite measurements [187]. Their method is based on a multipoint geostatistical procedure that uses the probabilities of precipitation patterns within a specified template to generate the rainy areas for individual replicates. Then, the precipitation intensity distribution within each rainy area is generated using a multiplicative cascade model. Although this method produces realistic replicates, the generated precipitation fields are stationary in space and the

method is not able to condition (or constrain) the replicates on point measurements of precipitation (e.g. in-situ measurements).

In a recent study, Paschalis et al. develop a new stochastic spatio-temporal precipitation simulation model [141]. This model, called Space-Time Realizations of Areal Precipitation (STREAP), uses a three-stage hierarchical structure to reproduce features of precipitation in space and time. The three stages are: estimating storm arrival time, evolving areal mean precipitation intensity in time, and evolving the spatial structure of the storm in time. However, the model has two main assumptions that makes it hard to be applicable to diverse climatic conditions. STREAP assumes that the spatial probability for precipitation occurrence and accumulation is homogeneous over the study domain, and the simulated precipitation fields are isotropic.

In this chapter, I propose an alternative stochastic method to produce an ensemble of spatially distributed precipitation storms that can be constrained by satellite and/or point measurements. The precipitation replicates are generated using a spatial filter developed in [63]. This so-called Steerable Filter decomposes an image to a set of subband coefficients that can be used to control the spatial structure of the generated replicates. In this method, the steerable filter coefficients are random variables that take on different values for every replicate. The statistical properties of the filter coefficients are inferred from a training image that defines the dominant spatial features (area, orientation, patchiness, intensity distribution, etc.) shared by all members of the precipitation ensemble. The resulting random precipitation snapshots are spatially discontinuous (intermittent in space) and non-stationary. Here, the focus is on generation of static precipitation images (snapshots) since the primary emphasis is on methods for producing spatially discontinuous rain storms with realistic intensity distributions. The basic concepts can be used to generate dynamic precipitation fields if the training image and steerable filter coefficients are allowed to vary over time.

The rest of this chapter is organized as follows. Section 3.2 outlines the data sources. Section 3.3 describes the steerable filter and the proposed replicate generation methodology in detail. Section 3.4 presents the results followed by their evaluation in Section 3.5. Section 3.6 presents summary and conclusions.

3.2 Datasets

In this chapter, ground-based radar estimates of precipitation are used as inputs to the replicate generation method. The study region is over central part of the Continental US (CONUS), and the measurements are from Next Generation Weather Radar (NEXRAD) data set obtained from the National Weather Service's ground-based WSR-88D radar network [64]. The hourly estimates produced by 12 River Forecast Centers (RFCs) in CONUS are mosaiced onto a 4km grid by the National Centers for Environmental Prediction (NCEP) into a national product (NEXRAD-IV) that is available through the National Center for Atmospheric Research (NCAR) Earth Observing Laboratory (EOL).

All the training images used in this study for replicate generation are selected from storms that occurred in Summer 2004 over CONUS. The set of training images captures different storm shapes and orientations as well as various ranges of rain rates. More details about the training images are presented in Section 3.4.

This method is capable of statistically constraining the precipitation replicates to cloudy areas. These ares are identified using the Geostationary Operational Environmental Satellites (GOES) data set produced by the U.S. National Oceanic and Atmospheric Administration (NOAA). The GOES infrared imager measures the cloud-top temperature, which is used to identify the clouds that are most likely to produce rain. Because rain-producing deep convective systems and non-rain-producing cirrus high-altitude clouds both have cold upwelling radiation, cloud-top temperature estimates from geostationary satellites visible/infrared instruments are not sufficient to quantify rain-rate magnitudes.There are different methods to classify the clouds to rainy and non-rainy [80, 168, 187, 20, 55]. In this study, I take a similar approach to [187]. Their method is based on thresholding the cloud-top temperatures and calculating the fraction of rainy areas that fall inside and outside of the cloudy area. Using NEXRAD data over central part of US, they found that a threshold of 258 K will result in a low probability of 0.01 for rain to occur outside the cloudy region. Since in this study we are using the precipitation storms over the same spatial and temporal domain, I use the same threshold to identify cloudy areas from GOES observations.

3.3 Methodology

The precipitation replicate generation method has 2 steps:

1. Scan a selected Training Image (TI) and calculate relevant spatial and intensity statistics;

2. Generate a randomized precipitation replicate that conforms to the TI statistics.

In each of the two steps, the corresponding TI or replicate is decomposed into several subbands using steerable filters. Further details of the steerable filters and each one of the generation steps are described in the following together with a description of the procedure used to constrain precipitation to specified areas or to reproduce specified point values.

3.3.1 Steerable Filters

Steerable filter, equivalently, steerable pyramid, is an image decomposition technique that is linear, multi-scale, multi-orientation, and self-inverting (meaning that the filters used to build the pyramid representation are the same as those used for reconstruction) [63]. The decomposition method uses spatial filters of arbitrary orientation that are synthesized from a linear combination of a set of basis filters. Since the steerable filter's basis functions are derived by dilation, translation, and rotation of a single function it can be considered to be a wavelet transform.

Steerable filter decomposes an image to a set of coefficients, or equivalently images, at different orientations and scales. Each of the set of coefficients at a specific scale and orientation is called a subband. Figure 3-1 illustrates an example of a one-scale image decomposition with four orientations. First, the image is filtered using a high-pass filter (called H_0). Then, the image is filtered using a low-pass filter (called L_0), and the output of the low-pass filter is further filtered using oriented filters. Figure 3-1 shows the output (equivalently subbands) of the high-pass and oriented filters at scale 1. Each of the four oriented subbands reveal different features of the image.

The second level of the steerable pyramid decomposition is created by first applying a low-pass filter (L_1) to the output of the low-pass filter L_0; then, down sampling the output

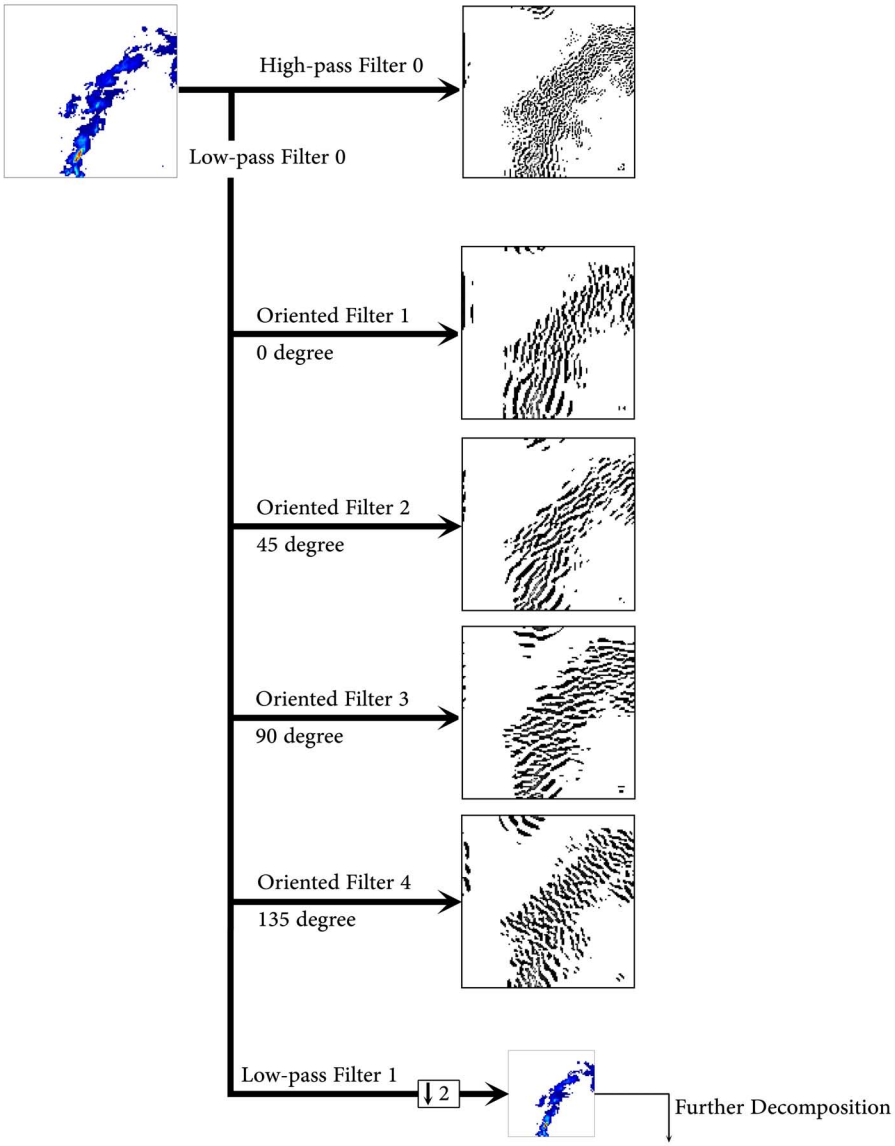


Figure 3-1: Steerable Pyramid decomposition of a sample precipitation image. The top panel is the high-pass band and lower panels show 4 sub-bands at level 1 of the pyramid, and the low-pass band that is down-sampled by a factor of 2 for further decomposition at level 2 of the pyramid. The subbands are shown in gray scale to highlight the orientational structure of the coefficients in different subbands.

by a factor of two (this is shown at the bottom of Figure 3-1). This will go through the oriented filters and generate the coefficients at the second level of the pyramid.

For the samples used in this study (images of 256 $\times$ 256 pixels), it was found that a two level pyramid with four orientational decomposition adequately captures precipitation spatial intermittency.

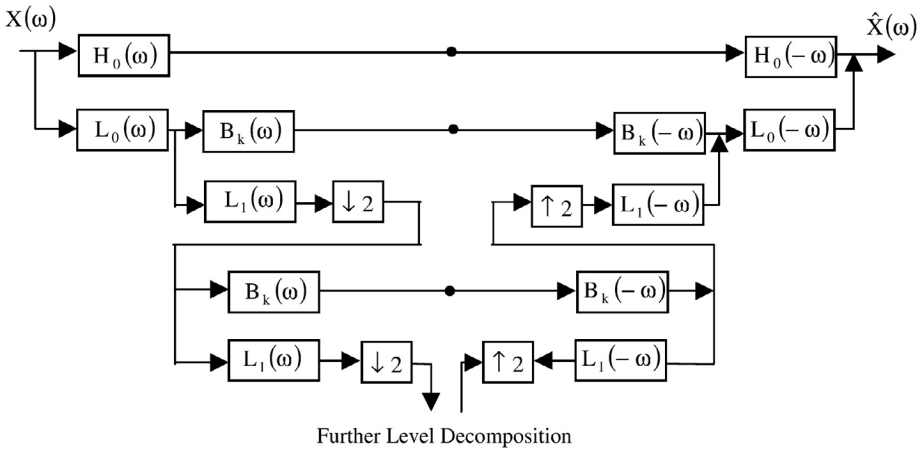


Figure 3-2: Structure of the Steerable Filter in frequency domain (source: Liu *et al.* [118])

3.3.2 Scanning the Training Image

The first step of the method is to scan a given TI to calculate relevant statistics that capture the texture of the precipitation in the TI. The statistics that are used here are based on [145]. They have used these statistics to simulate a wide range of common texture patterns in computer vision. The four statistics are briefly defined in the following. For further details, the reader is encouraged to read [145].

1. **Histogram of Intensities**: The histogram expresses the relative amount of each intensity in the image (at pixel or subband level). Here, three normalized sample moments (variance, skewness and kurtosis), together with the range are used for pixel-level intensities. The skewness and kurtosis are considered for subband-level intensities.

2. **Local Autocorrelation**: The steerable pyramid decomposition is overcomplete; moreover, covariances of subband coefficients can arise from spectral peaks or ridges in a TI. Therefore, the coefficients of this decomposition are typically correlated. Local autocorrelation of the low-pass subbands are measured at each level, and used as a statistics in the generation step.

3. **Coefficients' Magnitude Correlation**: "Features" in an image give rise to large coefficients in local spatial neighborhoods of a subband, as well as at adjacent scales

and orientational subbands. Therefore, the correlation of the magnitude of pairs of coefficients at adjacent positions, orientations and scales is used as the third statistics.

4. **Cross-scale Phase Statistics**: This statistics is used to measure the relative phase of coefficients of subbands at adjacent scales. This constraint will help in distinguishing edges from lines in the TI

3.3.3 Generating Replicates

The generation step is a recursive algorithm that iteratively converges towards a final replicate. First, a Gaussian noise image is generated. In order to be efficient, the variance and mean of the Gaussian generator is set to be the same as the variance and mean of the pixel-level intensities of the TI. Then, the noisy image is decomposed using the steerable filters, and the four subband statistics are matched to the values derived from the TI. Next, the pyramid is collapsed and the replicate is reconstructed. This is the first iteration of the method.

In order to have a more realistic replicate, the output of the first iteration is considered as input to the algorithm (instead of the Gaussian noise), and the whole procedure is implemented again. This is repeated until a realistic replicate is generated. Tests with different precipitation TIs showed that after 25 iterations there is insignificant change in the replicate, and the statistical characteristics of the replicate remain unchanged. Figure 3-3 shows a TI along with the results of different iterations for generating a replicate. After each iteration the replicate becomes more and more similar to the TI (i.e., more realistic).

3.3.4 Constraints

A replicate generation method should be able to limit precipitation storms to specified spatial supports, so that the replicates can be constrained by remote sensing or in-situ measurement data. This method is capable of both areal support and point precipitation constraining, which are described in the following.

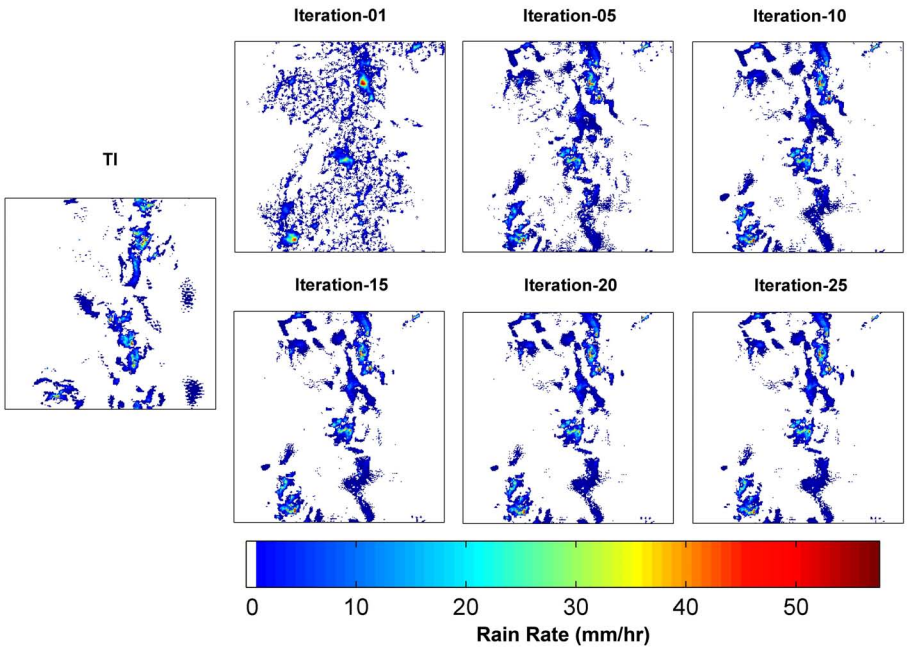


Figure 3-3: Outputs of different iterations in generating a replicate

Areal Support Constraint

GOES observations of cloud top temperature are used to identify potential precipitation areas and constrain the replicates with them (details of the cloud area detection are in Section 3.2). I multiply the initial Gaussian noise image by a binary support image. Figure 3-4 is an illustration of this masking process using geostationary cloud-top temperature observations. In Section 3.4, a comparison between the replicates that are generated constrained on the cloudy areas and the ones that are not constrained is presented.

Point precipitation Constraints

Gauge measurements of precipitation can be fused into this replicate generation method by limiting one of the pixels of the replicate to have the same value as the one measured by the gauge. This feature is implemented by defining a mask for the pixel that includes the gauge. The mask is designed to prevent discontinuity in the precipitation fields. Figure 3-5 shows the 3×3 pixel mask that is used in this method. In Section 3.4 the results of this type of constraining is presented.

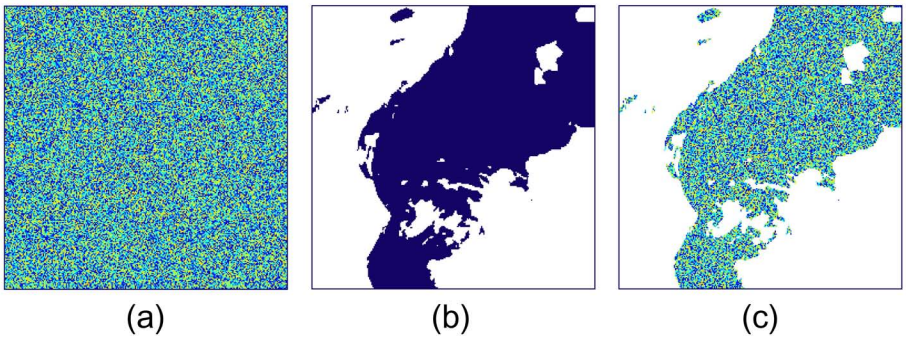


Figure 3-4: Constraining a replicate on cloud support. (a) is a Gaussian noise generated as an initial input for the method. (b) is the cloud support image in which dark areas are considered cloudy with potential precipitation. (c) a constrained input for the replicate generation method. Light green areas in this image have zero intensity.

3.4 Results

One hundred TIs are selected from the NEXRAD-IV database explained in Section 3.2. Figure 3-6 shows 16 of these TIs. Each of them is a 256×256 pixel image of an isolated precipitation storm. These TIs are selected in a way to capture different patterns of precipitation storms; they include both large and small clusters as well as storms orientated in different directions.

Three groups of replicates have been generated for each TI to evaluate the proposed replicate generation method. Each group tests one element of the overall algorithm:

1. Unconstrained replicates;
2. Replicates constrained by areal support, derived from GOES observations;
3. Replicates constrained by both areal support and point precipitation measurements.

0.5	0.5	0.5
0.5	1	0.5
0.5	0.5	0.5

Figure 3-5: Mask used for point precipitation constraint in the replicates

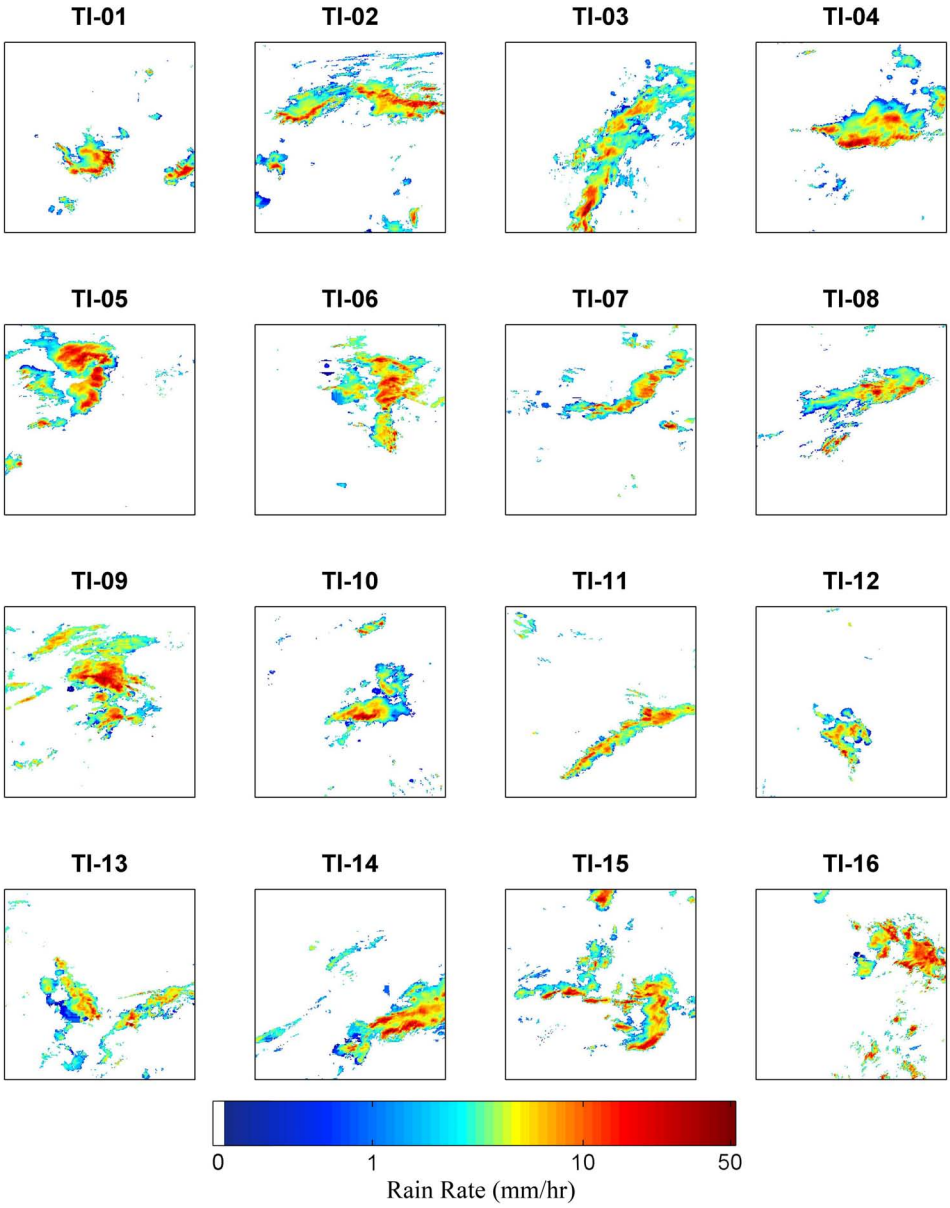


Figure 3-6: Sixteen precipitation Training Images (TI) used in this study

The areal support is from GOES observations and for the point precipitation data I have randomly selected 4% of the rainy pixels in the TI and used them to constrain the replicates. This subset of pixels are meant to represent possible surface gauge precipitation measurements. Alternatively, they can represent the finite swath and overpass coverage of a microwave instrument measurement (and its associated precipitation rate retrieval) that is to be merged with a continuous-in-time geostationary satellite visible/infrared could-top temperature measurement.

For each group of replicates, 300 replicates are generated from each TI, and the results are presented in the following.

3.4.1 Unconstrained replicates

Figure 3-7 shows a sample TI with its associated GOES support as well as 16 randomly selected replicates. These replicates are generated without any constraint on areal support or point precipitation measurement, and the GOES support is shown only for comparison. The replicates show the capability of this method in generating a set of diverse replicates while preserving the structural features of storms. Moreover, the method is generating high intensity rain clusters in the replicates similar to the ones in the TI. Further investigations revealed that these characteristics are preserved across all the replicates. The replicates are, however, diverse and random. They cannot be considered to be constrained by observations.

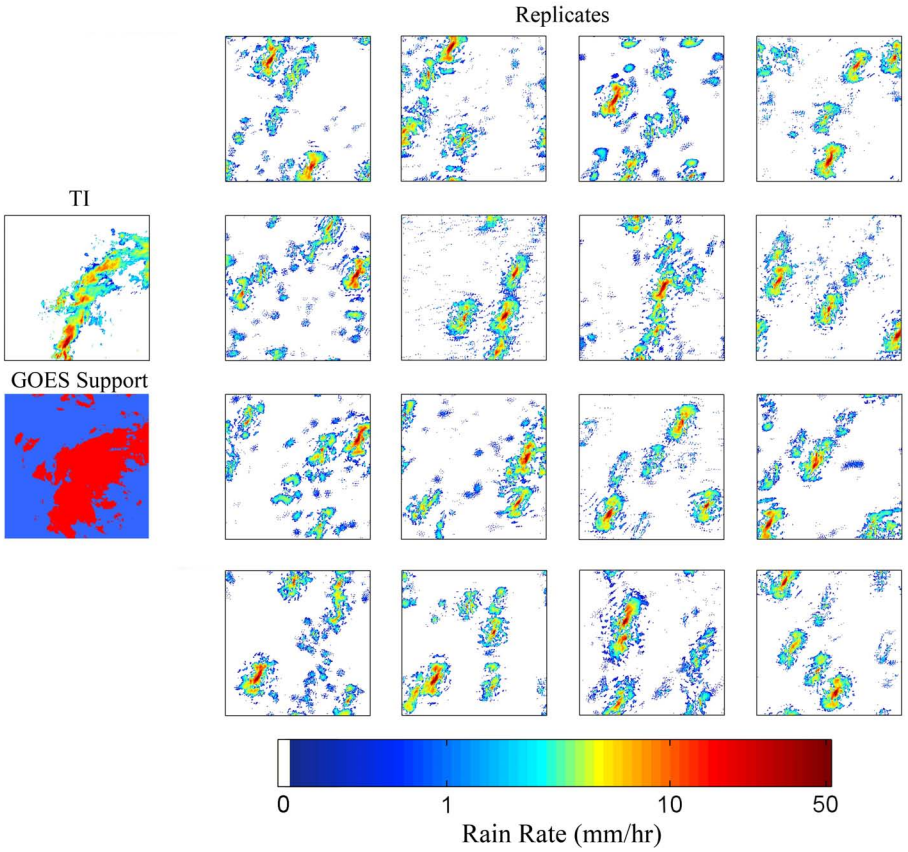


Figure 3-7: A sample TI, its GOES support and 16 selected replicates (no constraint)

3.4.2 Replicates constrained by areal support

Figure 3-8 shows the same TI with another set of replicates which are generated by constraining the replicates to the areal support derived from GOES observation. Comparison between Figure 3-8 and the figure without constraints (Figure 3-7) shows that the position of the rain storms are better captured in the replicates which are constrained by the areal support while containing some randomness. Moreover, the overall diagonal/arc shape of the storm is better preserved.

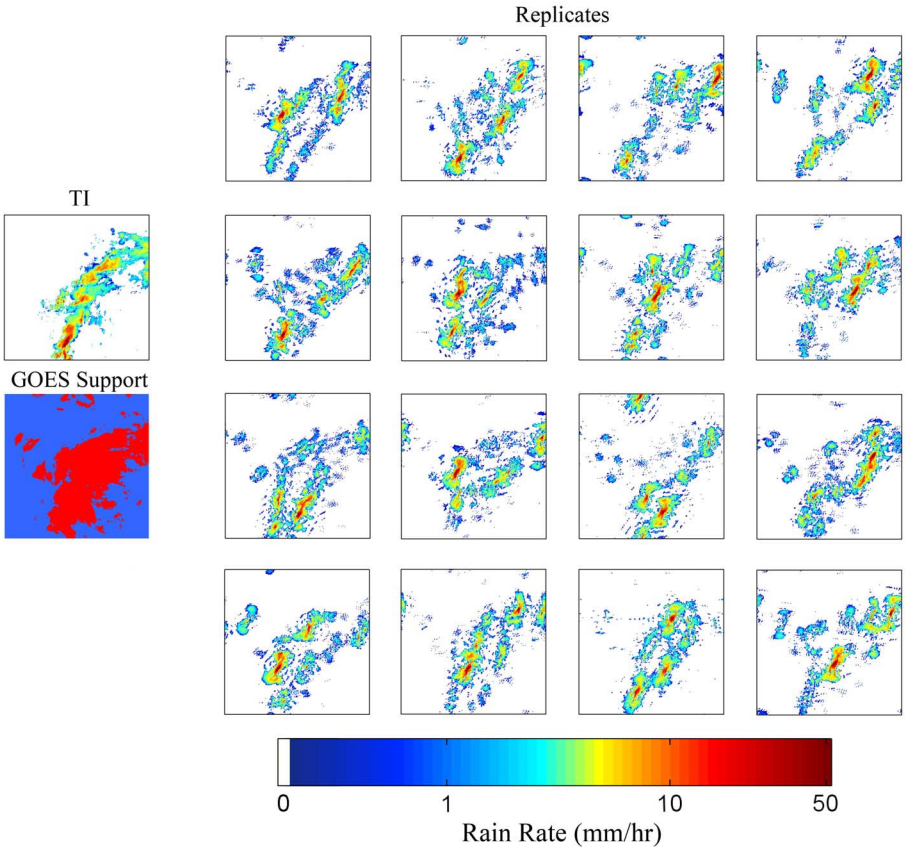


Figure 3-8: A sample TI, its GOES support and 16 selected replicates (areal constraint)

3.4.3 Replicates constrained by both areal support and point precipitation measurements

Figure 3-9 illustrates a new set of replicates from the same TI that are generated by constraining the replicates on both areal support and 4% point precipitation measurements.

The 4% point means that 4% of the rainy pixels have been selected randomly and have been used to constrain the precipitation fields. Moreover, different percentages of points (between 1% up to 5%) were tested for point precipitation measurements, and the resulting replicates have comparable statistical properties.

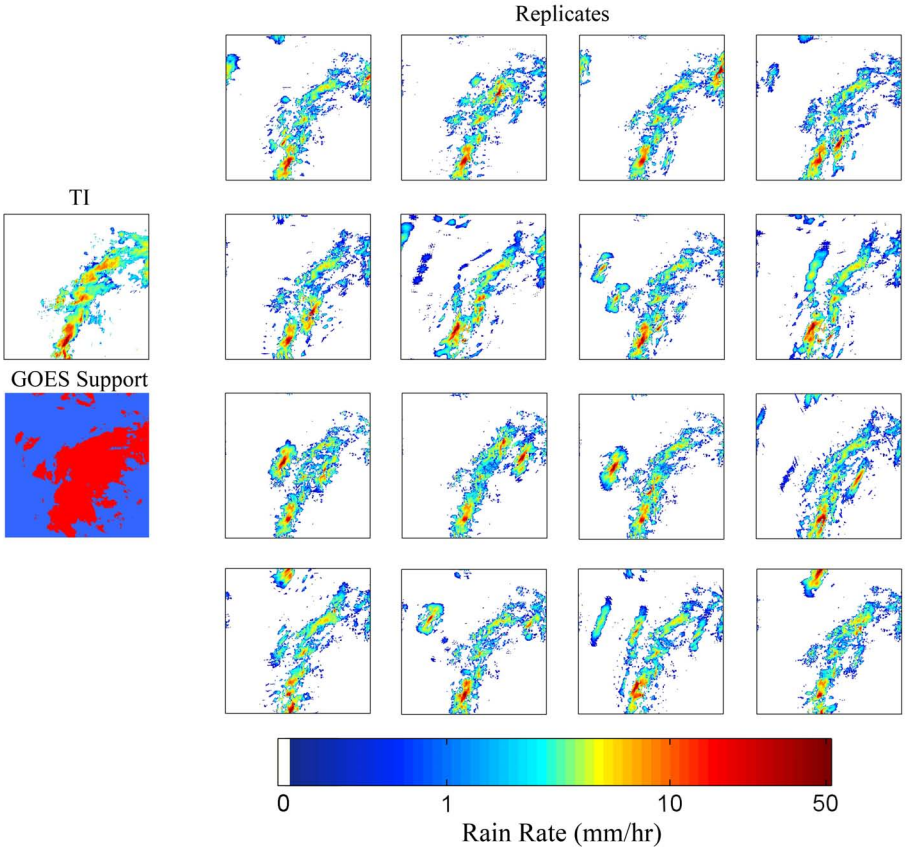


Figure 3-9: A sample TI, its GOES support and 16 selected replicates (areal and point constraints)

Figure 3-9 shows that the replicates are more spatially constrained compared to the second group of replicates, and the structure of the storm is to a high degree the same as the TI in all the replicates. This is reasonable as these replicates have a higher degree of spatial constraining. Although these replicates are more similar to the TI, they still have some variability in the spatial structure and precipitation rate distribution.

Here, I have used randomly selected points from the TI for point precipitation constraining to evaluate the capability of the method; however, in real applications, these points can be pixels that have gauge measurements of precipitation and their value can be derived

from the in-situ measurement.

In general, the results show that it is possible to generate a diverse set of replicates that can capture uncertainties in the given TI. The degree of diversity that should be used depends on the application. For example, in a Bayesian framework it might be necessary to have a wide range of prior population, then a population of replicates with less constraints should be generated. However, it might be useful to include both constrained and non-constrained replicates in the prior population.

3.5 Evaluation

The quality of an ensemble cannot be adequately measured by a single metric, several different attributes of prediction skill need to be used ([137]).

Common skill attributes for evaluating ensemble replicates include the Brier Score (BSS), Ranked Probability Skill Score (RPSS) and Rank Histogram among others ([33]; [136];). However, these skill measures are designed either for scalar values or categorical variables and they can not be applied to evaluate ensemble of spatial replicates of precipitation. Therefore, three alternative indices are considered to evaluate the replicates here: the cluster size distribution, the cumulative distribution function of the precipitation intensity, and the spatial similarity of replicates to the TI. These indices give insight into how well the replicates represent the spatial structure as well as precipitation intensity of storms.

3.5.1 Cluster Size Distribution

Precipitation storms are generally distributed in different spatial clusters each with different sizes. Therefore, it would be essential for a precipitation replicate generator to preserve the distribution of cluster sizes. Three histogram groups are used to illustrate the distribution of the logarithm of cluster size for the three groups of replicates that were presented in Section 3.4. Figures 3-10 & 3-11 & 3-12 show the histogram of logarithm of cluster sizes in the replicates generated with no constraint, generated by areal support constraint and generated by both areal support and point rainfall constraint. The histogram from the TIs generally matches with the histogram from replicates. The replicate generation method,

with and without constraints, preserves the rainfall cluster sizes in the replicates to an adequate degree. Although there is a low frequency of small clusters in Figure 3-12 (which is for the replicates that have areal and point constraint), the replicates are more realistic spatially. There is a trade-off in which the more the replicates are constrained by points, the more they are similar spatially. However, the cluster size distribution might become biased to some degree.

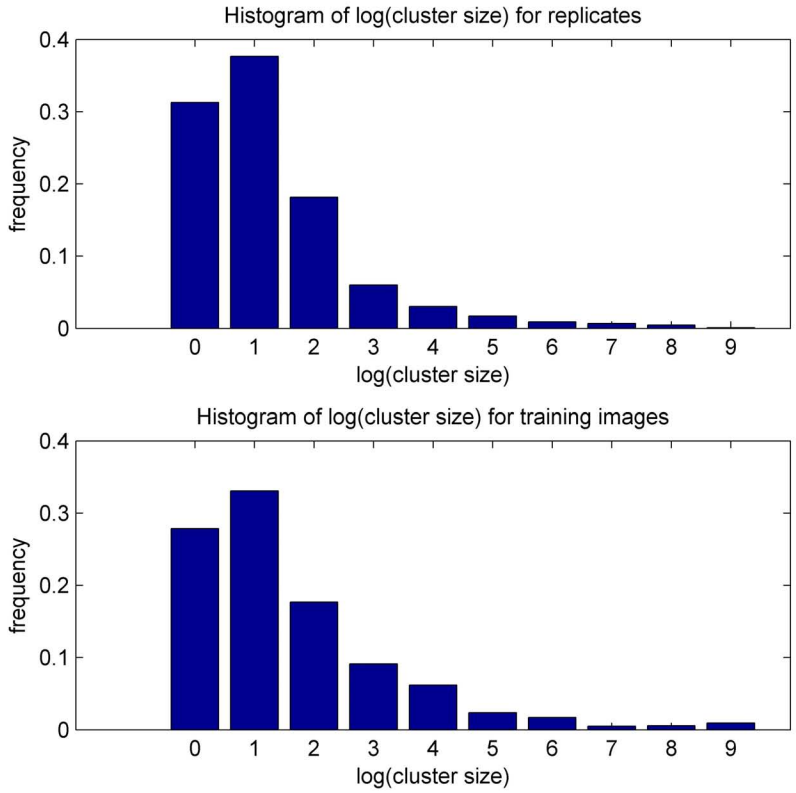


Figure 3-10: Histogram of cluster size distribution in the TIs and the replicates (No Constraining)

3.5.2 Distribution of Rain Rate

Realistic precipitation replicates should both have correct spatial distribution and correct precipitation rate. As a test, we evaluate the cumulative distribution function (CDF) of precipitation rate in the TIs and replicates. Figure 3-13 illustrates the CDF of rain rate for TIs and replicates for group of replicates that are generated with areal support and point

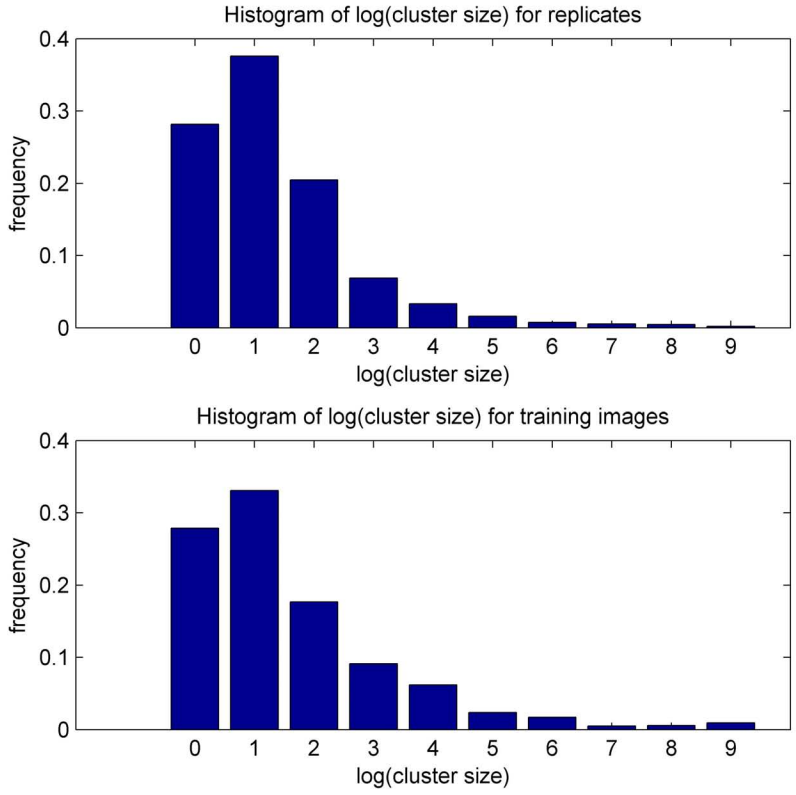


Figure 3-11: Histogram of cluster size distribution in the TIs and the replicates (Areal Constrained)

precipitation constraints. Moreover, Figure 3-14 shows the CDFs for three individual TIs' and their associated replicates.

As these figures show, the replicates appropriately capture the precipitation rate distribution in the TIs. Further investigations showed that all three groups of replicates have the same shape and coverage of rain rate distributions.

3.5.3 Spatial Similarity of Replicates

In this section the spatial similarity of replicates to the corresponding TIs are investigated using Jaccard distance. Jaccard distance was introduced in Equation 2.3. For the comparisons presented here, the TIs and replicates are threshold at 0 mm/hr to create binary images appropriate for use with Jaccard distance.

The Jaccard distances between the TIs and their corresponding replicates have been

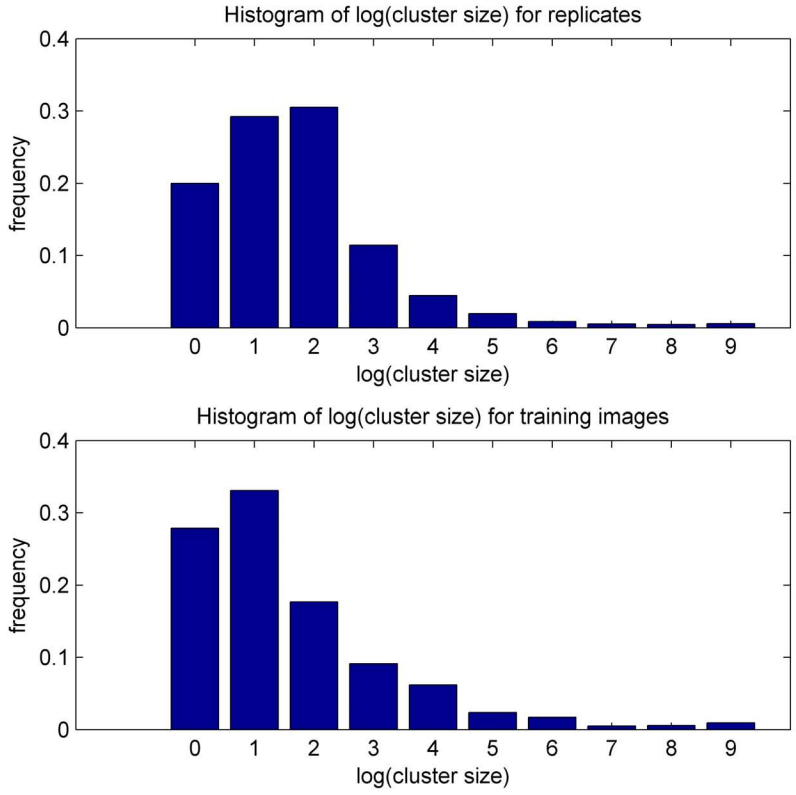


Figure 3-12: Histogram of cluster size distribution in the TIs and the replicates (Areal and Point Constrained)

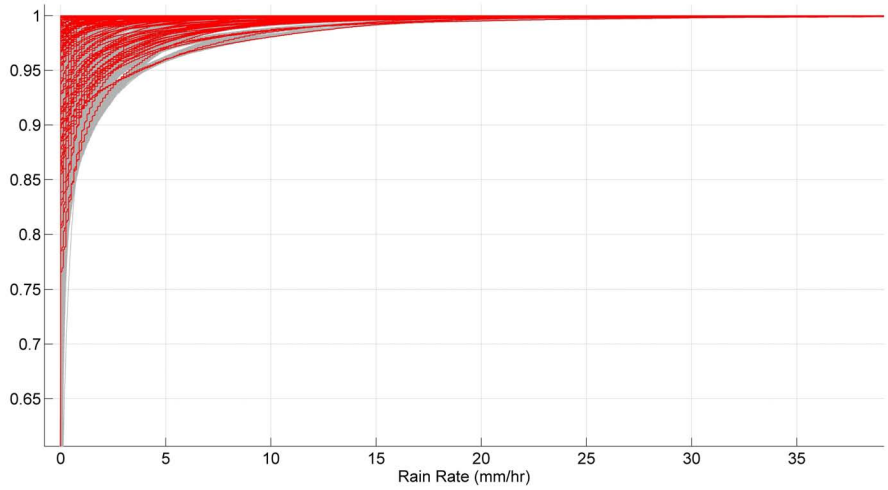


Figure 3-13: CDF of precipitation rate for the replicates and TIs (areal and point constraints). Red colors are CDF of TIs and gray colors are CDF of replicates.

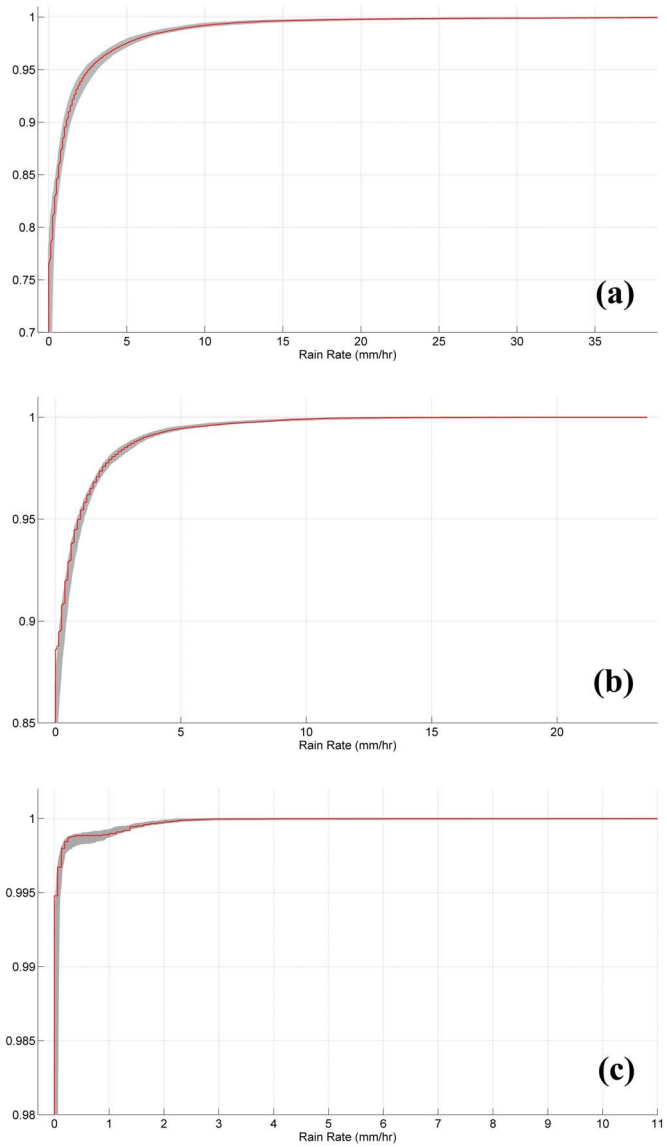


Figure 3-14: CDF of rain rate for three TIs and their associated replicates (areal and point constraints). Red colors are CDF of TIs and gray colors are CDF of replicates.

calculated for each group of replicates. The first group contains unconstrained replicates, the second group contains replicates with areal support constraint, and the third group contains replicates with both areal support and point rainfall constraint. Figure 3-15 presents the mean and one standard deviation of the Jaccard distance for each group in the 16 selected TIs. With greater constraints on the replicates, the Jaccard distances are smaller and replicate similarities are improved. On average the mean of Jaccard distances for group two are smaller than group one, and similarly the mean of Jaccard distances for group three

are smaller than group two for each TI. The variance of Jaccard distances within the three groups are almost the same indicating that the replicates have the same diversity around their mean in each group.

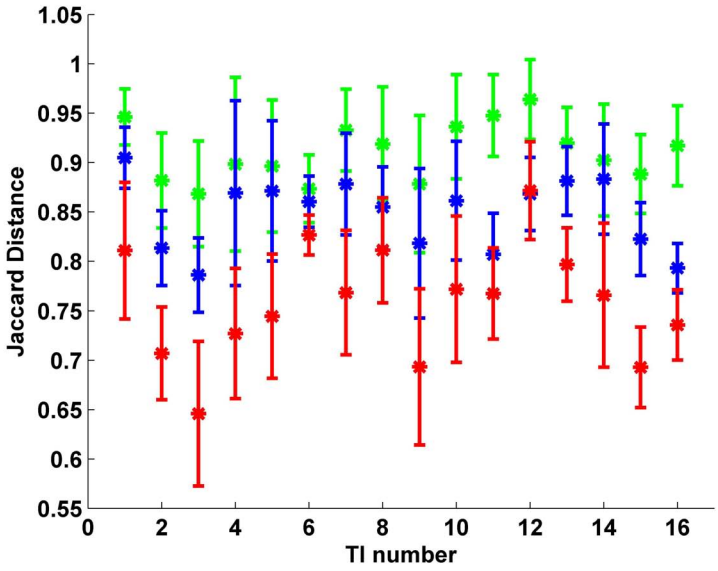


Figure 3-15: Mean and one standard deviation of the Jaccard distance between the replicates and the corresponding TI.

Based on the three skill attributes used in this section, the proposed replicate generation method may reasonably be considered capable of producing realistic precipitation replicates that are intermittent in space. Moreover, the constraint capability provides a tool to generate a diverse set of replicates that differ while containing similar structural features.

3.6 Conclusions and Future Work

A new method for generating realistic precipitation replicates is presented in this chapter. The method that is based on joint statistics of subband coefficients of a training image, uses steerable filters to derive four statistics from a training image that captures features of spatial structure and precipitation rate distribution. These four statistics are used to generate a diverse set of replicates.

This method is also capable of constraining the replicates on areal support (e.g. cold cloud top temperature from geostationary satellite instruments) and point precipitation

measurements or finite swath and orbiting microwave instrument precipitation rate retrievals that are necessary tools for generating precipitation replicates which are constrained on satellite and/or in-situ measurements. The method generates results with adequate similarity to observations and random diversity to capture uncertainty. The replicates can be used as inputs to hydrological and meteorological data assimilation models, as well as flood forecasting models and many other applications.

The method is tested on a population of training images measured by ground-based radar (NEXRAD-IV product) from 2004 summer storms over CONUS. The results show the performance of this method in generating realistic rainfall replicates that are designed to be statistically similar to their corresponding TI. Moreover, the results indicate that the constrained options improve the replicates' spatial distribution and structure. The method has parameters that can be altered to produce the desired or target level of diversity in the replicates. Jaccard distance, a proximity measure for binary fields, is used to evaluate similarity of replicates and corresponding TIs. The results of tests with this skill metric further confirm the way in which replicates become less diverse by applying more constraints.

As further improvement, this method can be extended to generate 3D precipitation replicates from 3D training images. These 3D images can be 3D precipitation patterns in space, or a 3D image that is created by putting together 2D snapshots of precipitation over one region; therefore, it has two dimensions of space and one dimension of time. The former, can be used to investigate vertical patterns of precipitation and the latter can be used to investigate temporal correlations of spatial precipitation fields. It can also be a reasonable path forward for merging microwave-based precipitation retrievals based on multiple low-Earth orbit satellites (with limited swath and temporal overpass time coverage) with precipitating area (without magnitude information) temporally-complete mapping based on visible/infrared measurements from instruments on board geostationary satellites.

Chapter 4

Ensemble-Based Characterization of Precipitation Features

Abstract

This chapter introduces a new method to characterize uncertainties in spatial precipitation retrievals from space-born instruments. Unlike previous studies, this method derives the error likelihood using an archive of historical measurements and provides an ensemble characterization of retrieval error. The characterization process is formulated as a Bayesian sampling problem and solved with a non-parametric version of importance sampling. A novel dimensionality reduction scheme using Singular Value Decomposition (SVD) is implemented to describe images in a problem-specific low-dimensional attribute space, and the importance sampling operations are formulated entirely in terms of the attribute vectors. Examples are presented using retrievals from operational passive microwave instruments and the performance of the method is assessed using ground validation measurements from surface weather radar. Results indicate that this ensemble estimation approach is able to provide an improved description of precipitation features by giving a posterior ensemble that is narrower than the prior.

4.1 Introduction

High-resolution precipitation retrievals from space-born instruments contain uncertainties that necessitates a comprehensive assessment of them before they can be input to hydrological and meteorological models. In this chapter, I address this assessment by developing an ensemble-based method that incorporates historical retrieval errors with observations using Bayes theorem. In this method, a set of equally probable replicates are generated from the current uncertain estimate. Next, each replicate is assigned a probability based on the historical error likelihood. These replicates and the associated probabilities constitute the posterior distribution of the true precipitation. Depending on the application, different characteristics of the posterior distribution (such as mode, mean, etc) can be evaluated.

As it was reviewed in Section 3.1, ensemble approaches have a wide range of applications. There are also different methods to generate ensembles, and several studies have used them to explore uncertainties in precipitation products [25, 40, 69, 187, 46, 188, 11]. In the present study, I use the method developed in Chapter 3 to generate realistic and spatially discontinuous precipitation fields.

The focus of this study is to characterize uncertainties in single satellite retrievals of precipitation, and no merged product is considered here. The reason for taking this approach is that the algorithms that are used to merge different retrievals to produce spatially and temporally continuous precipitation estimates over the globe introduce some errors and uncertainties into the final product that are not of the input retrievals. Therefore, evaluation of the merged product does not provide useful insight about the quality of the retrieval algorithms. By focusing on quantifying errors in single satellite retrievals, I aim to provide information on how the future retrieval algorithms should be improved and how new instruments should be designed.

Figure 4-1 shows four examples of the type of error that this method aims to characterize. Each of these examples show a pair of satellite-based retrieval (from AMSU-B on NOAA-16) and ground-validation (from NEXRAD-IV) for a precipitation storm. Each pair have the same spatial resolution and time of measurement. These snapshots show how complex the errors are in the satellite-based retrievals. An accurate uncertainty esti-

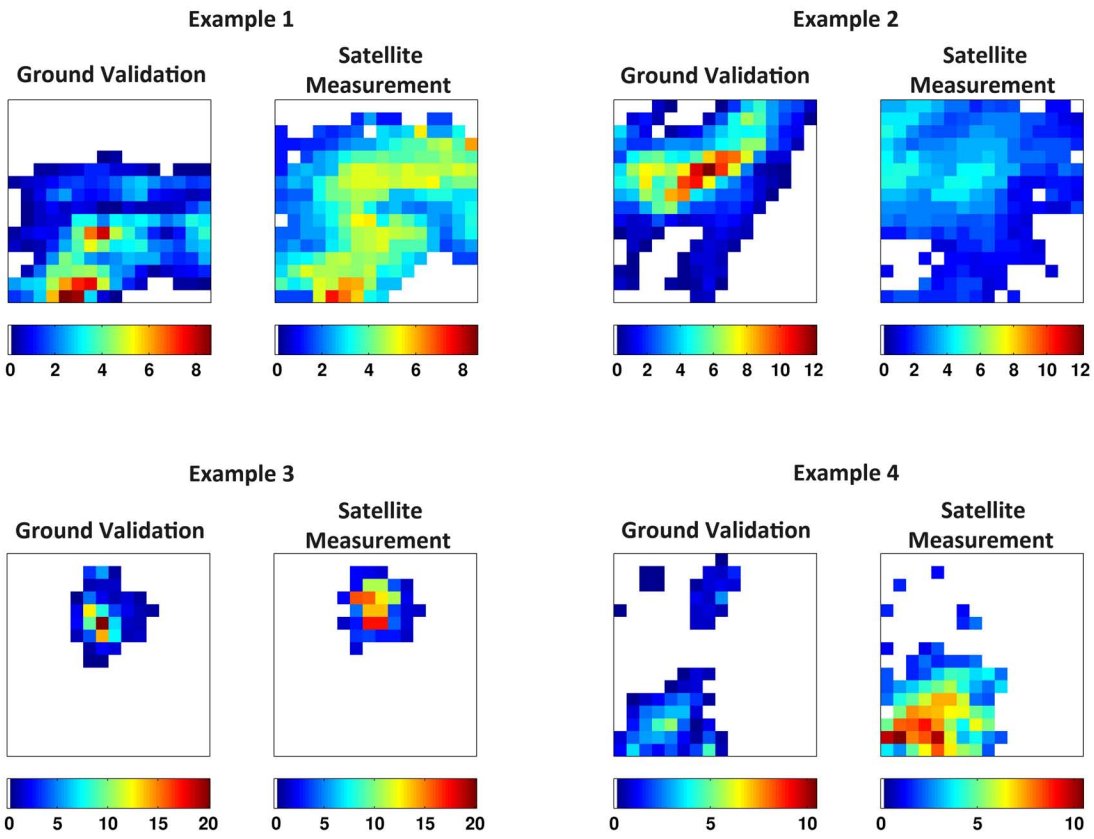


Figure 4-1: Four pairs of precipitation measurement from space-born instruments and their corresponding ground validation. Colors show the precipitation rate in mm/hr

mation of these retrievals needs to take into account both phase (location) and magnitude (precipitation rate) errors.

In the rest of this chapter an introduction to Bayes' Theorem is presented, then the datasets are introduced. Next, the methodology including the dimensionality reduction technique is explained in detail. Finally, examples of this uncertainty quantification along with evaluation of the method are presented.

4.2 Bayes' Theorem

Bayes' theorem provides a tool to measure the degree of belief in the outcome of an event using probability distributions. Based on Bayes' theorem, the degree of belief changes as new observations of the event become available. This theorem provides a great platform for

merging prior knowledge with current observation(s) to infer the state of a system. Assume X is the state of the system to be inferred and Z is an observation of X. Based on Bayes' theorem, we can write:

$$p_{X|Z}(x|z) = \frac{p_{Z|X}(z|x) \times p_X(x)}{p_Z(z)} \tag{4.1}$$

in which $p_X(x)$ is the prior distribution, $p_{Z|X}(z|x)$ is the measurement model, $p_Z(z)$ is the marginal distribution of $p_{Z|X}(z|x)$ and $p_{X|Z}(x|z)$ is the posterior distribution [185, 67]. In this formula x, z are deterministic values of the random variables X, Z, respectively and have m dimensions. This means that x, z represent a $m \times 1$ vector.

In many application $p_Z(z)$ is not known and $p_{X|Z}(x|z)$ is usually estimated to a normalizing factor since $p_Z(z)$ is independent of x. This is the case in the present study as well; therefore:

$$p_{X|Z}(x|z) = C \times p_{Z|X}(z|x) \times p_X(x) \tag{4.2}$$

in which C is a constant. The most common approach to use the Bayes' theorem is to assume Gaussian distribution for $p_X(x)$ and $p_{Z|X}(z|x)$ [30]. This results in an estimation of $p_{X|Z}(x|z)$ similar to using Kalman Filter, which is the best linear estimator [90, 126, 135]. However, as it was mentioned earlier, the goal in this study is to go beyond simple descriptors and parametric distributions. Therefore, I will derive measurement model and prior distribution as empirical distributions using the data itself.

In this study, I use three types of information that are introduced in the following. The first one is the *current uncertain precipitation estimate* (z_c). This is the retrieval from a PMW instrument to be characterized. The characterization aims to infer the true precipitation and the uncertainties associated with the true estimates given z_c. The second one are *prior replicates*. These replicates are generated using a training image (TI) and the details are described in Section 4.4. Prior replicates are denoted as $x_i, i = 1, \ldots, N_{prior}$. The prior TI itself is denoted as x_p. The third set of information are *archival pairs of satellite and ground-based radar measurement*. These are collocated estimates of precipitation from satellite and ground-based radar. The ground-based radar product is used as a benchmark for the true precipitation, and these pairs are used to estimate the error likelihood of each measurement. Details of likelihood estimation are presented in Section 4.6.1. There are a

total number of N_a pairs in this group. In general, N_T data points are used in this study, in which $N_T = N_{prior} + 2 \times N_a + 1$ (1 is for the current measurement z_c).

4.3 Datasets

Datasets used in this study include ground-based radar measurements of precipitation from NEXRAD-IV product, precipitation estimates from the Advanced Microwave Sounding Unit (AMSU-B) on-board NOAA-16 satellite, hereafter referred to as NOAA-16, and the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI)[1], hereafter referred to as TMI.

The temporal coverage of this study is from Jul. 2003 until Dec. 2010. I only use data from summer months (April - October, inclusive), since the storm patterns and the errors have different characteristics in different seasons. The spatial coverage of the study is part of the central continental US ranging $31°37'30''$N and $47°37'30''$N latitudes and $104°37'30''$W and $80°37'30''$W longitudes as shown in Figure 4-2. The domain is selected in a way to avoid measurements in mountainous regions of western US in which ground-based radars are prone to beam blockage and considerable errors in estimation [127]. I have also excluded areas over Great Lakes since the retrieval algorithm over water bodies is different than the one over land areas. To avoid the swath limitations and missing pixels and at the same time maximize the number of samples, the study domain is divided into 23 sub-regions each of which covering a $4° \times 4°$ region. Samples are generated over each sub-region using the collocated datasets. All of the three products are posted on a $0.25°$ lat-lon grid; therefore, the size of each sample is 16 pixel by 16 pixel.

The NEXRAD-IV product used here is the same as the one used in Chapter 2. The processing is the same as the one reported in Section 2.2. I have only matched the time of sampling to the overpass of NOAA-16 or TRMM as needed.

The uncertain PMW-based estimates of precipitation are obtained from NOAA-16. The precipitation rates from this sensor are available through Microwave Surface and Precipi-

[1]Since this study uses data from 2002 until 2010, the mission changes explained in Section 1.1 do not have any impact here.

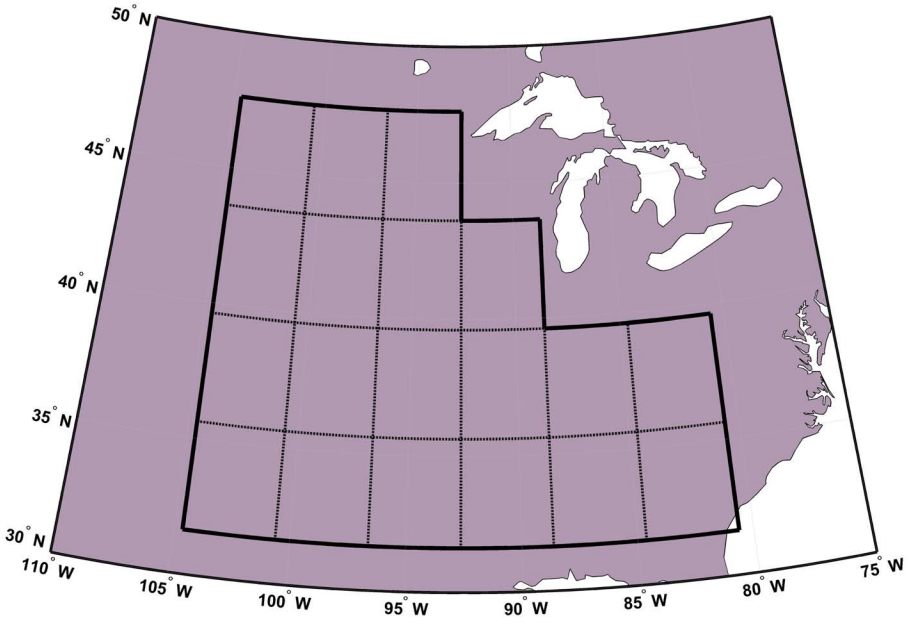


Figure 4-2: Spatial coverage of the study domain. Grids show the 23 sampling sub-regions.

tation Products System (MSPPS) Orbital Data at NOAA. These estimates are available on orbital grids, and I mapped them into a $0.25°latitude \times 0.25°longitude$ grid using nearest neighbor sampling that does not affect the marginal distribution of precipitation intensities.

The precipitation rate parameter from TRMM 2A12 product (version 6) is used in this study to generate prior replicates as it's described in Section 4.4 . The TRMM 2A12 product is produced on an orbital grid, and I mapped it to the into a $0.25°latitude \times 0.25°longitude$ grid using nearest neighbor sampling.

In total, 4,640 pairs of precipitation estimates from NOAA-16 and NEXRAD-IV are generated. These pairs will be used to derive the likelihood function in the Bayesian update framework.

4.4 Prior Generation

In order to generate an informative prior replicate population, the method developed in Chapter 3 is used. This method needs a training image (TI) as an input together with the values of some parameters such as number of orientations and scales for image decompo-

sition.

The TI for the Bayesian update problem should be a precipitation image with the same size and spatial resolution as the uncertain measurement. TI should be selected in a way to reflect the prior knowledge of the user about the state of precipitation at time *t* before the uncertain measurement is made. Prior TI selection is problem-specific and it is based on the expert's judgment to provide an informative and diverse prior population.

In this study, I use TMI-based precipitation estimates from the times preceding the time of uncertain measurement over the same geographical region. TMI measurements are one of the most accurate estimations of precipitation from microwave instruments. This choice is made here for the sake of presenting the application of the method. In general, the user can choose any appropriate prior information to generate a set of reliable and informative replicates as prior samples. Another source of information is to use the output of a forecast model (physical or stochastic) for time *t* based on observations from preceding times that was not available for the temporal domain of this study.

As it was mentioned in Chapter 3, the proposed replicate generation method has parameters that can be tuned to generate population of replicates with different diversities. The level of diversity in the prior population is really important as too much diversity will make the prior uninformative, and a very narrow prior probability density will bias the estimation.

The goal of this section is to evaluate how informative and diverse are the replicates generated from the TMI-based TI and how they capture the uncertainty in the true precipitation of the succeeding hour. To this end, I use a set of TIs from TMI observations and generate replicates using a set of parameters for the replicate generation method. Then, the replicates are evaluated against the true precipitation of the succeeding hour based on the NEXRAD-IV data (assumed to be the benchmark product representing true precipitation).

The priors generated from set of 166 TIs are used to evaluate the prior generation procedure. The parameters that are needed to be set in the prior replicate generation are number of orientations (N_{or}) and scales (N_{sc}) in the decomposition as well as number of replicates (N_{prior}). Conducting several evaluations, the following numbers are used for these parameters: $N_{or} = 4$, $N_{sc} = 2$, $N_{prior} = 1000$. From the 1000 replicates, half are not constrained

by any pixel values but the other half are constrained to preserve 5% of the pixel values in the TI . Moreover, all of the replicates are constrained by cloud observations from GOES satellites. Details of this constraining are described in Section 3.3.4.

The first analysis is to evaluate how the distribution of precipitation intensity in the replicates capture the precipitation intensity distribution of the benchmark product. Figures 4-4 and 4-3 show two sets of Cumulative Distribution Functions (CDF) of precipitation intensity from two examples. These CDF plots include CDF of precipitation from the TI, NEXRAD-IV and the 1000 replicates. As it can be seen the diversity of replicates' CDFs creates an envelope around the CDF of the benchmark estimate (NEXRAD-IV) and provides a diverse and informative set of prior replicates (meaning that the benchmark product can be a member of the ensemble prior replicates). Similar observation was made for other TIs used in this study.

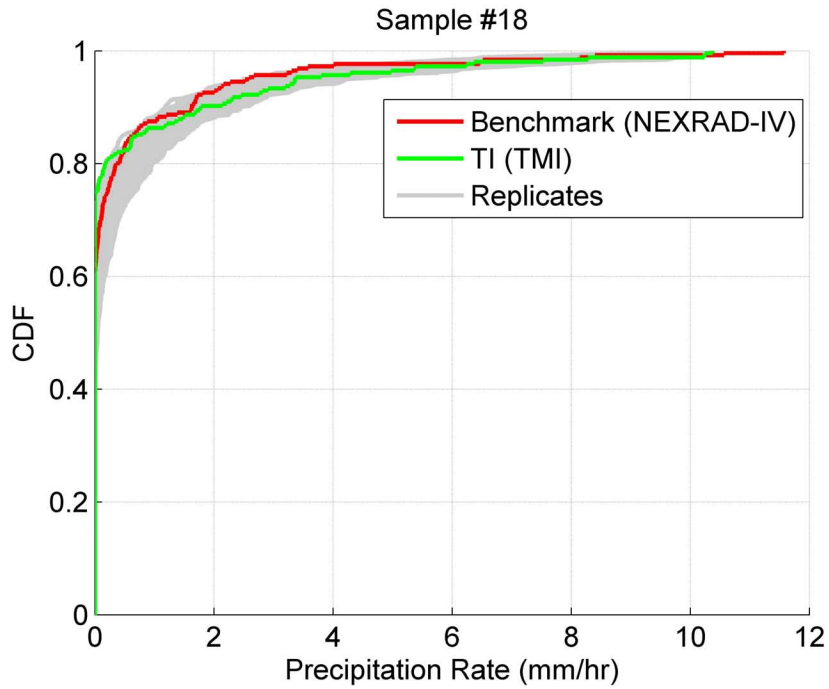


Figure 4-3: Cumulative Distribution Functions (CDF) for the TI, NEXRAD-IV benchmark estimate and the priors generated from the TI (this is for example #18 that is used in Section 4.8)

Another evaluation is carried out on the distribution of the size of clusters of precipitation. Precipitation usually occurs in different patches or clusters. It's necessary for the prior

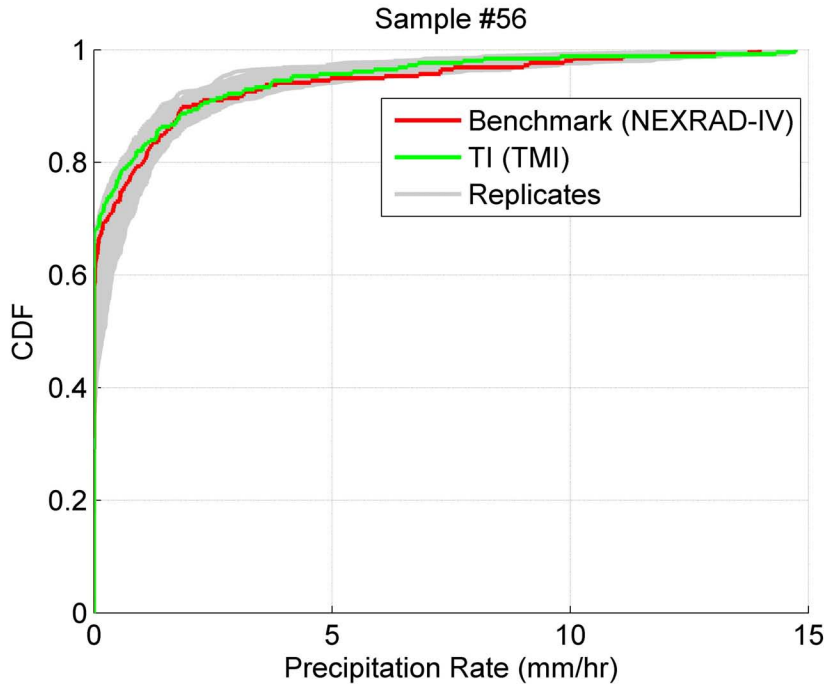


Figure 4-4: Cumulative Distribution Functions (CDF) for the TI, NEXRAD-IV benchmark estimate and the priors generated from the TI (this is for example #56 that is used in Section 4.8)

population replicates to have similar cluster sizes to those of the benchmark product. Figure 4-5 shows the histogram of logarithm of cluster sizes in both the TIs and the replicates. This figure shows that the histograms are very similar, indicating that the prior replicates have similar cluster sizes of precipitation to those of the benchmark product. Moreover, A Kolmogorov-Smirnov test revealed that the null hypothesis that these two histograms are from the same distribution is accepted at 5% significance level.

The analysis provided in this section shows that the prior generation procedure used for this problem produces a set of reliable and informative prior samples that can be used in the Bayesian update scheme to characterize the true distribution of precipitation (posterior distribution). Next section presents the scheme developed for mapping data points to the low-dimensional space.

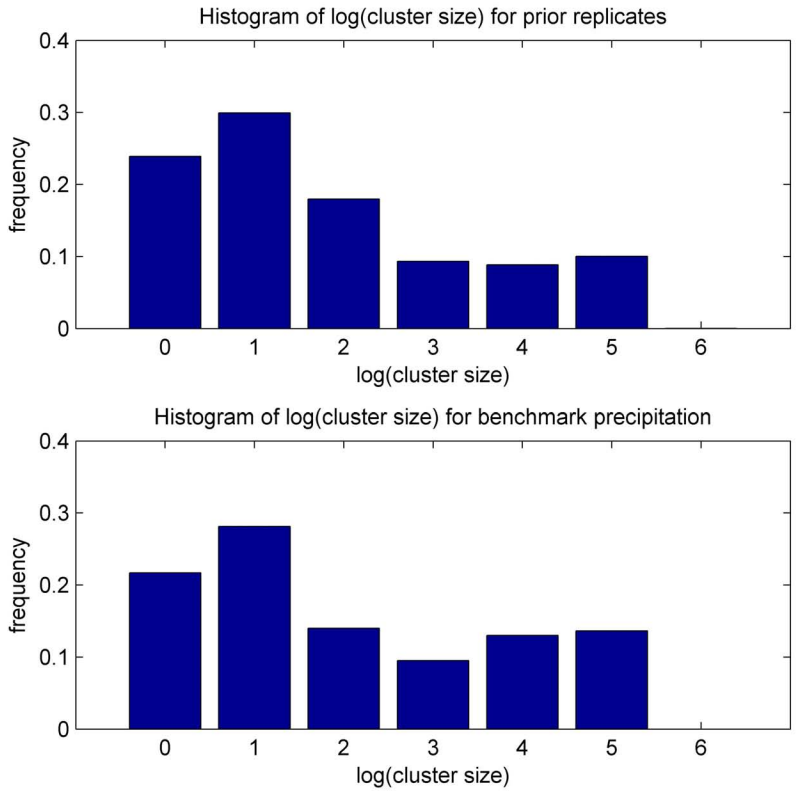


Figure 4-5: Histogram of logarithm of precipitation cluster sizes in prior population and benchmark precipitation product

4.5 Mapping Precipitation Data to a Reduced Attribute Space

It is impractical and computationally inefficient to implement the Bayesian theorem in high-dimensional space. Estimating probability density distributions needs millions of samples in a very high-dimensional space. In this study, the precipitation images are 16×16 pixel, which is a 256 dimensional space. Therefore, it is necessary to map all of the data (uncertain measurement, prior replicates, and archival pairs) to a low-dimensional space, hereafter called *attribute space*.

The goal is to identify dimensions that can best represent the data in low-dimensional space. After identifying the dimensions, data are mapped to the attribute space, and each data point in the original pixel space (or image space) will have a set of attributes that

corresponds to dimensions in the attribute space. Inevitably, there is some information loss in this mapping. However, the goal is to minimize this information loss and develop the most appropriate dimensionality reduction procedure. This dimensionality reduction is problem specific and many approaches are proposed in previous research studies [178, 146].

One common approach is to use Multi-Dimensional Scaling (MDS) techniques to map data into the attribute space [174, 45]. MDS techniques are a family of techniques that aim to map data from an original high-dimensional space to low-dimensional space by preserving the pairwise distances between data as much as possible [31]. There are a range of algorithms for this purpose, including but not limited to: Sammon Mapping [152], Curvilinear Component Analysis (CCA) [50], Generative Topographic Mapping (GTM) [29], Isomap [166], Locally Linear Embedding (LLE) [151], Diffusion Maps [41], Data Driven-High Dimensional Scaling (DD-HDS) [114], and RankVisu [113].

Applying MDS techniques to map precipitation images to a low dimensional space has three challenges. First, there is a need for a distance metric in pixel space. MDS techniques preserve the pairwise distance between samples; therefore, we should be able to measure the distance between two textured images. To the best of knowledge of the author, there is no such distance metric that can measure distances between textured images. The common Euclidean distance is not applicable here as the Euclidean metric does not consider the spatial correlation of pixels in the input image. Second, the quality of mapping from these techniques reduces as the number of samples increases. Beside the mathematical proof for the mapping error (presented in [188, 125]), it makes common sense that the more samples we have, the more pairwise distances are that should be preserved in the low dimensional space, and this would be harder. This limits the number of samples that can be mapped to the attribute space. Consequently, the probability distributions estimated from these limited data will have a lower accuracy. Third, the mappings from these techniques are data dependent. This means that when the data are mapped to the low dimensional space, any new sample that arrives can not be mapped individually and all of the samples should be mapped again together. Considering these three challenges, the MDS techniques are not appropriate for use in this problem.

Other techniques such as Discrete Cosine Transform (DCT) are also available for dimensionality reduction. Using this technique data are mapped to a set of coefficients in the frequency domain; then, the non-significant set of coefficients are truncated and the high-dimensional image is represented using a set of low-dimensional coefficients [6, 86, 85]. This is the technique that is used for compressing JPEG and MP3 files too. However, further investigations show that this is not appropriate for the precipitation application. The basis functions that are used in this method are set of different cosine functions (with different amplitudes and wavelengths). In order to be able to properly describe the highly non-homogenous precipitation images using this technique it is necessary to keep large number of coefficients in the low-dimensional space. However, this is contradictory with the fact that we cannot have a large number of dimensions in the attribute space.

It is expected that the dimensionality reduction scheme provides an attribute space that can distinguish between pairs of ground validation and satellite-based estimated of precipitation. These pairs are very similar and a set of generic basis functions (such as cosine) cannot distinguish between them using few number of attributes (equivalently coefficients).

The dimensionality reduction problem in this study is defined as following. Given a set of data points in pixel space (denote $y_i, i = 1, \ldots, N_T$), find the set of attributes $\hat{y}_i$ by solving the following optimization problem

$$\hat{y}_i = \arg\min_{\hat{y}_i} \sum (y_i - \mathbf{\Psi}\hat{y}_i)^2, \qquad i = 1, \ldots, N_T \tag{4.3}$$

in which $\mathbf{\Psi}$ is the matrix of basis vectors defined as following:

$$\mathbf{\Psi} = \begin{bmatrix} \psi_{11} & \psi_{12} & \ldots & \psi_{1k} \\ \psi_{21} & \psi_{22} & \ldots & \psi_{2k} \\ \vdots & & \ddots & \vdots \\ \psi_{m1} & \psi_{m2} & \ldots & \psi_{mk} \end{bmatrix}$$

in which each column is one basis vector denoted as $\Psi_i, i = 1, \ldots, k$, m is the dimension of data in the pixel space, and k is the number of dimensions in the attribute space.

The objective function of this problem (Equation 4.3) minimizes the square difference

between the original image in pixel space and the reconstructed image from the attributes. The goal is to find the set of $\hat{y}_i$ given $\mathbf{\Psi}$ so that there is minimum information loss as a result of mapping. This information loss is formulated as the difference between input samples and reconstructed samples from attributes.

The choice of basis vectors is problem specific. The most common approach here is to use Singular Value Decomposition (SVD) over all of the samples and derive the eigenvectors of the samples. These eigenvectors can be used as the basis vectors in this formulation. However, more evaluations show that if I use all of the samples to derive the basis vectors using SVD, this will result in a set of smooth basis vectors that can not properly distinguish the pairs of archival satellite and ground-based radar measurement. These pairs have differences, and there is an error in the satellite measurement that I want to quantify; however, their differences is small and it is needed to find a set of problem-specific basis vectors that can distinguish between them. In the following, I present the approach for defining the set of basis vector.

4.5.1 Basis Vector Selection

Basis vectors represent dimensions in the attribute space. The goal is to identify dimensions along which data exhibit the most variability. This enables us to keep more information in the attribute space. Here, I derive three sets of vectors from three population of samples; then, use them together as basis vectors to map all of the data points to the attribute space.

The first population is a censored set of all the archival pairs. I apply censoring here to find a set of archival pairs that are more similar to the current precipitation storm (based on the satellite measurement). The population of precipitation storms from all the years over this region is diverse, and using all of them as samples to derive basis vectors results in smooth basis functions. Figure 4-6 shows a comparison between eignevectors derived from all of the archival samples and the ones derived from a censored population. Smooth basis vectors cannot distinguish between the pairs of the ground validation and satellite estimates; therefore, they give rise to zero error in the attribute space. However, there are differences between the paired measurements, and a measurement-specific set of basis vec-

tors are needed to be able to capture these differences (errors). I implemented the censoring algorithm to derive these basis vectors.

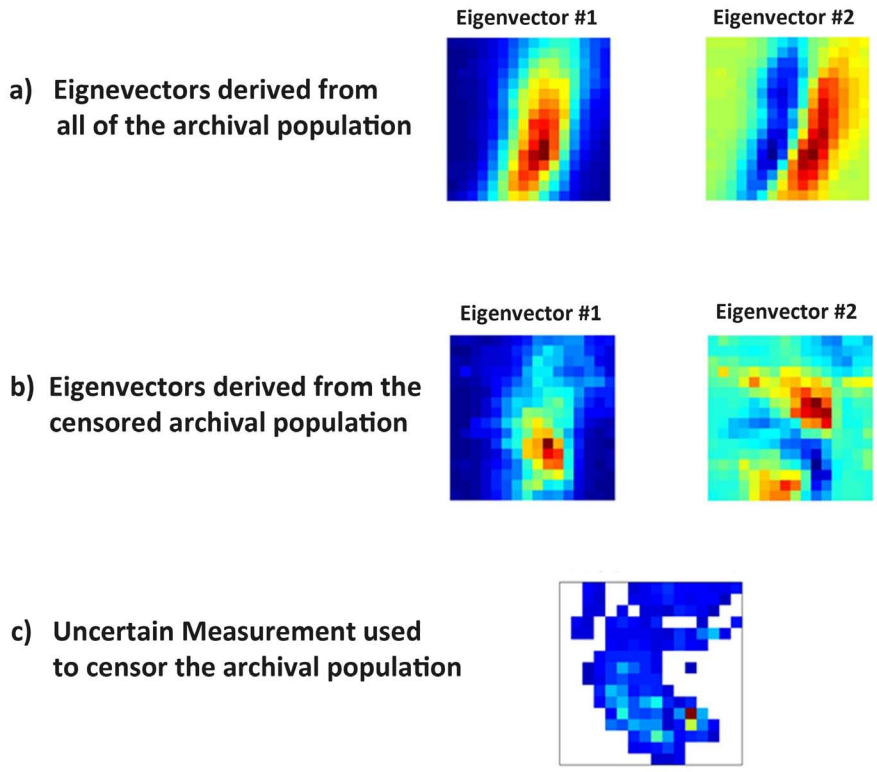


Figure 4-6: Comparison between the eignevectors derived from censored and uncensored population of archival pairs.

Four geometric features are used to censor the archival pairs: Orientation, Solidity and Center of Mass in two directions. These four features are selected based on an extensive evaluation of several features (texture-based, geometric, and morphological ones [111, 107]). Orientation is the angle between x-axis and the major axis of the ellipse that has the same second moments as the precipitation area in the image. Figure 4-7 shows a sample precipitation storm and the orientation angle. Solidity is the ratio of the area of the storm to the area of the convex hull which is the smallest convex polygon that can contain the storm area. Center of Mass are coordinates of the center of the precipitation storm in x and y direction.

For each uncertain satellite measurement, these four features are calculated. Then, the

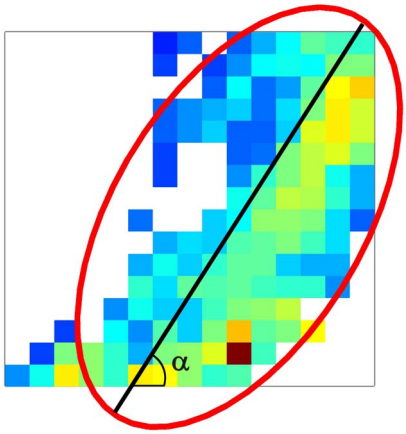


Figure 4-7: A sample precipitation storm and the orientation of the storm (α) defined as the angle between the x-axis and the major axis of the ellipse

archival pairs are monitored and each pair that has a value within $\pm 10\%$ of the values calculated for the features in the uncertain measurement is accepted. The rest of pairs are censored out. Next, the SVD algorithm is applied to the censored population and the leading eigenvectors are derived. I will use between 1 to 3 leading eigenvectors from this population as is described later.

The second population, is the set of prior samples generated for each uncertain measurement. Section 4.4 provides detail explanations of the prior generation procedure. The SVD algorithm is applied to this population and the leading eigenvectors are calculated. Similar to the censored archival data, between 1 to 3 eigenvectors from this population are used as basis vectors.

The third set is the uncertain measurement itself. In order to use the uncertain measurement as a basis vector, first the mean of whole measurement is removed from each pixel, then the uncertain measurement is normalized with respect to it's length:

$$\bar{z_c} = \frac{z_c - mean(z_c)}{|z_c|} \tag{4.4}$$

The eigenvectors from censored archival population and priors together with the normalized measurement ($\bar{z}_c$) form the basis vectors ($\mathbf{\Psi}$) of this method. In order to show the sensitivity of the results to the number of basis vectors (equivalently the number of dimen-

sions in the attribute space), three combinations are used for $\mathbf{\Psi}$. The first one has three dimensions and uses one eigenvector from censored archival population, one eigenvector from priors and the normalized measurement itself. The second one has five dimensions and uses two eigenvectors from censored archival population, two eigenvector from priors and the normalized measurement itself. The last one has seven dimensions and uses three eigenvectors from censored archival population, three eigenvector from priors and the normalized measurement itself. The sensitivity analysis is presented in Section 4.8.2.

4.6 Importance Sampling

One of the novel contributions of this study is that it does not assume any parametric distribution or additive/multiplicative form for the error. Indeed, the error likelihood function in the Bayes' theorem is estimated using pairs of satellite and benchmark ground-based radar measurement that are mapped to the attribute space. This section presents the non-parametric importance sampling algorithm that is implemented to characterize the posterior distribution in Equation 4.1. Figure 4-8 also shows the overall structure of this methodology.

Importance sampling is a method to generate samples from a target distribution (π) by weighting samples that are generated from a second distribution (q, called importance distribution). This approach is necessary when there are complexities in directly simulating samples from π. In Bayesian update problem, it is not possible to directly sample from the posterior distribution $p_{X|Z}(x|z)$; therefore, samples are generated from the prior distribution ($p_X(x)$) and are weighted using the likelihood function ($p_{Z|X}(z|x)$) [148, 43]. These weighted prior samples constitute the posterior distribution that can be characterized for different applications.

4.6.1 Likelihood Function and Posterior Weights

This section presents details of calculating the posterior weights using the likelihood function. As it was mentioned earlier in this chapter, no parametric distribution is assumed for the likelihood function. Instead, all of the pairs of archival data that are mapped into the

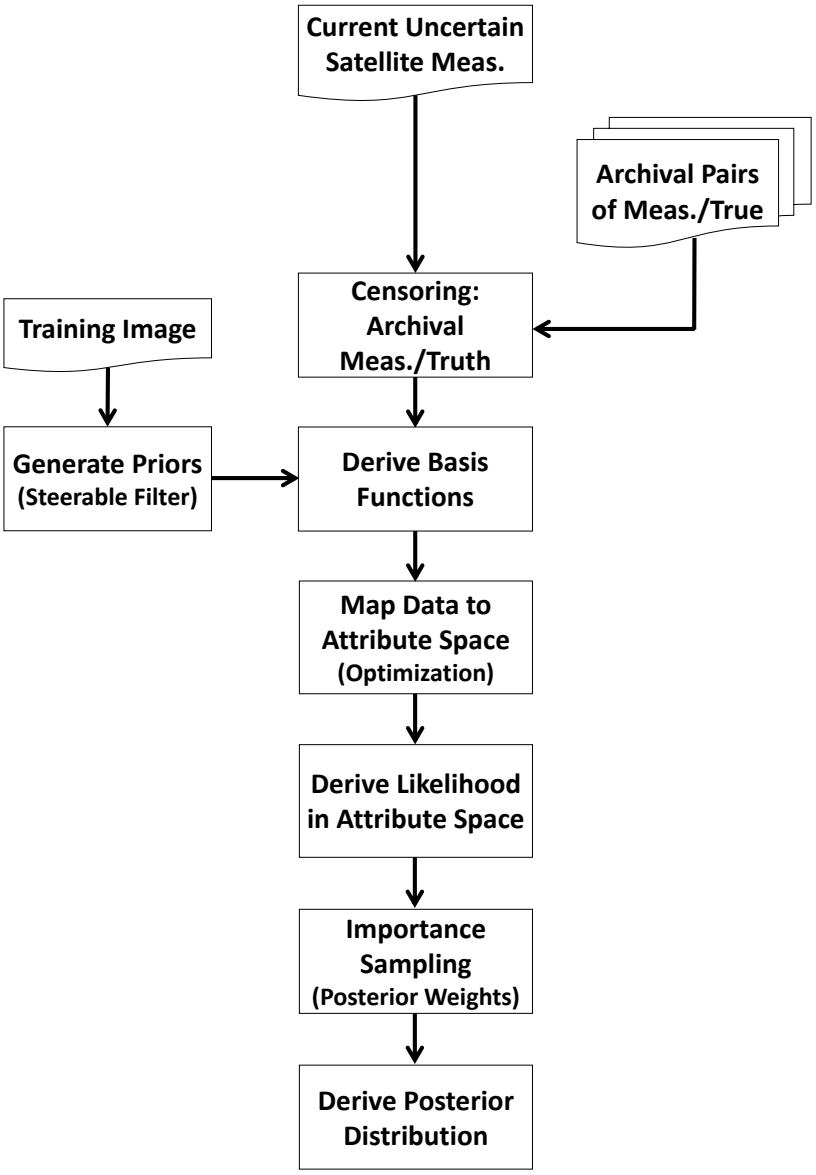


Figure 4-8: Flow Chart of the Ensemble-Based Characterization proposed in this chapter

attribute space are used to estimate the likelihood distribution. I fit a Mixture of Gaussian distributions to the pairs of archival satellite and ground-based radar measurement to find the likelihood distribution. The number of mixtures that is used is adaptive and it is determined using the Akaike Information Criterion (AIC) [7]. This distribution provides the joint probability of $p_{\hat{Z},\hat{X}}(\hat{z},\hat{x})$ between the uncertain measurement and truth (all of the

calculations of importance sampling are implemented in attribute space; therefore, the data are denoted by$\hat{}$). The joint distribution will be used in the following to derive the posterior weights.

Using the conditional probability definition, the likelihood value for each of the prior samples $\hat{x}_i$ is defined as:

$$p_{\hat{Z}|\hat{X}}(\hat{z}_c|\hat{x}_i) = \frac{p_{\hat{Z},\hat{X}}(\hat{z}_c, \hat{x}_i)}{p_{\hat{X}}(\hat{x}_i)}; \qquad i = 1, \ldots, N_{prior} \tag{4.5}$$

The numerator of this equation is evaluated using the probability density function that was estimated by fitting a kernel density to the pairs of archival data in the attribute space. However, the denominator is the marginal distribution of the joint density $p_{\hat{z},\hat{x}}(\hat{z}_c, \hat{x}_i)$ that is impractical to calculate numerically by discretizing the grid (there is no analytical solution for it). Therefore, $p_{\hat{x}}(\hat{x}_i)$ is calculated by applying Monte Carlo integration technique using importance sampling algorithm [148]. To this end, a q_i distribution is selected for each prior sample $\hat{x}_i$ and the integration is calculated using the following relationship:

$$p_{\hat{x}}(\hat{x}_i) = \int p_{\hat{z},\hat{x}}(\zeta, \hat{x}_i)\mathrm{d}\zeta = \int \frac{p_{\hat{z},\hat{x}}(\zeta, \hat{x}_i)}{q_i(\zeta)} q_i(\zeta)\mathrm{d}\zeta = E_{q_i}\left[\frac{p_{\hat{z},\hat{x}}(\zeta, x_i)}{q_i(\zeta)}\right] \tag{4.6}$$

in which E_{q_i} is the expectation over q_i. Therefore, $p_{\hat{X}}(\hat{x}_i)$ is evaluated by generating samples ζ from q_i and calculating the expectation on the right side of Equation 4.6. For this problem q_i is chosen to be Gaussian distribution with mean μ_{q_i} and variance $\sigma^2_{q_i}$. The values of μ_{q_i} and $\sigma^2_{q_i}$ are estimated based on a population of $\hat{z}_j$ samples from archival pairs that have their corresponding $\hat{x}_j$ close to the prior sample $\hat{x}_i$. Further evaluations reveal that the choice of q_i is appropriate and the value of the integral converges by increasing the number of ζ samples.

This likelihood estimation is also applied to the current uncertain measurement $\hat{z}_c$. Uncertain measurement is itself a potential replicate for the true precipitation; therefore, it is used as a prior replicate in this problem. Among the examples presented in Section 4.7, there is also an example that the uncertain measurement will get the highest probability in the posterior distribution.

4.6.2 Characterizing the Posterior Distribution

The goal of this study is to generate samples from the posterior distribution $p_{\hat{X}|\hat{Z}}(\hat{x}|\hat{z}_c)$ that have the highest probability and can represent the range of uncertainty in estimating the true precipitation. Therefore, I use the weights assigned to each of the prior samples, to estimate the mode of the posterior distribution using the following steps:

- Resample the prior replicates using the importance weights calculated in Equation 4.5. This resampled population would represent samples from the posterior distribution;
- Estimate $p_{\hat{X}|\hat{Z}}(\hat{x}|\hat{z}_c)$ using a kernel density fit (mixture of Gaussian distributions) to the resampled population;
- Identify mode of the posterior distribution;
- Find replicates that are closest to the mode of the posterior distribution.

The last step is conducted because the mode of the distribution is not necessarily a point in the pixel space (i.e. an image). All of these calculations are carried out in the attribute space and not all of the points in the low dimensional attribute space can be mapped back to the pixel space. Only points that correspond to a prior replicate have a representation in the pixel space. Therefore, it is necessary to find replicates that are closest to the mode of the distribution.

In the next section, I use this approach to characterize and present posterior samples for different precipitation estimates from AMSU-B instrument on NOAA-16.

4.7 Examples

This section presents examples of applying the ensemble-based characterization method (proposed in this chapter) to real retrievals of precipitation from PMW sensor introduced in Section 4.3. Here, I present results based on mapping data points to a 3D attribute space. The effect of the number of dimensions in the attribute space on the quality of the posterior is evaluated in the next section.

For each example, the precipitation estimate from NOAA-16 is used as the uncertain measurement; prior replicates are generated using TMI estimates from preceding times as prior TI; and archival pairs of NOAA-16 and NEXRAD-IV are used to estimate the likelihood. Moreover, the NEXRAD-IV estimate at the time of NOAA-16 measurement is used for evaluation. This estimate is also mapped to the attribute space (denoted $\hat{x}_T$) for evaluation purpose explained in Section 4.8.

Figure 4-9 shows the uncertain measurement (from NOAA-16) and the ground validation (from NEXRAD-IV) precipitation estimates for example 1.

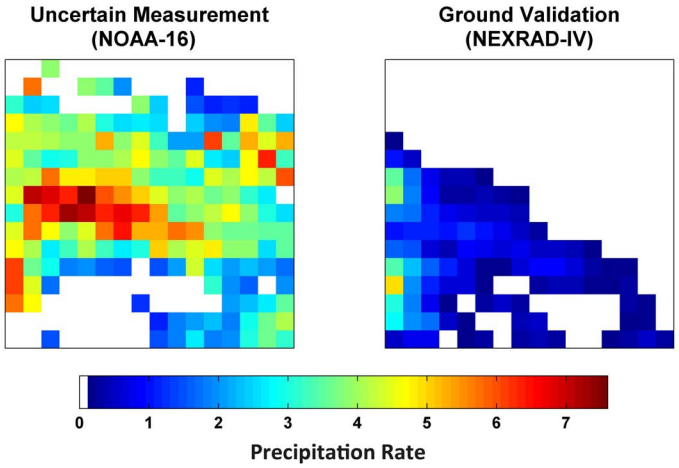


Figure 4-9: Precipitation estimates for example 1 from NOAA-16 (uncertain measurement) and NEXRAD-IV (ground validation).

Figure 4-10 shows the TI (from TMI) and 16 of the prior replicates generated from this TI for example 1. The TMI estimate is from one hour before the time of NOAA-16 measurement. The 16 prior replicates are randomly selected among 1000 replicates.

Figure 4-11 shows 16 of the most probable replicates from the posterior distribution for example 1. These replicates have the closest distance to the mode of the posterior distribution. This figure also presents the Jaccard distance between each posterior replicate and the ground validation (NEXRAD-IV). Jaccard distance is used here to evaluate how close are the posterior replicates to the ground validation compared to the measurement. As it can be seen in this example all of the posterior samples have a Jaccard distance less than the one between the measurement and the ground validation.

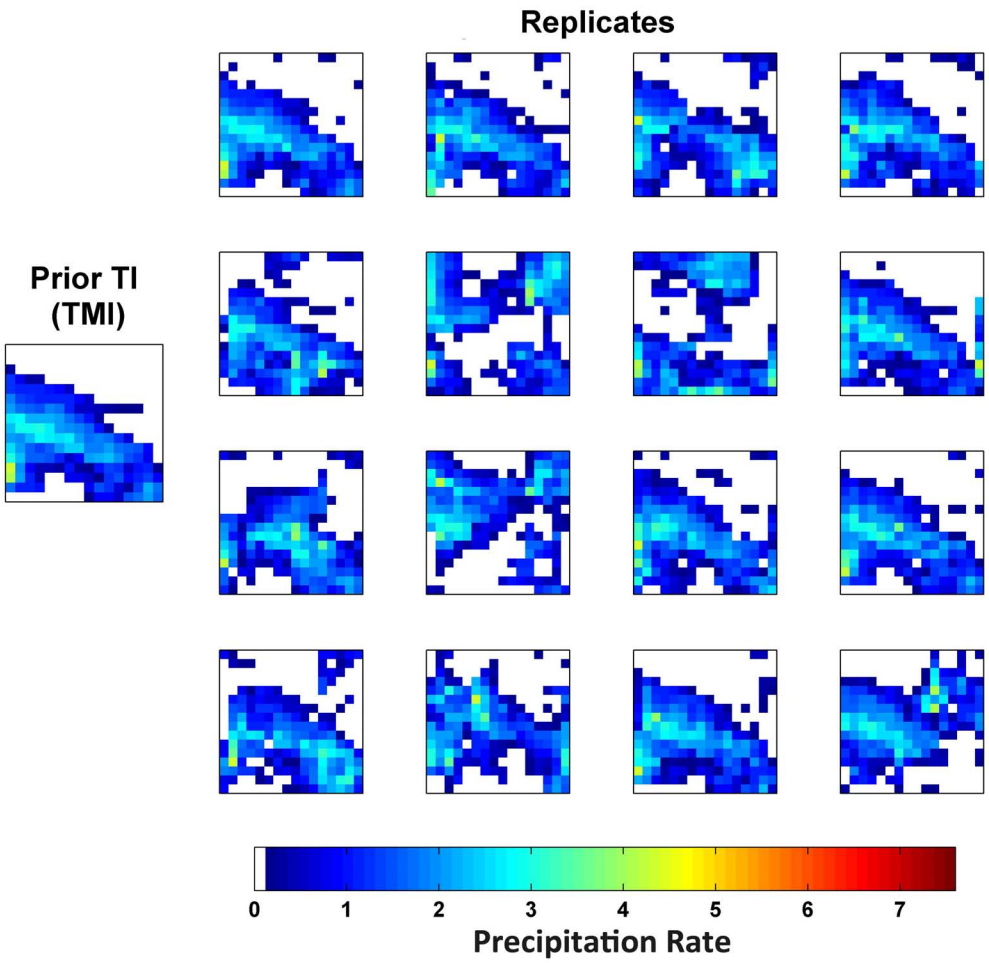


Figure 4-10: Prior TI and 16 of the prior replicates for example 1

Figures 4-12 through 4-14 and 4-15 through 4-17 show two other examples and the results. In example 2, the most probable sample from the posterior (closest sample to the mode) is the uncertain measurement itself. This means the uncertain measurement has a better quality relative to the prior; therefore, the measurement gets the highest probability. In other words, the prior has not been informative enough to improve our knowledge of the true precipitation above the uncertain measurement itself. This shows how important it is to have a good prior. This is absolutely relative. If the uncertain measurement is very good, then providing an informative prior that can improve the uncertain measurement is hard. However, the prior should be chosen in a way that in most of the cases provides an informative knowledge of the true state of the system. In the next section, I will evaluate

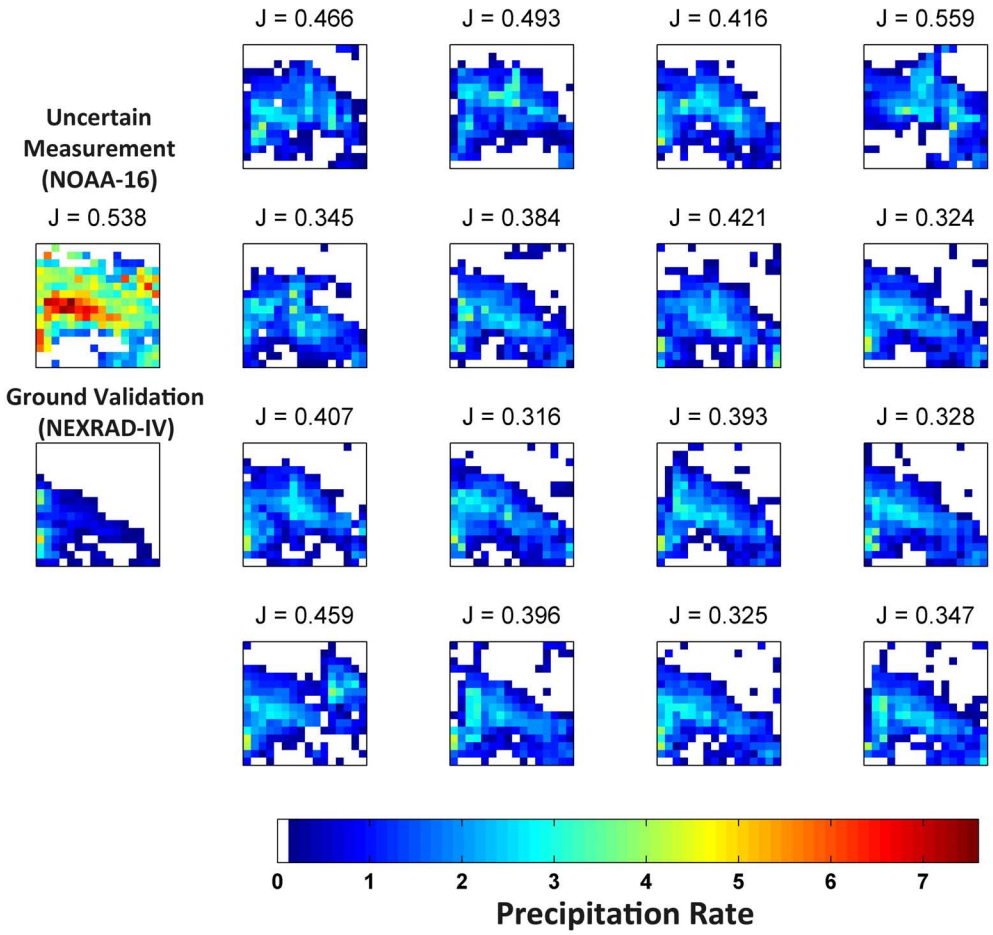


Figure 4-11: The most probable replicates from the posterior distribution for example 1. In the left column the uncertain satellite measurement and the ground validation is also shown. On top of each of the replicates as well as the uncertain measurement the Jaccard distance (J) between the replicate and the ground validation is also reported.

this over many examples and investigate how good are the priors.

4.8 Evaluation

Evaluating an ensemble of replicates is a challenging task. This becomes more challenging when the replicates are images and not scalar values. There are several studies that define indices and skill scores to evaluate quality of an ensemble forecast, but they are focused on scalar variables [33, 123, 16, 186, 136, 149]. Here, the goal is to evaluate an ensemble of

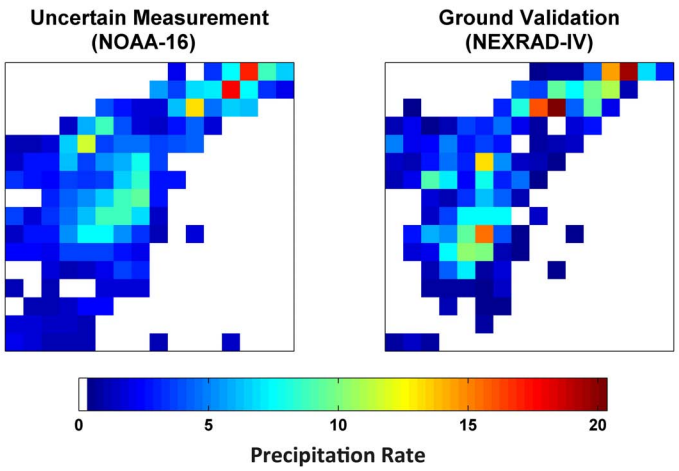


Figure 4-12: Precipitation estimates for example 2 from NOAA-16 (uncertain measurement) and NEXRAD-IV (ground validation).

posterior samples that represent the true precipitation. This is indeed an evaluation of the posterior distribution without knowing the true posterior distribution. In the examples that were presented in Section 4.7 the benchmark estimate (true) is known; therefore, we can use that to evaluate the posterior distribution. However, the benchmark itself is only one sample of the true distribution.

Based on the evaluation provided in Section 4.4 the prior distribution (or the ensemble of prior replicates) is informative and the true precipitation can be a sample form that. Therefore, in this section I want to evaluate the posterior distribution to see if it is narrower than the prior, and if it provides a better estimate of the true distribution.

In the following, three evaluations are presented. First, a relative distance measure is calculated (in pixel space and attribute space) to evaluate the closeness of Maximum a Posteriori (MAP) estimation to the benchmark estimate. Then, Rank Histogram is calculated for scalar variables of total mass and percentage of rainy area in the posterior replicates. Finally, The Jaccard distance at different thresholds (based on quantiles) is evaluated.

4.8.1 Relative Distance Index

In this section, I define two relative indices to evaluate the posterior distribution. These indices measure how good is the mode of posterior in comparison to the measurement and

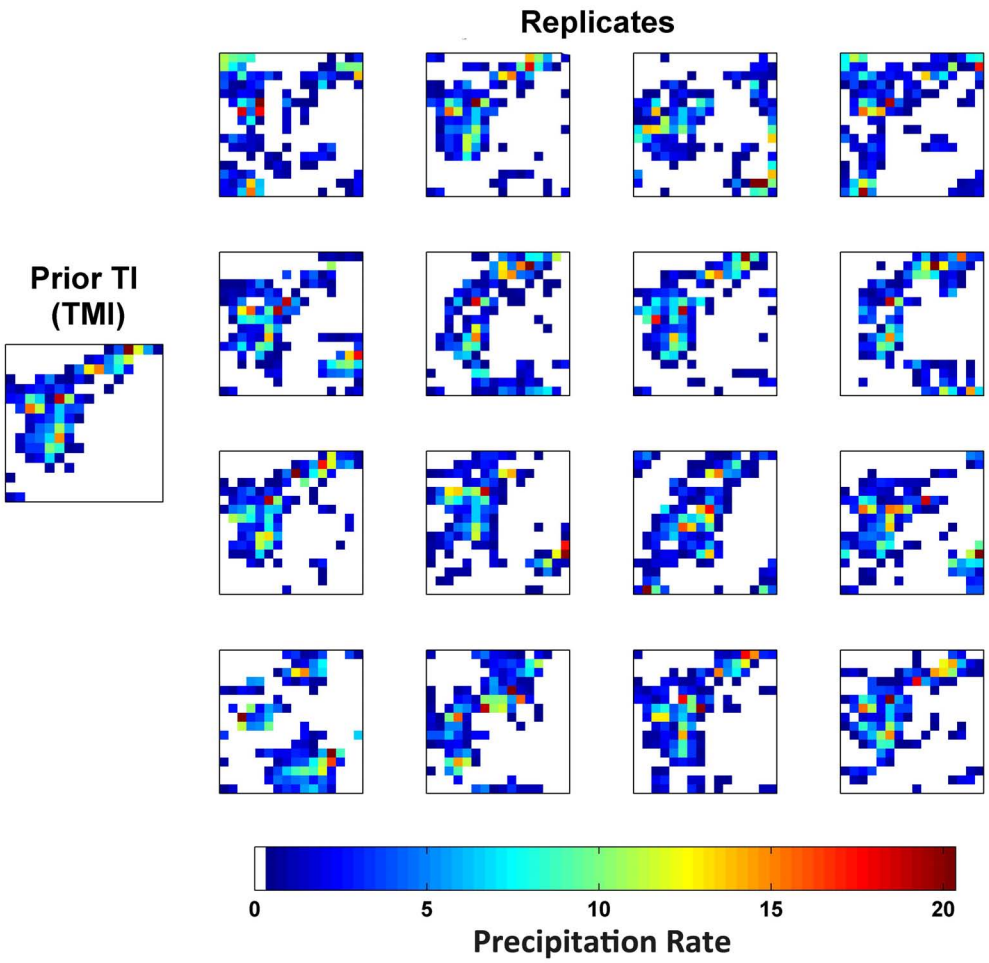


Figure 4-13: Prior TI and16 of the prior replicates for example 2

to the prior TI, respectively. They are defined as:

$$R_1 = \frac{J(x_T, mode(p_{X|Z}(x|z_c)))}{J(x_T, z_c)} \tag{4.7}$$

$$R_2 = \frac{J(x_T, mode(p_{X|Z}(x|z_c)))}{J(x_T, x_p)} \tag{4.8}$$

in which J is the Jaccard distance function. A similar notation is used to define two other relative distance indices in attribute space:

$$\hat{R}_1 = \frac{D(\hat{x}_T, mode(p_{\hat{X}|\hat{Z}}(\hat{x}|\hat{z}_c)))}{D(\hat{x}_T, \hat{z}_c)} \tag{4.9}$$

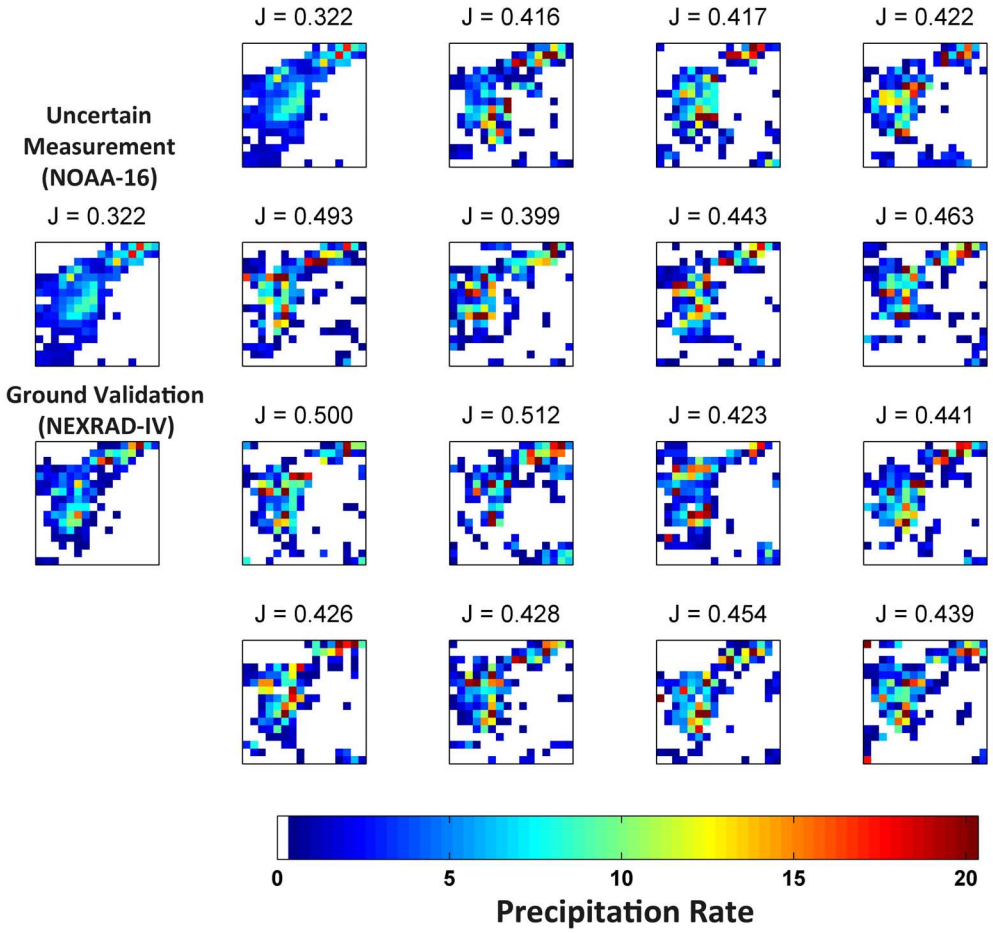


Figure 4-14: The most probable replicates from the posterior distribution for example 2. In the left column the uncertain satellite measurement and the ground validation is also shown. On top of each of the replicates as well as the uncertain measurement the Jaccard distance (J) between the replicate and the ground validation is also reported.

$$\hat{R}_2 = \frac{D(\hat{x}_T, mode(p_{\hat{X}|\hat{Z}}(\hat{x}|\hat{z}_c)))}{D(\hat{x}_T, \hat{x}_p)} \tag{4.10}$$

In this case D is the Euclidean distance in the pixel space. If R_1 (or $\hat{R}_1$) is less than one, it means the mode of the posterior is a better estimate for the truth compared to the uncertain measurement. If R_2 (or $\hat{R}_2$) is less than one, it means the mode of the posterior is a better estimate for the truth compared to the prior TI. It is preferred that all four indices are less than one. However, this is not the case always, as in some situations the choice of

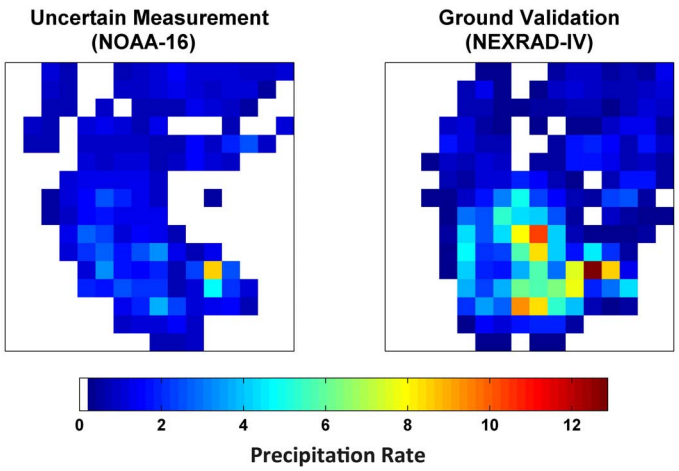


Figure 4-15: Precipitation estimates for example 3 from NOAA-16 (uncertain measurement) and NEXRAD-IV (ground validation).

prior might not be informative enough or maybe the likelihood function is not good.

I used 166 examples and evaluated all four indices. The results are summarized in Figures 4-18 and 4-19. These figures show that in 82% of the examples (in 3D attribute space) R_1 is less than or equal to one. Therefore, in 23% of the examples the posterior samples generated by this method are not better than the measurement itself. This can be caused by a misleading prior or likelihood distribution. The prior can be misleading because the storm might have evolved quickly between the TMI measurement (used as prior TI) and the time of uncertain measurement. The likelihood might be wrong if the precipitation storm is a rare event and no storm similar to that has occurred before. In this case the likelihood function has limited information about the error and it might bias the estimation.

In case of R_2, in 65% of the examples the index is less than one. Again in 35% of the cases the posterior samples are not better than the prior TI, and we can use the prior TI as an estimate for the truth. However, the goal of this study is not to provide a single deterministic estimate of the true precipitation. This method aims to provide replicates that represent the uncertainty in our knowledge of the true precipitation. This uncertainty neither can be summarized in a single scalar value nor can be evaluated using a single true value. Even though the prior TI is in some cases better than the mode of the posterior dis-

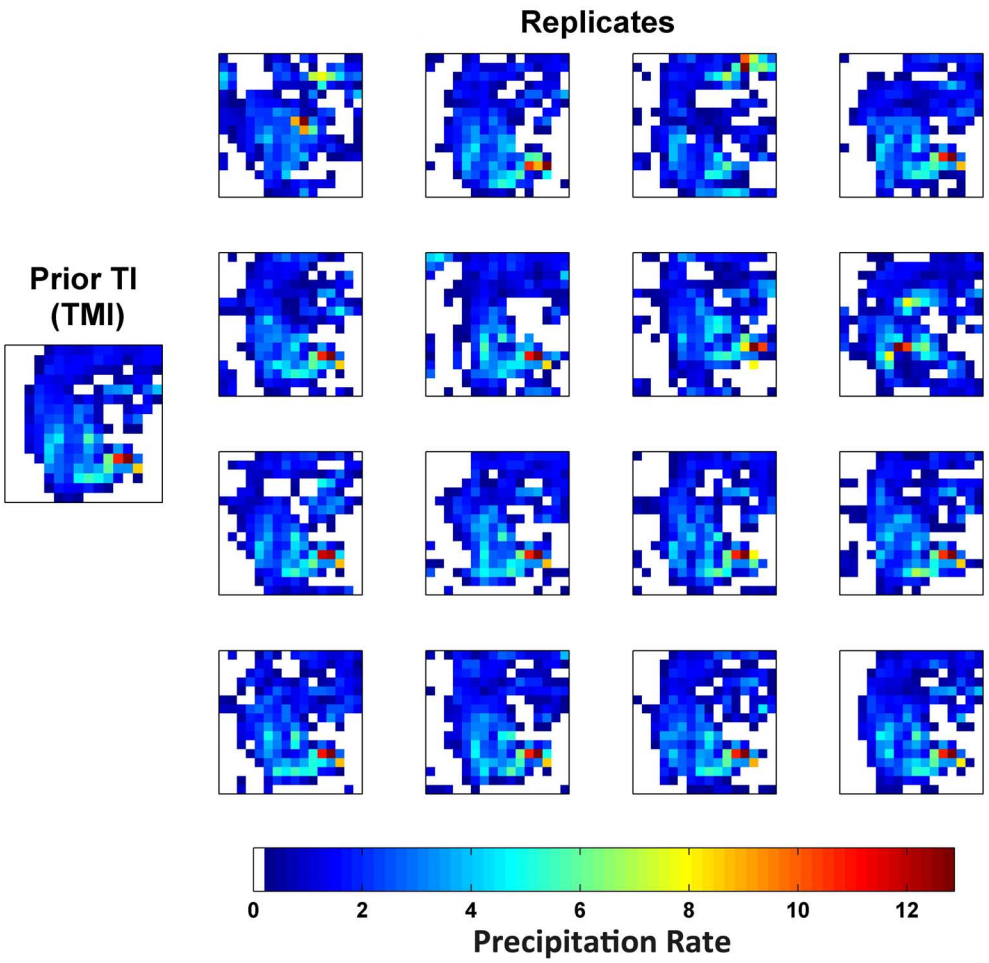


Figure 4-16: Prior TI and16 of the prior replicates for example 3

tribution, the ensemble of posterior replicates are representative of a range of uncertainty in the true precipitation. This ensemble can be used as a random forcing in many hydrological and meteorological models to derive uncertainty estimates for other variables of the hydrological cycle.

Similar results are obtained for $\hat{R}_1$ and $\hat{R}_2$. For these two indices the percentage of the cases that have values less than or equal to one is slightly lower; however, similar trends is observed. This also shows that the dimensionality reduction method applied here, properly projects data into a low dimensional space and the attributes are representative of images in the poxel space.

Another analysis is carried out to evaluate effectiveness of the proposed approach in

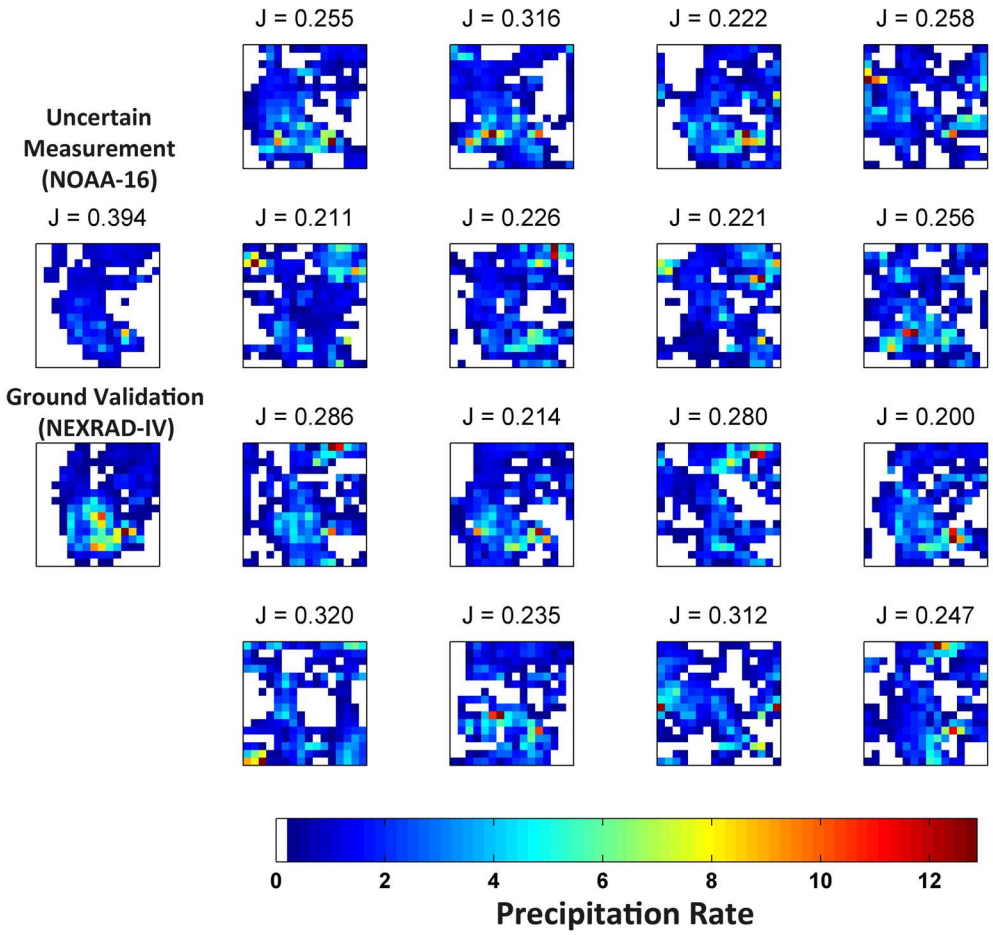


Figure 4-17: The most probable replicates from the posterior distribution for example 1. In the left column the uncertain satellite measurement and the ground validation is also shown. On top of each of the replicates as well as the uncertain measurement the Jaccard distance (J) between the replicate and the ground validation is also reported.

this chapter in comparison to the common approaches in error characterization of remotely-sensed precipitation estimates. Histogram matching is a common technique to reduce errors in an estimate using a reference estimate. For example, in generating global precipitation datasets usually the less accurate precipitation estimates from instruments like SSM/I are histogram matched with the more accurate TRMM combined radiometer and radar estimate.

Here, I will compare how good would be the uncertain current measurement if it is just histogram matched with the prior TI (denoted as $\acute{z}_c$). I use a revised definition of R_1 and

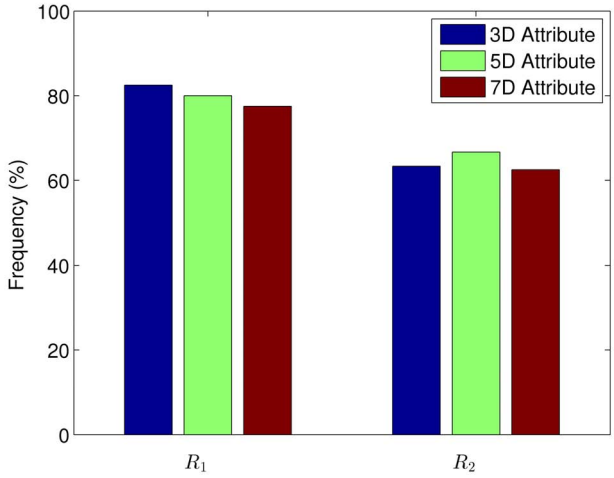


Figure 4-18: Percentage of R_1 and R_2 that are ≤ 1 over 166 examples

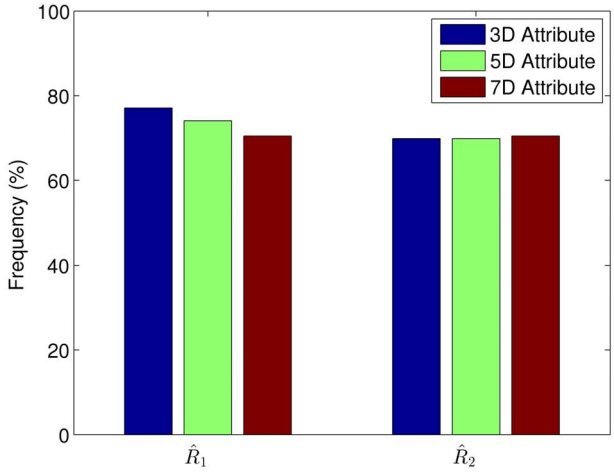


Figure 4-19: Percentage of $\hat{R_1}$ and $\hat{R_2}$ that are ≤ 1 over 166 examples

R_2:

$$\acute{R_1} = \frac{J(x_T, \acute{z_c})}{J(x_T, z_c)} \tag{4.11}$$

$$\acute{R_2} = \frac{J(x_T, \acute{z_c})}{J(x_T, x_p)} \tag{4.12}$$

Evaluating this index across all the 166 examples shows that the percentage of the

examples that $\acute{R}_1$ and $\acute{R}_2$ are less than or equal to one less than the similar percentage in R_1 and R_2 as well as $\hat{R}_1$ and $\hat{R}_2$ (Figure 4-20). This indicates that a simple histogram matching is not effective in reducing the uncertainty in precipitation estimates from space-born instruments. The proposed approach in this chapter takes advantage of historical observed error patterns and merges that with prior information to provide a comprehensive characterization of these uncertainties. This comparison shows that the proposed approach has a better performance compared to the histogram matching approach if one wants to only use the top posterior replicate as a deterministic characterization of error.

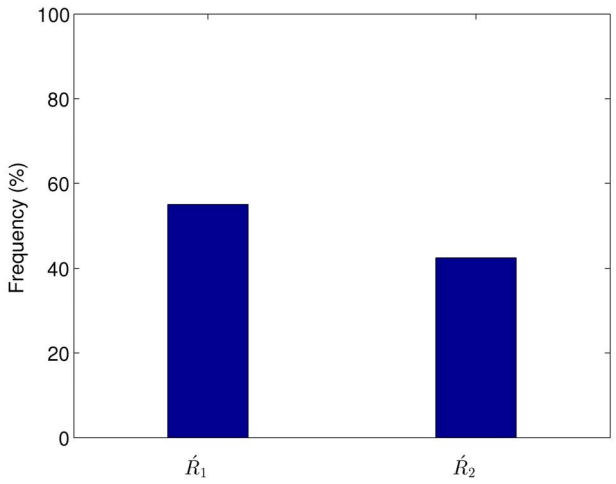


Figure 4-20: Percentage of $\acute{R}_1$ and $\acute{R}_2$ that are ≤ 1 over 166 examples

4.8.2 Sensitivity Analysis

Figures 4-18 and 4-19 show the results of evaluating the method if the attribute space is 5D and 7D. This evaluation is carried out to do a sensitivity analysis on the choice of the number of dimensions in the attribute space. In Section 4.5.1 the combination of eigenvectors from three different populations that form the basis vectors for dimensionality reduction was introduced. For 5D, I have used two eigenvectors from the censored archival population, two eigenvectors from the prior replicates and the normalized uncertain measurement as the fifth basis vector. In 7D, three eigenvectors from the censored archival population,

and three eigenvectors from the prior replicates are used together with the normalized uncertain measurement to form the 7 basis vectors.

Results for R_1 and $\hat{R}_1$ show that by increasing the number of dimensions in the attribute space, the results get worse and in fewer number of examples the posterior mode is better than the uncertain measurement. The changes are small, but noticeable. At first this might seem contradictory because the more dimensions we keep in the attribute space, the less is the information loss in the mapping. Therefore, the method should be able to provide a better estimate. However, when the number of dimensions increases more data points are needed to be able to estimate the probability distributions in the attribute space. This is a limitation in this case. The number of paired archival samples is small (4640 samples) compared to the number of dimensions (5 and 7). Therefore, by increasing the number of dimensions the error in estimating the probability distributions diminishes the information gained from keeping more dimensions.

Results for R_2 and $\hat{R}_2$ do not show a meaningful change in the quality of the posterior compared to prior TI as the number of dimensions increase.

4.8.3 Rank Histogram

Rank histogram is a measure to evaluate the spread of an ensemble. This measure is only applicable to scalar values and has been used in many hydrological and meteorological applications [77, 36, 12, 34, 39, 65, 52]. In this technique a set of bins is defined by distributing the ensemble replicates on the real axis. Then the probability of occurrence of the observation within each bin is evaluated. The result is expected to be a uniform distribution, meaning that the observation has equal probability of occurrence in each of the bins. Details of this technique can be found in [16, 77].

The ensemble replicates in this problem are high dimensional, so it is not possible to apply rank histogram to the ensembles themselves. However, two scalar quantities are considered for this analysis: The total mass of precipitation in each replicate and the rainy area. These two quantities gives us a comprehensive analysis of the spread of the ensembles both with respect to rain detection and rain intensity. Figures 4-21 and 4-22 show the resulting

rank histograms by considering the top 20 posterior replicates. The rank histogram for the rainy area shows that the more probable posterior replicates (low rank ones) are more similar to the benchmark estimate (since they have a higher frequency in the histogram). The rank histogram for the total mass of precipitation shows that the benchmark estimate has equal probability to have a total mass in any of the 20 bins (between top 20 posterior replicates). These two histograms show that the posterior ensemble replicates have a reasonable spread with respect to rain detection and rain intensity.

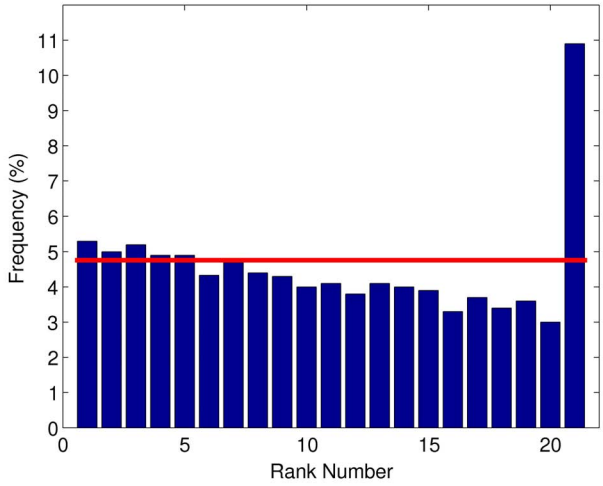


Figure 4-21: Rank Histogram for rainy area. The red line shows the uniform distribution.

4.8.4 Multi-Level Jaccard Distance

In the examples shown in Section 4.7 the Jaccard distance between each posterior replicate and the benchmark estimate was reported. It was shown that in most of the cases these distances are lower than the Jaccard distance between the uncertain measurement and the benchmark estimate. However, Jaccard distance was applied only by thresholding the replicates at 0 mm/hr. Here, I will present the result of comparing the posterior replicates and the benchmark estimates using Jaccard distance but with thresholds higher than 0 mm/hr.

This evaluation is carried out over all the 166 examples and the results are shown in Figure 4-23. The plots show mean and one standard deviation of Jaccard distance at different thresholds. These thresholds are based on the quantiles of the precipitation rate in all the

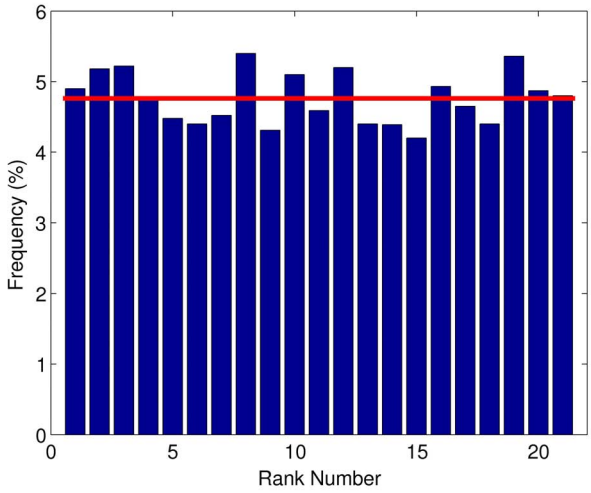


Figure 4-22: Rank Histogram for total amount of precipitation. The red line shows the uniform distribution.

examples that were sued. They are respectively the 5%, 10%, 50%, 90% and 95% quantiles. This figure shows that at higher precipitation rates both the uncertain measurements and the posterior replicates (the top 20 replicates) have a higher Jaccard distance; however, for all the thresholds the posterior replicates have a better performance. Moreover, at higher thresholds and unlike the measurements, the posterior replicates have a lower standard deviation (i.e. lower variability). This indicates that on average the posterior replicates provide a better representation of the true precipitation.

Figure 4-23 also shows the Jaccard distance between all the prior replicates and the benchmark estimate. The results indicate that there is a large variability in the prior replicates (high standard deviation); however, the mean is similar to the top posterior replicates at higher thresholds. At lower thresholds the priors have a higher mean of Jaccard distance. This comparison shows that the ensemble characterization developed in this chapter is properly characterizing the posterior distribution by assigning correct posterior weights to the prior replicates.

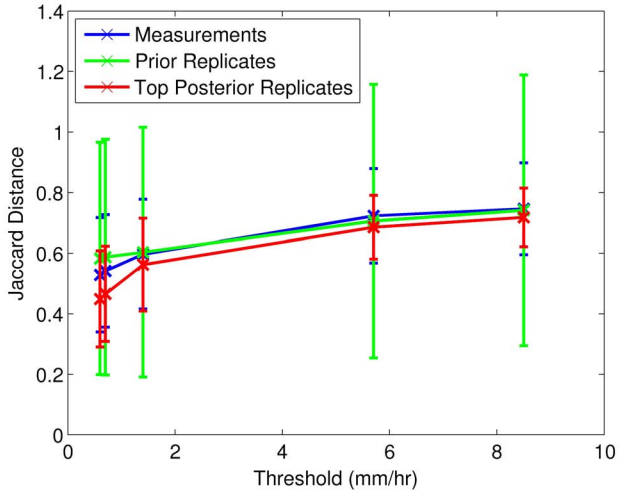


Figure 4-23: Jaccard Distance between uncertain measurement and their associated benchmark estimates (in blue) and between top 20 posterior replicates and their benchmark estimates (in red). The error bars show one standard deviation data.

4.9 Summary

This chapter presents a Bayesian framework to update an uncertain satellite-based estimate of precipitation using historical error likelihoods. This framework provides a population of replicates from the posterior distribution that represent the uncertainty in the true precipitation.

The error likelihood is derived by comparing historical observations of the satellite and ground-based radar. The likelihood does not assume any parametric distribution for the error, and it is based on an empirical distribution. This approach provides a more realistic characterization of error and it takes advantage of a rich dataset of collocated satellite and ground-based radar measurements.

The Bayesian update technique is impractical in high dimensional space. Since the precipitation images in this problem were 256D, I implemented a dimensionality reduction algorithm to map all of the data to a low-dimensional attribute space. The dimensionality reduction is measurement specific and it properly describes the images in the attribute space.

The main goal of this chapter was to show the capability of using the Bayesian theorem

in a low-dimensional space to characterize errors in precipitation estimates. The choice of prior replicate generation (including prior TI) and the dimensionality reduction is user dependent. These two elements can be improved or replaced with other algorithm.

Applying this method to a population of storms observed by AMSU-B instrument on-board NOAA-16 satellite and comparing the results with the estimates from ground-based NEXRAD-IV radar product reveal that the proposed method is capable of characterizing the uncertainty in true precipitation by providing a posterior distribution that is narrower than its prior.

Chapter 5

Conclusions and Future Research

5.1 Conclusions

Monitoring precipitation using space-born instruments has great advantages including near-real time and continuous spatial measurements. Space-born instruments have evolved over time and new sensors with new algorithms that are capable of precipitation retrieval at high spatial resolutions are built. However, these high resolution measurements have high uncertainties that need to be characterized before they can be fully incorporated into decision making and data assimilation tools.

The new Global Precipitation Measurement (GPM) mission focuses on merging precipitation estimates from each of the constellation of GPM satellites to provide global precipitation estimates with high temporal and spatial resolution. However, error quantification of precipitation estimates from each of the satellites is a vital step towards merging the estimates.

In the first part of the thesis, I introduce a new metric for evaluating errors in precipitation estimates. Applying the metric to precipitation products from Special Sensor Microwave/Imager (SSM/I) shows marked seasonality and precipitation intensity dependence. Results show that SSM/I-based precipitation products are positively biased in magnitude, and the bias is less at higher intensities and in geographic locations where precipitation rates are generally higher.

In the rest of the thesis, I develop a new method for characterizing uncertainties in

remotely-sensed precipitation estimates using historical observations and the Bayes' theorem. This method provides a framework for optimally merging prior knowledge of the true precipitation with error likelihoods conditioned on the current uncertain estimates. This data merging provides a characterization of the posterior distribution of precipitation.

A new stochastic-based method was also presented for generating non-stationary prior precipitation replicates. This method is capable of generating realistic precipitation replicates that capture the true uncertainty in the measurement. Moreover, the replicates can be constrained on the cloud measurements from the more frequent geostationary satellites. The prior replicates generated by this method are realistic and yield better results than priors generated from parametric distributions.

This ensemble-based characterization uses a non-parametric version of importance sampling to derive samples from the posterior distribution. Importance sampling is relatively easy to implement; however, it needs adequate ensemble sizes, accurate measurement error models, and realistic priors to be useful in practical applications. For precipitation retrievals that are high-dimensional a very large prior ensemble is needed to provide a representative characterization of the posterior distribution. Generating such a large ensemble is impractical; therefore, a novel dimensionality reduction scheme is introduced to map all the data to a low-dimensional attribute space. The dimensionality reduction scheme is problem-specific and is able to properly describe the data in the attribute space. Then, the importance sampling is implemented in terms of the attribute space.

In order to properly represent measurement errors, and unlike previous studies, this approach relies on non-parametric probability densities estimated from an archival dataset rather than hypothesized parametric distributions. This increases the applicability of the method and insures that the posterior probabilities properly reflect the effects of measurement uncertainty. Moreover, it enables the characterization procedure to make best use of information from different instruments with different strengths and weaknesses.

The primary limitation of implementing importance sampling in a low-dimensional attribute space is the inevitable information loss. Mapping data from a high dimensional space to a much lower dimension attribute space is accompanied by some information loss. This is a function of the number of dimensions in the attribute space, the lower the

dimensions the more the information loss. However, increasing number of dimensions in the attribute space increases the needs for more samples to properly characterize the probability densities. The compromise is to select the number of dimensions in the attribute space in a way that can accurately represent the data and probability distributions.

Results indicate that the proposed attribute-based ensemble estimation approach using importance sampling is a promising option for characterizing complex precipitation images. the method provides an improved description of precipitation storms by giving a posterior ensemble that is narrower than the prior. The ultimate goal of this ensemble-based approach is to provide a range of possible precipitation replicates that represent the retrieval uncertainty.

5.2 Suggestions for Future Research

The proposed error characterization in this thesis has two elements that can be improved in future studies: the dimensionality reduction scheme and the prior generation approach. Indeed, these two elements are application specific and should be changed if this method is going to be used for another application.

The dimensionality reduction scheme that is developed in this study is measurement-specific, and for each measurement it derives a different set of basis vectors. This is advantages because the attribute space is tuned to be able to optimally represent the errors in the current uncertain measurement. However, if this approach is going to be used in real-time application it would be useful to examine alternative options that perform faster. The current scheme runs at a reasonable time on a regular PC for sample sizes that were presented in Chapter 4 but for larger populations and sample sizes this needs to be improved.

The prior replicate generation that is developed in this thesis uses a training image (TI). The choice of the TI is application specific and the more informative is the TI the more informative is the prior population. A better prior population results in a more accurate characterization of the uncertainties in the measurement. It would be useful to use other TIs with this approach, especially those that are derived from physical model forecasts. Physical models now produce hourly forecasts up to 18 hours with high spatial resolution

(over continental US) that are appropriate for this application. These forecasts can be a better choice for prior TI compared to TRMM measurements that are not available every hour.

Another topic worth further effort is the archival sample size. The archival pairs are used to derive the likelihood function in the Bayesian framework. A more diverse and rich archival data will provide a better estimate of the error likelihood. Here, I used around 7 years of archival data; however, with ever increasing measurements from space-born instruments this population can be enriched.

Finally, applying this error characterization to the estimates from a wide range of instruments provides a great tool to compare relative performance of different instruments. This comparison is useful for future instrument and retrieval algorithm developments as well as calibration studies.

Appendix A

Comparing Jaccard Distance with POD and FAR

Definitions

Assume two measurements - one is the true measurement (called T hereafter) and the other is a second measurement to be validated versus the true one (called V hereafter). Based on the definition provided in Section 2.3.2 the Jaccard distance (J), POD and FAR are:

$$J = 1 - \frac{f_{11}}{f_{10} + f_{01} + f_{11}} = \frac{f_{10} + f_{01}}{f_{10} + f_{01} + f_{11}} \tag{A.1}$$

$$POD = \frac{f_{11}}{f_{11} + f_{10}} \tag{A.2}$$

$$FAR = \frac{f_{01}}{f_{11} + f_{01}} \tag{A.3}$$

The main difference between J and POD is that J considers both the hits and misses in the measurements while POD only captures hits. On the other hand FAR only deals with misses. Therefore, the combined outcome of the POD and FAR can be inferred from J itself. However, depending on application it might be advantageous to distinguish between POD and FAR.

It is possible to relate the Jaccard distance uniquely to POD and FAR. Based on A.2:

$$f_{10} = \frac{(1 - POD) \times f_{11}}{POD} \tag{A.4}$$

Using A.3:

$$f_{01} = \frac{FAR \times f_{11}}{1 - FAR} \tag{A.5}$$

Substituting A.4 and A.5 into A.1 results in:

$$J = \frac{\frac{(1-POD)\times f_{11}}{POD} + \frac{FAR\times f_{11}}{1-FAR}}{\frac{(1-POD)\times f_{11}}{POD} + \frac{FAR\times f_{11}}{1-FAR} + f_{11}} \tag{A.6}$$

The f_{11} terms can be eliminated to write J only in terms of POD and FAR:

$$J = \frac{FAR \times POD + (1 - POD)(1 - FAR)}{FAR \times POD + (1 - POD)(1 - FAR) + POD(1 - FAR)} \tag{A.7}$$

Figure A-1 shows the Jaccard distance as a function of POD and FAR. The Jaccard distance changes with different combinations of POD and FAR. The Jaccard distance captures both types of errors measured by POD and FAR.

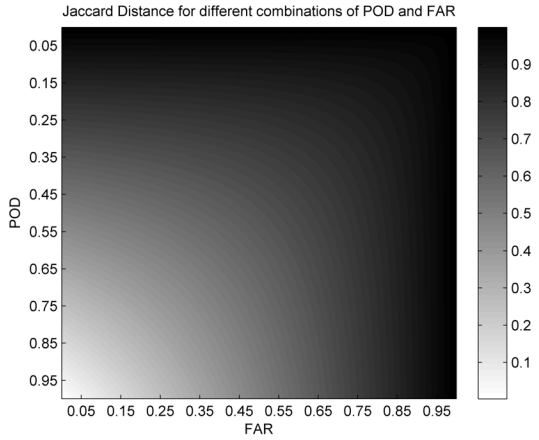


Figure A-1: Jaccard Distance for different combinations of POD and FAR

Examples

In this section, I provide a series of synthetic examples to compare J, POD and FAR. In all the cases the true measurements is the same. Figure A-2 shows 3 examples. In this figure, the first column shows the true measurement (T), the second column shows the second measurement (V) to be compared and the last column shows the overlapped measurements. In the first row, T and V are the same, so J = FAR = 0 and POD = 1. For the second row, V is shifted as not to overlap with T. This is case with the most difference between T and V, so J = FAR = 1 and POD = 0. Row 3 shows a situation in between. Both T and V are overlapping in some pixels, and not all of them. As a result, POD, FAR and J are all intermediate between 0 and 1.

Figure A-3 shows another set of examples with the same T. Here, the first row has a V that has a smaller true area with respect to T, so there is no false detection, i.e. FAR = 0. However, POD and J are not equal to 1 as the two images are different. If we add some false area to V (as in the second row) the POD does not change, but J and FAR change. This shows that POD is not sensitive to false detected areas. So, if there is a V that is just overestimating the true area, it will have a POD=1, although the V and T will not be the same. At the same time, J changes between rows 1 and 2. Thus J is sensitive to both hits and false detections. Row 3 shows how POD, FAR and J change when there is spatial overestimation.

It should be noted that in special cases FAR will have the exact value as J. For example, if $f_{10} = 0$, then J = FAR. This happens when there is no misses in the detection.

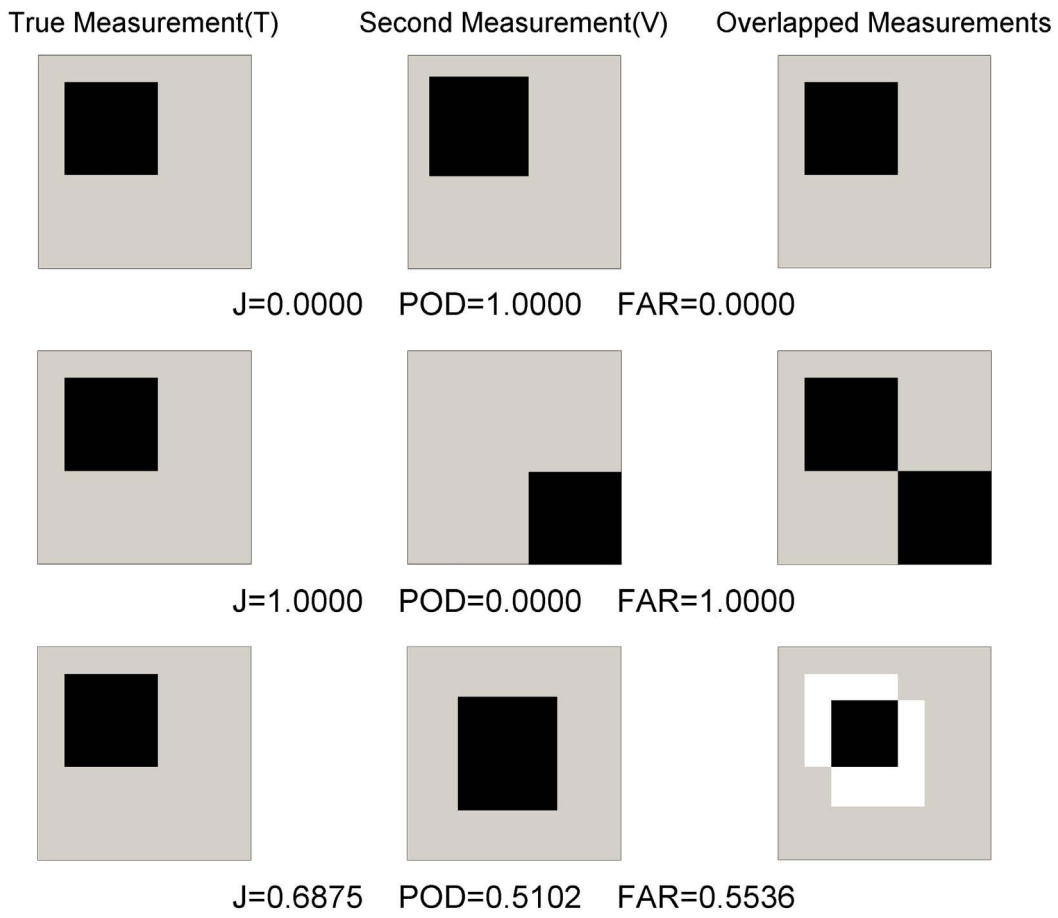


Figure A-2: Example 1 comparing J and POD and FAR (Black color: Feature in T or V, and overlapping feature in the overlapped measurement. White color: Non-overlapping feature in the overlapped measurement. Gray color: No feature)

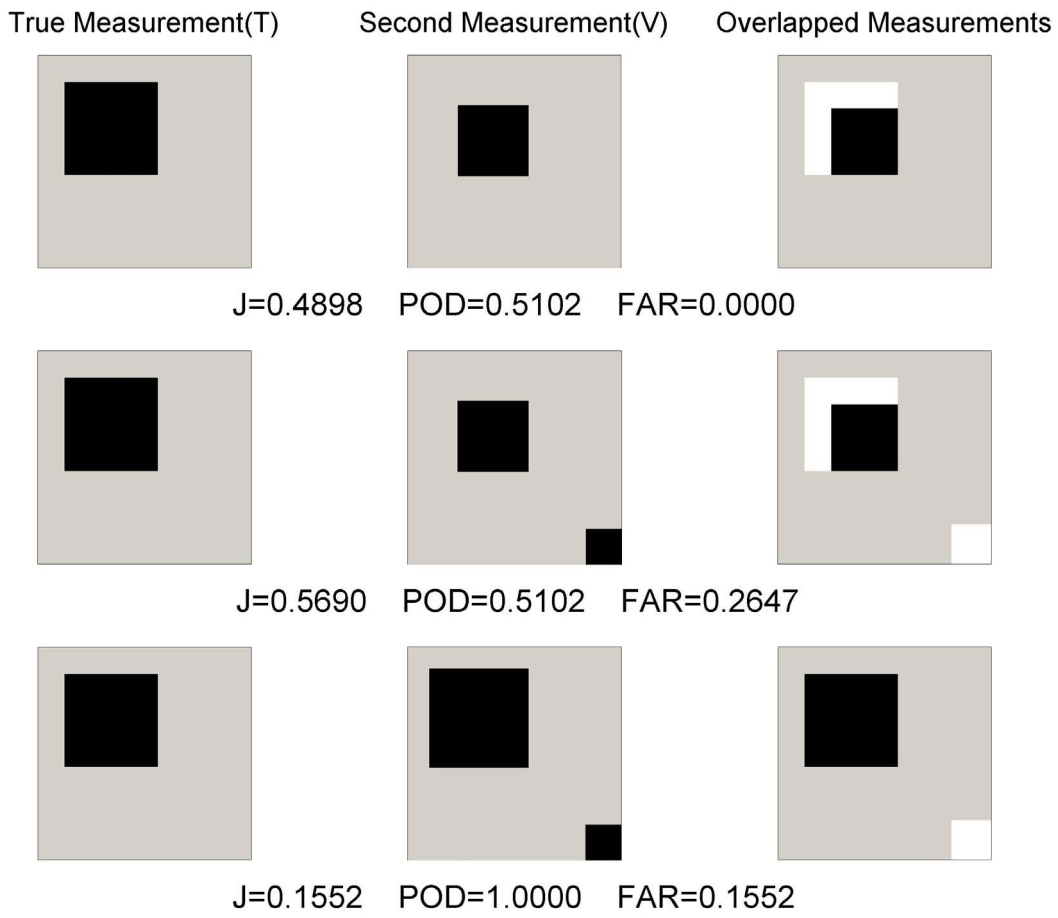


Figure A-3: Example 2 comparing J and POD and FAR (Black color: Feature in T or V, and overlapping feature in the overlapped measurement. White color: Non-overlapping feature in the overlapped measurement. Gray color: No feature)

Appendix B

Merging Satellite Measurements of Rainfall Using Multi-scale Imagery Technique

Abstract

Several passive microwave satellites orbit the Earth and measure rainfall. These measurements have the advantage of almost full global coverage when compared to surface rain gauges. However, these satellites have low temporal revisit and missing data over some regions. Image fusion is a useful technique to fill in the gaps of one image (one satellite measurement) using another one. The proposed algorithm uses an iterative fusion scheme to integrate information from two satellite measurements. The algorithm is implemented on two datasets for 7 years of half-hourly data. The results show significant improvements in rain detection and rain intensity in the merged measurements.

B.1 Introduction

Satellite rainfall retrievals have enormous potential benefits as input to hydrological and agricultural models because of their real time availability, low cost, and full spatial coverage [69]. PMW measurements of rainfall have their own unique advantages as well as

shortcomings. These sensors have a low temporal revisit (usually two times a day) because of their orbital characteristics. Moreover, in some cases they have missing data over some regions where retrieval is confounded by other influences. Therefore, in order to obtain reasonably accurate rainfall estimation and with a good spatial and temporal sampling resolution, useful for studies of climate change and large-scale hydrological processes, it is necessary to merge different types of measurements.

It is not possible to fill in the gaps of one measurement (hereafter, called image) using another one because none of these images are perfect measurements, and this will cause discontinuities in the final image. Moreover, the comparison of results from the proposed fusion algorithm (presented later here) with the results of an interpolation algorithm (which produces a merged image by interpolating between two input images) shows that a simple interpolation is not enough to merge the two images, and the results are not accurate.

In order to produce a better estimation of the rainfall, researchers have tried to combine different types of rainfall measurement. Several models have been developed that combines satellite measurements with ground-base measurements, and provide estimation of rainfall for both missing pixels and for the times that there has not been a satellite measurement. Some of these models include: Global Precipitation Climatology Project (GPCP) [112], Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) [83], Climate prediction center MORPHing method (CMORPH) [87], and Climate prediction center Merged Analysis of Precipitation (CMAP) [190]. Shen et al. (2010) investigated the errors and biases of these merged measurements [156] and showed that they are not reliable in some regions. Here, I will only focus on improving the retrievals when there are at least two satellite retrievals available over one region. However, this method can be applied to any two sets of rainfall measurement to fill in the missing regions or provide a better estimate.

One of the methods that can be used for image fusing is multi-resolution pyramids method [1]. This method is based on fusing the images in a transformed domain, created by decomposing images into sub-images. Here, I propose an algorithm for merging two rainfall images using the multi-resolution pyramid method and an interpolation scheme. The proposed algorithm has two section: *Texture Production* and *Shape Production*. In

order to produce the texture, it uses the multi-resolution pyramid method, and for producing the shape, it uses an interpolation scheme. The results show that the distinction between shape and texture in the algorithm improves the results both in detecting the rainfall and in producing the accurate rain intensity.

In the rest of this chapter, first, the datasets that are used are described. Then, the algorithm is presented in detail. Next, results are provided with comparisons with other methods. Finally, conclusions along with a discussion on the results is presented.

B.1.1 Datasets

The satellite measurements of rainfall for this study are obtained from AMSU-B sensor on-board NOAA-15 and NOAA-16 as well as TMI. The former is produced by Microwave Surface and Precipitation Products System (MSPPS) in NOAA and the latter is produced by Goddard Earth Sciences Data and Information Services Center in National Aeronautics and Space Administration (NASA). An independent set of surface-based radar measurements from NEXRAD-IV product is assumed as the true measurement and used to evaluate the results of the model. This assumption can be further investigated by incorporating the errors of the radar measurements.

For this study, all the data are mapped to a common spatial gird of $0.25°latitude \times 0.25°longitude$ covering the central part of the continental U.S. ranging $26.10° - 42.10°N$ and $107.85° - 91.85°W$. The period of study is from Jan. 2004 until Dec. 2010. All the data are half-hourly measurements.

In order to obtain good samples, I selected images that had a reasonable number (at least 40) of non-missing pixels. Imposing this criteria, a set of 219 coincident images for each of the NOAA-16 and TRMM, and a set of 221 coincident images for each of the NOAA-15 and TRMM was created. Each sample is a 64×64 pixel image.

B.2 Methodology

The procedure for producing the merged images consists of two steps: *texture* simulation and *shape* simulation. Here, texture is the rain rate (or intensity) and shape is the rain

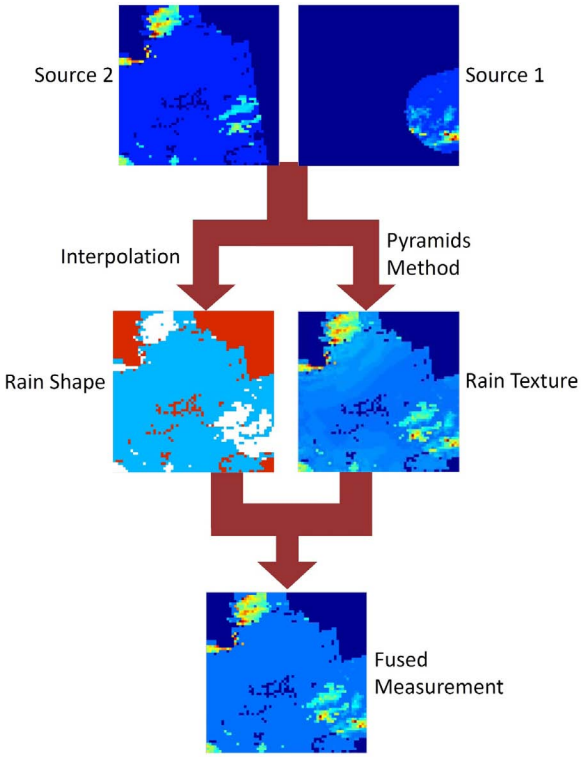


Figure B-1: Procedure of the fusion algorithm. Rain Shape is an image of [-1, 0, 1]. Other images show rain intensity (from blue to red the intensity increases from zero to its maximum; dark blue are the missing regions)

support (i.e. rain / no rain areas.) The complete procedure is illustrated in Figure B-1. Using the two input images, a texture and a shape is produced using the methods described in the following. Then, the fused image is created by multiplying the texture and the shape.

B.2.1 Texture Production

Image fusion using pyramids transformation is a useful method to fill in the gaps of one image using another one(s). This method that was first proposed by Adelson et al., 1984 uses an iterative fusion scheme to integrate information from multiple images [1]. Figure B-2 shows the algorithm for this method. First, a filter is applied to the two input images to decompose them into sub-images (or equivalently to coefficients). Then, a fusion algorithm is applied to merge the subimages and produce a fused subimage. Finally, an inverse transformation will produce the merged image.

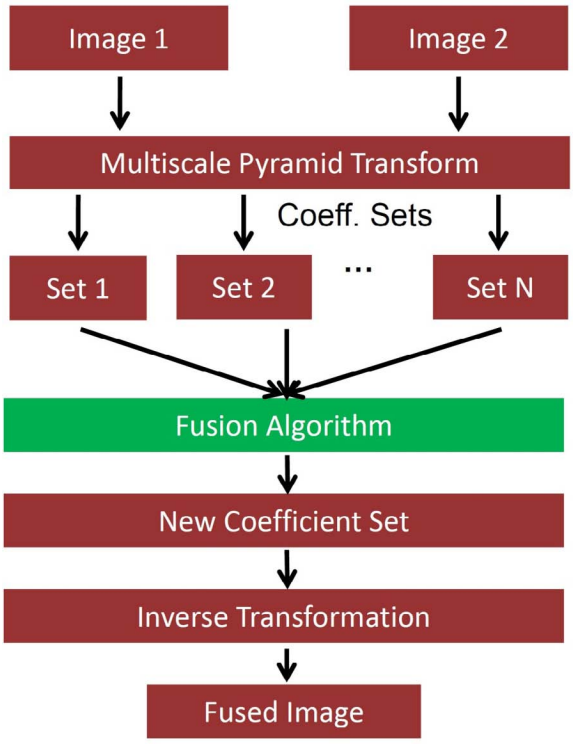


Figure B-2: The pyramids algorithm for producing texture (the Fusion Algorithm is described in Figure B-3)

Rainfall is an intermittent event in space, so the choice of the filter is one of the important aspects in this work. It is necessary to have a decomposition that can obtain the most information possible from the images. Preliminary experiments using Gaussian, Laplacian, and Steerable pyramids showed that Steerable pyramids works better in Rainfall is an intermittent event in space, so the choice of the filter is one of the important aspects in this work. It is necessary to have a decomposition that can obtain the most information possible from the images. Preliminary experiments using Gaussian, Laplacian, and Steerable pyramids showed that Steerable pyramids works better in decomposing the images, and the resulting fused images are more accurate. Thus, I will only use the Steerable pyramids in the following.

Steerable pyramid was introduced in detail in Section 3.3.1. Steerable pyramid has been widely used in vision problems and image analysis. Here, I use it to transform the initial images to a sub-image domain and inverse-transform the fused sub-images to the merged final image. Investigating different number of levels and orientational decomposi-

tions, it was found that a 4 level decomposition (which is the maximum possible) and 16 orientational decompositions are good enough to describe the images in sub-image sapce.

Figure B-2 shows that a fusion algorithm is needed to combine the two sub-images (or coefficient sets) in the transformed domain. A simple criteria for this integration is absolute value maximum selection (AVMS). This criteria selects the larger absolute values in the sub-images as the corresponding value for the fused sub-image. Although this criteria works well in many applications, it is so simple and it might miss some of the details in the integration process. Liu et al. (2001) proposed a fusion algorithm using the Laplacian pyramids [118]. In this algorithm each pair of subimages from the two input images are decomposed by using the Laplacian filters. Then, an integration scheme is applied to merge the two pyramids of the subimages. This integration scheme can be AVMS or any other defined criteria. Next, the fused subimage is produced by inverse transformation of the integrated pyramids of the input subimages. This fusion algorithm allows for more information to be fused into the merged subimages, and for the rainfall images that are so intermittent is really necessary. Therefore, I use this fusion algorithm for merging the subimages. Figure B-3 shows the structure of the fusion algorithm.

B.2.2 Shape Production

This part of the algorithm will produce shapes for the rainfall fields (i.e. support). Early investigations on using the pyramids method to merge the two rainfall images revealed that the method will produce a kind of continuous texture in the resulting image, which is not the case in reality. Moreover, the pyramids method produces high values of false alarm in the fused images. Therefore, I implemented this part to constrain the rainy areas in specific regions of the final fused image. Nevertheless, these regions are still defined by the two input images, but using an interpolation scheme.

The procedure is that a linear interpolation is applied to create a merged image of the two input ones. Then, the pixels that were missing in both of the input images are set to be missing in the final image too. This is because of the high intermittency of rainfall patterns that makes it hard to interpolate the value of the missing pixels from the non-missing ones.

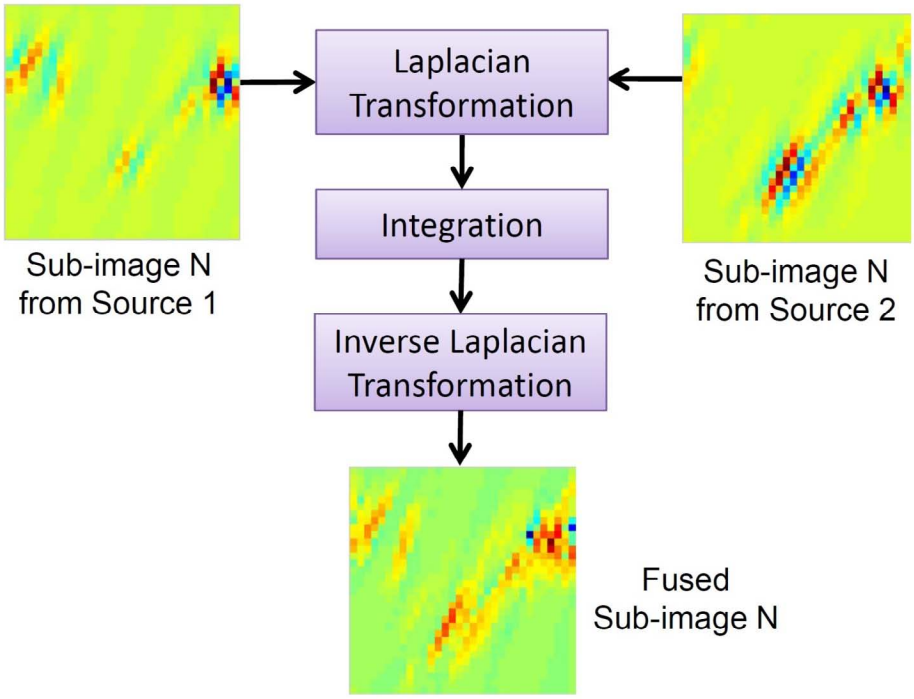


Figure B-3: Fusion scheme for sub-images

This part produces an image of values of [-1, 0, 1] in which -1 means missing pixel, 0 means no rain can occur, and 1 means rain can occur. Conditioning the texture to be only inside the possible rain areas, the fused measurement is produced.

B.2.3 Fused Image Production

This step is a multiplication of the produced texture and the produced shape. Therefore, the regions that have a shape value equal to 1 will have their values from the texture product, the regions that have a shape value equal to zero will be zero in the fused image, and the regions that their shape value is -1 (missing data) will remain missing in the fused image.

B.3 Results

In this section the results of the proposed algorithm are presented and evaluated for both rain detection and rain intensity. Ebert (2007) provides a comprehensive review on the statistics useful for validating the results of a rainfall estimation methods [53]. Here, I

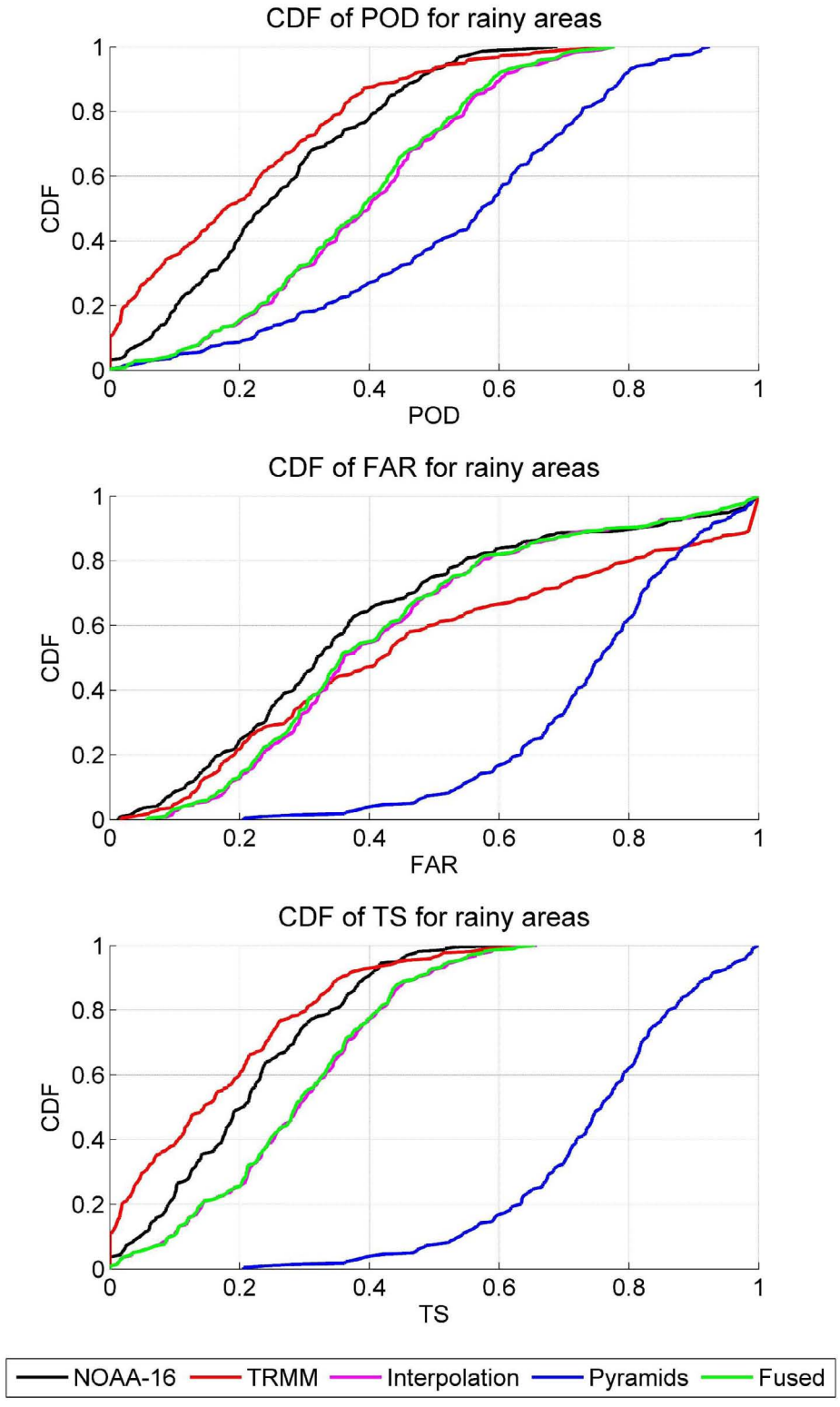


Figure B-4: CDF plot of detection statistics for NOAA-16 and TRMM dataset

selected three statistics to evaluate the capability of the algorithm in detecting rainfall. The statistics are: Probability of Detection (POD), False Alarm Ratio (FAR), and Threat Score (TS). They are defined as following:

$$POD = \frac{hits}{hits + misses} \tag{B.1}$$

$$FAR = \frac{false\ alarms}{hits + false\ alarms} \tag{B.2}$$

$$TS = \frac{hits}{hits + misses + false\ alarms} \tag{B.3}$$

in which *hits* is defined as the number of instances that both of the true measurement and the proposed measurement detect rain; *misses* is the number of instances that the true detects rain, but the proposed one does not; and *false alarms* is defined as the number of instances that the true does not detect rain, but the proposed one does. Note that I have only considered the detection of rain.

The preferred values for each of these statistics are POD = 1, FAR = 0, and TS = 1. Figure B-4 shows the Cumulative Distribution Function (CDF) of the above mentioned statistics for NOAA-16 and TRMM dataset (with the characteristics described in section B.1.1) using the developed algorithm in this paper, the single interpolation method, and the pyramids method. As this figure shows, the pyramids method works well in improving the POD and TS, but at the same time it worsens the FAR to a high degree. Therefore, it is detecting the rainy areas by overestimating the rainy regions and giving too much false alarm that is not acceptable. The interpolation method seems to have a well matched result with the fused algorithm. They both improve the POD and TS with respect to the two input measurements (NOAA-16 and TRMM). Moreover, they provide a reasonable FAR which is somehow similar to the two input images. However, the difference between the interpolation method and the fused method becomes clear by comparing the Probability Distribution Function (PDF) and CDF of the rain intensity of different methods. This is an important comparison because a method that is only capable of detecting rainfall and does not provide reasonable intensities will not be useful. Figure B-5 shows the PDF and CDF of rain

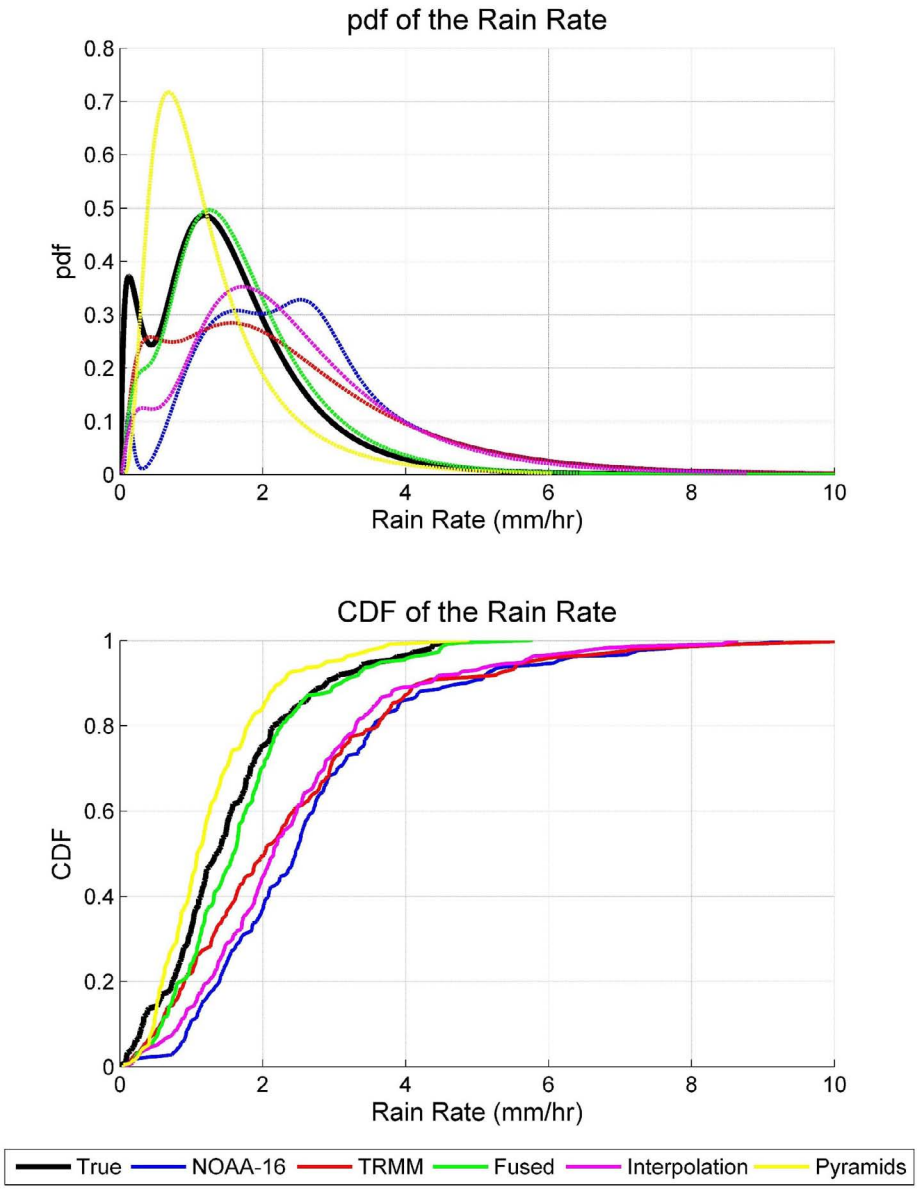


Figure B-5: PDF and CDF of rain intensity for NOAA-16 and TRMM dataset

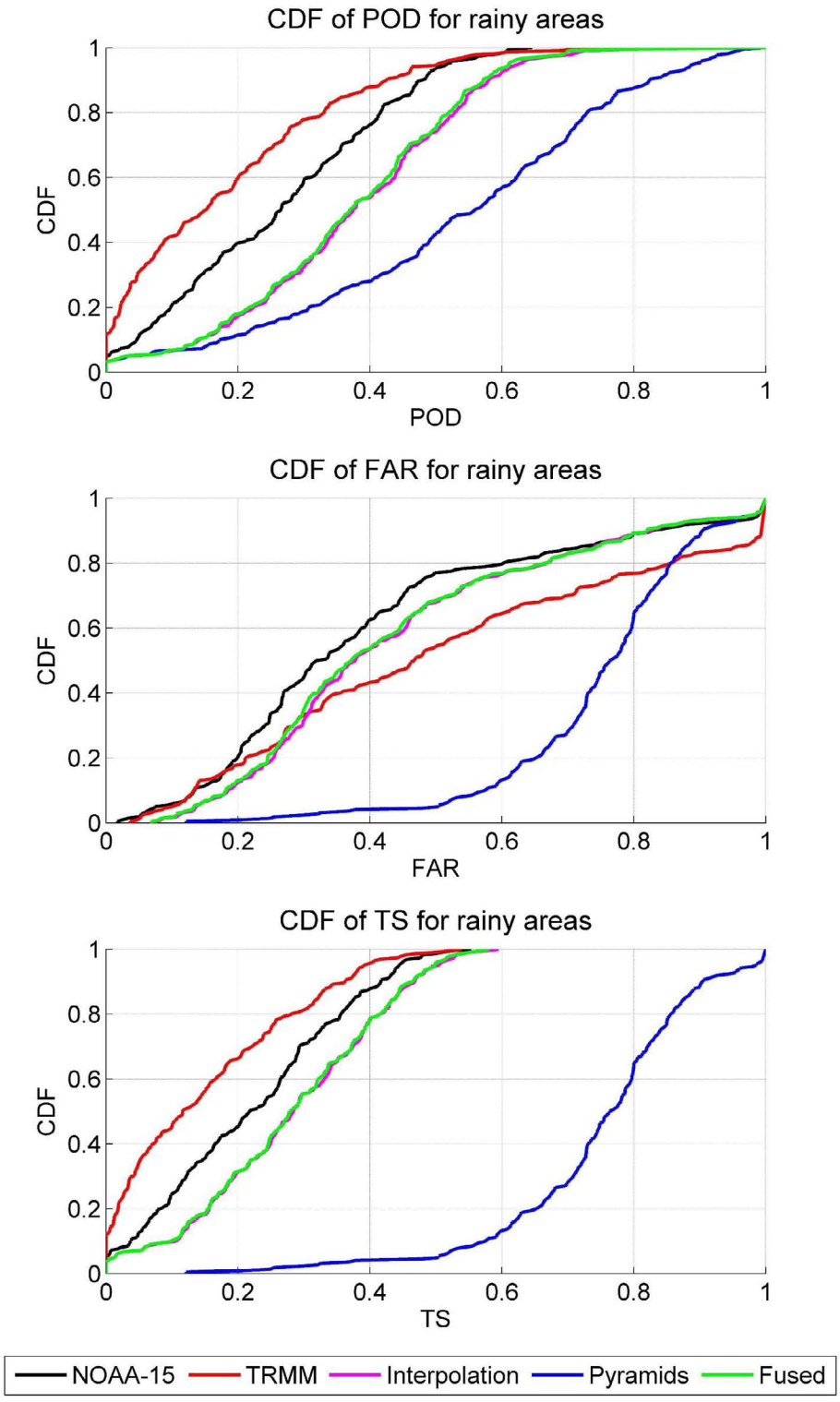


Figure B-6: CDF plot of detection statistics for NOAA-15 and TRMM dataset

intensity for NOAA-16 and TRMM dataset. Note, these plots are for the pixels that were non-missing in all the measurements (i.e. True, NOAA-16, TRMM, Fused, Interpolation, Pyramids.)

As can be seen in Figure B-5, the PDF and CDF of the fused measurement match well with the ones for true measurement, which is the surface-based radar measurement as discussed in section B.1.1. In addition, to test which of these CDFs are from the same distribution as the true one, I run the Kolmogorov-Smirnov test [124]. This test evaluates the hypothesis that two given datasets are from the same continuous distribution. The result with 5% significance level showed that only the fused measurement can be from the same distribution as the true one. This shows that use of the fused algorithm, proposed in this chapter, improves each of the pyramids and interpolation methods and provides better estimation with respect to both rain detection and rain intensity.

In order to further test the method, I run the method on another dataset which was the NOAA-15 and TRMM dataset (with the characteristics described in section B.1.1.) The results are illustrated in Figures B-6 and B-7. These two figures reveal that the proposed fusion algorithm produces promising results in this case too. The Kolmogorov-Smirnov test again showed that only the fused measurement CDF can be from the same distribution as the true one, at 5% significance level.

Figures B-8 and B-9 show four samples of the sources, fused, and true images of rainfall intensity in NOAA-16 and NOAA-15 dataset. The purple color in these images represents the missing regions in the images.

B.4 Discussion

This chapter explored the possibility of using a multi-scale imagery technique in accordance with an interpolation scheme to merge two satellite measurements of rainfall and produce a better estimation of rain intensity with less missing points in the fused image. The proposed algorithm improves both rain intensity and rain detection of the two input images. Moreover, it was shown that only the results of the rain intensity from the fused image can be considered to be from the same distribution as the true measurement.

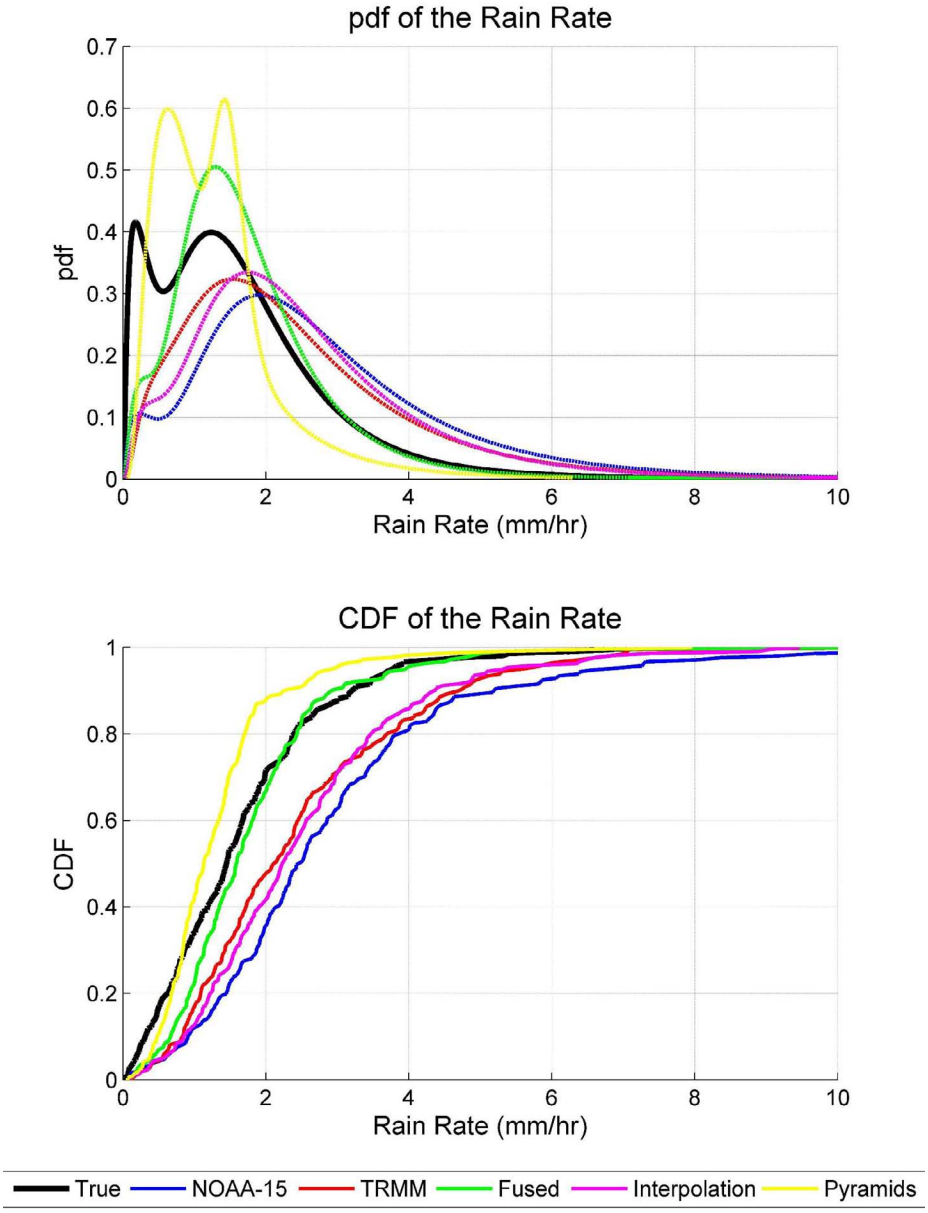


Figure B-7: PDF and CDF of rain intensity for NOAA-15 and TRMM dataset

In addition, as can be seen in Figure B-9, there are regions that there is no true measurement, and this algorithm can improve the satellite measurements and provide accurate estimates over them. A few percentage of the Earth surface is covered by ground-based radars to measure rain rate. Therefore, this method can be used across the globe to fuse different satellite measurements and produce better estimations of rainfall.

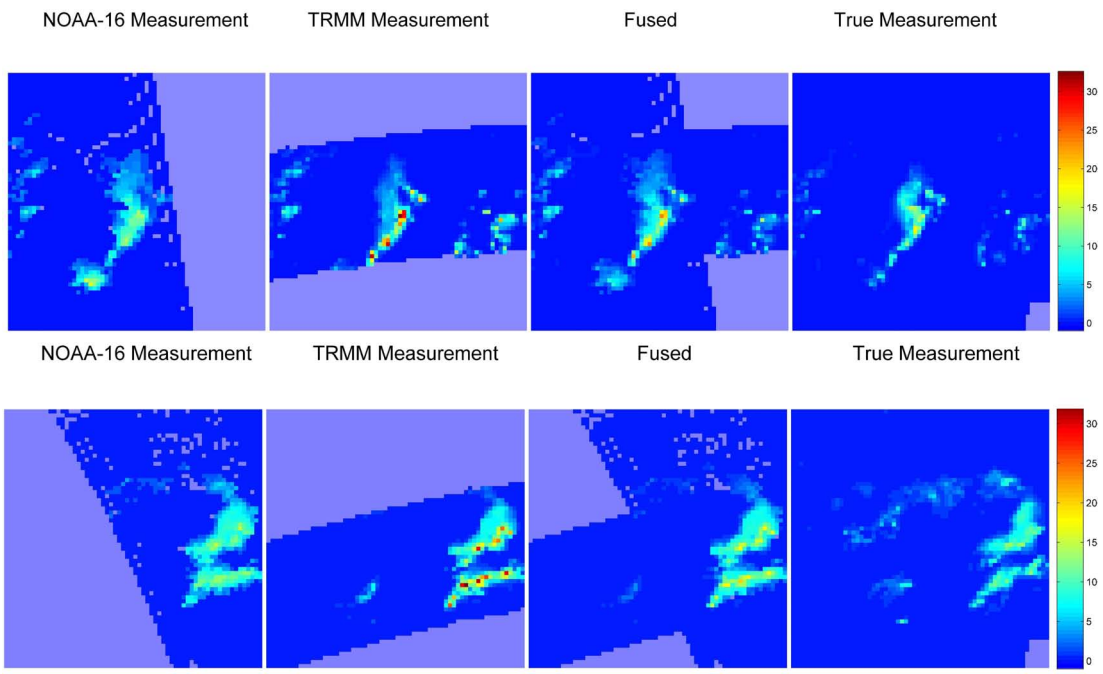


Figure B-8: Sample source, fused, and true measurements of NOAA-16 and TRMM dataset (the scale shows rain intensity in mm/hr)

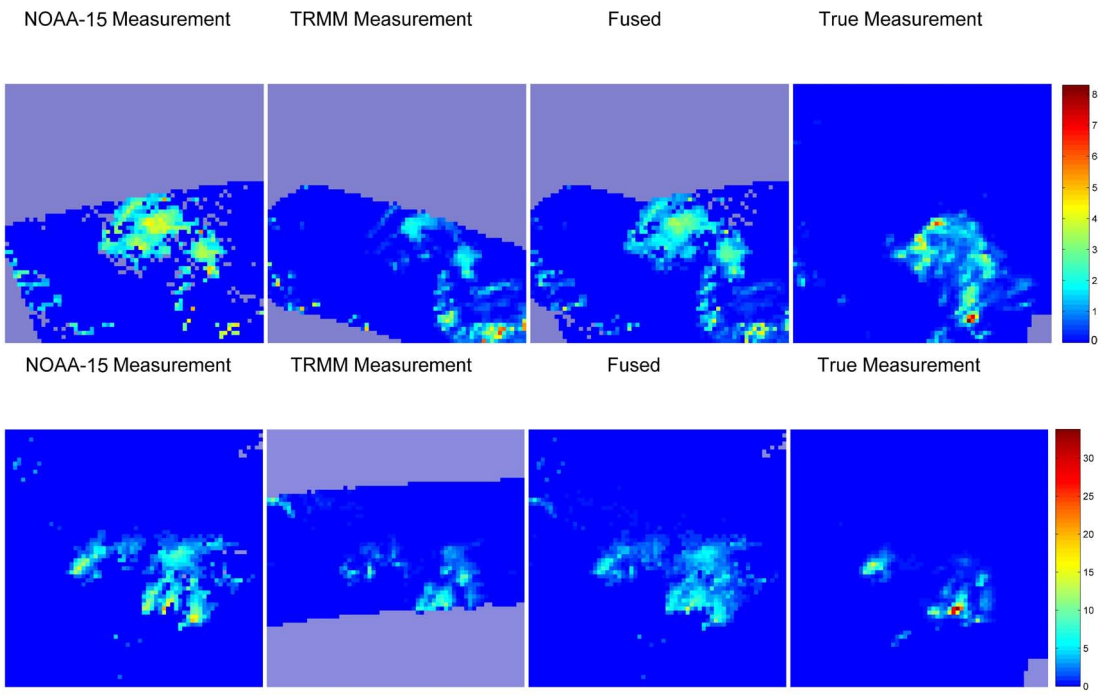


Figure B-9: Sample source, fused, and true measurements of NOAA-15 and TRMM dataset (the scale shows rain intensity in mm/hr)

Furthermore, daily estimates of satellite rainfall can be improved by using this algorithm. There are places on the world that are not covered in the daily passes of one satellite, but they are covered by another satellites which misses some regions in the first one. The proposed algorithm can be applied to these daily measurements to provide full daily rainfall estimates that are useful for water balance studies and rainfall forecasting.

Bibliography

[1] E. H. Adelson, C. H. Anderson, J. R. Bergen, P. J. Burt, and J. M. Ogden. Pyramid methods in image processing. *RCA Engineer*, 29(6):33–41, 1984.

[2] Robert F. Adler, George J. Huffman, Alfred Chang, Ralph Ferraro, Ping-Ping Xie, John Janowiak, Bruno Rudolf, Udo Schneider, Scott Curtis, David Bolvin, Arnold Gruber, Joel Susskind, Philip Arkin, and Eric Nelkin. The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979-present). *J. Hydrometeor*, 4(6):1147–1167, December 2003.

[3] A. AghaKouchak, E. Habib, and A. Bardossy. Modeling radar rainfall estimation uncertainties: Random error model. *Journal of Hydrologic Engineering*, 15(4):265–274, 2010.

[4] A. AghaKouchak, N. Nasrollahi, and E. Habib. Accounting for uncertainties of the TRMM satellite estimates. *Remote Sensing*, 1(3):606–619, 2009.

[5] Amir AghaKouchak, Andras Bardossy, and Emad Habib. Conditional simulation of remotely sensed rainfall data using a non-gaussian v-transformed copula. *Advances in Water Resources*, 33(6):624 – 634, 2010.

[6] N. Ahmed, T. Natarajan, and K.R. Rao. Discrete cosine transform. *Computers, IEEE Transactions on*, C-23(1):90–93, Jan 1974.

[7] H. Akaike. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6):716–723, 1974.

[8] Seyed Hamed Alemohammad, Reza Ardakanian, and Akbar Karimi. A framework for modelling probabilistic uncertainty in rainfall scenario analysis. In *Advances in Hydro-Science and Engineering*, volume VIII, pages 472–479. Nagoya Hydraulic Research Institute for River Basin Management (NHRI), 2008.

[9] Seyed Hamed Alemohammad and Dara Entekhabi. Merging satellite measurements of rainfall using multi-scale imagery technique. *WCRP Open Science Conference*, 2011.

[10] Seyed Hamed Alemohammad, Dara Entekhabi, and Dennis McLaughlin. Evaluation of long-term SSM/I-based precipitation records over land. *Journal of Hydrometeorology*, online, 2014.

[11] Seyed Hamed Alemohammad, Dara Entekhabi, and Dennis McLaughlin. Generating spatially distributed rainfall replicates for ensemble forecasting and estimation. *Water Resources Research*, In Reivew, 2014.

[12] Eric A. Aligo, William A. Gallus, and Moti Segal. Summer rainfall forecast spread in an ensemble initialized with different soil moisture analyses. *Wea. Forecasting*, 22(2):299–314, 2007.

[13] Emmanouil N. Anagnostou. Overview of overland satellite rainfall estimation for hydro-meteorological applications. *Surveys in Geophysics*, 25(5-6):511–537, 2004.

[14] E.N. Anagnostou, V. Maggioni, E.I Nikolopoulos, T. Meskele, F. Hossain, and A Papadopoulos. Benchmarking high-resolution global satellite rainfall products to radar and rain-gauge rainfall estimates. *Geoscience and Remote Sensing, IEEE Transactions on*, 48(4):1667–1683, April 2010.

[15] J. Anderson and S. Anderson. A monte carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Weather Rev.*, 127(12):2741–2748, 1999.

[16] J. L. Anderson. A method for producing and evaluating probabilistic forecasts from ensemble model integrations. *J. Climate*, 9:1518–1530, 1996.

[17] I. Astin. A survey of studies into errors in large scale space-time averages of rainfall, cloud cover, sea surface processes and the earthŠs radiation budget as derived from low orbit satellite instruments because of their incomplete temporal and spatial coverage. *Surv. Geophys.*, 18:385–403, 1997.

[18] Andras Bardossy. Generating precipitation time series using simulated annealing. *Water Resources Research*, 34(7):1737–1744, 1998.

[19] Idar Barstad and Ronald B. Smith. Evaluation of an orographic precipitation model. *J. Hydrometeor*, 6(1):85–99, 2005.

[20] A. Behrangi, K. Hsu, I. Bisher, and S. Sorooshian. Daytime precipitation estimation using bispectral cloud classification system. *J. Appl. Meteor. Climatol.*, 49:1015–1031, 2010.

[21] T. L. Bell. A space-time stochastic-model of rainfall for satellite remote-sensing studies. *J. Geophys. Res.-Atmos.*, 92:9631–9643, 1987.

[22] T. L. Bell, A. Abdullah, R. L. Martin, and G. R. North. Sampling errors for satellite-derived tropical rainfall - monte-carlo study using a space-time stochastic-model. *J. Geophys. Res.-Atmos.*, 95:2195–2205, 1990.

[23] Thomas L. Bell, Prasun K. Kundu, and Christian D. Kummerow. Sampling errors of SSM/I and TRMM rainfall averages: Comparison with error estimates from surface data and a simple model. *J. Appl. Meteor.*, 40(5):938–954, May 2001.

[24] T. J. Bellerby. Satellite rainfall uncertainty estimation using an artificial neural network. *J. Hydrometeor.*, 8:1397–1412, 2007.

[25] T. J. Bellerby and J. Sun. Probabilistic and ensemble representations of the uncertainty in an IR/microwave satellite precipitation product. *J. Hydrometeo.*, 6:1032–1044, 2005.

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

[27] Tim Bellerby, Kuo-lin Hsu, and Soroosh Sorooshian. Lmodel: A satellite precipitation methodology using cloud development modeling. part i: Algorithm construction and calibration. *J. Hydrometeor*, 10(5):1081–1095, 2009.

[28] Wesley Berg, Christian Kummerow, Mathew Sapiano, Nereida Rodrigues-Alvarez, and Fuhzong Weng. A fundamental climate data record of microwave brightness temperature data from 25 years of SSM/I and SSMIS observations. *GEWEX News*, 22:4–5, August 2012.

[29] C.M. Bishop, M. Svensen, and C.K.I. Williams. GTM: The generative topographic mapping. *Neural Computation*, 10(1):215–234, 1998. cited By (since 1996)599.

[30] Marc Bocquet, Carlos A. Pires, and Lin Wu. Beyond gaussian statistical modeling in geophysical data assimilation. *Mon. Wea. Rev.*, 138(8):2997–3023, April 2010.

[31] Ingwer Borg and Patrick J.F. Groenen. *Modern multidimensional scaling: Theory and applications*. Springer, 2005.

[32] R. L. Bras and I. Rodríguez-Iturbe. Rainfall generation: A nonstationary time-varying multidimensional model. *Water Resour. Res.*, 12(3):450–456, 1976.

[33] G. W. Brier. Verification of forecasts expressed in terms of probability. *Mon. Wea. Rev.*, 78:1–3, 1950.

[34] Elisa Brussolo, Jost von Hardenberg, Luca Ferraris, Nicola Rebora, and Antonello Provenzale. Verification of quantitative precipitation forecasts via stochastic downscaling. *J. Hydrometeor*, 9(5):1084–1094, 2008.

[35] Nataliya Bulygina and Hoshin Gupta. How bayesian data assimilation can be used to estimate the mathematical structure of a model. *Stochastic Environmental Research and Risk Assessment*, 24(6):925–937, 2010.

[36] G. Candille and O. Talagrand. Evaluation of probabilistic prediction systems for a scalar variable. *Quarterly Journal of the Royal Meteorological Society*, 131(609):2131–2150, 2005.

[37] S. Chen, P. E. Kirstetter, Y. Hong, J. J. Gourley, Y. D. Tian, Y. C. Qi, Q. Cao, J. Zhang, K. Howard, J. J. Hu, and X. W. Xue. Evaluation of spatial errors of precipitation rates and types from TRMM spaceborne radar over the southern CONUS. *J. Hydrometeor*, 14(6):1884–1896, July 2013.

[38] Sheng Chen, Yang Hong, Jonathan J. Gourley, George J. Huffman, Yudong Tian, Qing Cao, Bin Yong, Pierre-Emmanuel Kirstetter, Junjun Hu, Jill Hardy, Zhe Li, Sadiq I. Khan, and Xianwu Xue. Evaluation of the successive V6 and V7 TRMM multisatellite precipitation analysis over the continental united states. *Water Resources Research*, 49(12):8174–8186, 2013.

[39] Adam J. Clark, William A. Gallus, Ming Xue, and Fanyou Kong. A comparison of precipitation forecast skill between small convection-allowing and large convection-parameterizing ensembles. *Wea. Forecasting*, 24(4):1121–1140, 2009.

[40] Martyn P. Clark and Andrew G. Slater. Probabilistic quantitative precipitation estimation in complex terrain. *J. Hydrometeor*, 7(1):3–22, 2006.

[41] R. R. Coifman and S. Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis*, 21(1):5–30, 2006.

[42] Brian A. Colle, Clifford F. Mass, and Kenneth J. Westrick. Mm5 precipitation verification over the pacific northwest during the 1997âĂŞ99 cool seasons*. *Wea. Forecasting*, 15(6):730–744, 2000.

[43] Jean-Marie Cornuet, Jean-Michel Marin, Antonietta Mira, and Christian P. Robert. Adaptive multiple importance sampling. *Scandinavian Journal of Statistics*, 39(4):798–812, 2012.

[44] P. S. P. Cowpertwait, C. G. Kilsby, and P. E. O'Connell. A space-time neyman-scott model of rainfall: Empirical analysis of extremes. *Water Resources Research*, 38(8):6–1–6–14, 2002.

[45] Trevor F. Cox and Michael A.A. Cox. *Multidimensional scaling*. CRC Press, 2010.

[46] C. M. DeChant and H. Moradkhani. Improving the characterization of initial condition for ensemble streamflow prediction using data assimilation. *Hydrology and Earth System Sciences*, 15(11):3399–3410, 2011.

[47] Caleb M. DeChant and Hamid Moradkhani. Examining the effectiveness and robustness of sequential data assimilation methods for quantification of uncertainty in hydrologic forecasting. *Water Resources Research*, 48(4), 2012.

[48] Caleb M. DeChant and Hamid Moradkhani. Toward a reliable prediction of seasonal forecast uncertainty: Addressing model and initial condition uncertainty with ensemble data assimilation and sequential bayesian combination. *Journal of Hydrology*, In Press(0), 2014.

[49] Jacques W. Delleur and M. Levent Kavvas. Stochastic models for monthly rainfall forecasting and synthetic generation. *J. Appl. Meteor.*, 17(10):1528–1536, October 1978.

[50] P. Demartines and J. Herault. Curvilinear component analysis: A self-organizing neural network for nonlinear mapping of data sets. *IEEE Transactions on Neural Networks*, 8(1):148–154, 1997. cited By (since 1996)293.

[51] Tufa Dinku and Emmanouil N. Anagnostou. TRMM calibration of SSM/I algorithm for overland rainfall estimation. *J. Appl. Meteor. Climatol.*, 45(6):875–886, June 2006.

[52] Jeffrey D. Duda, Xuguang Wang, Fanyou Kong, and Ming Xue. Using varied microphysics to account for uncertainty in warm-season qpf in a convection-allowing ensemble. *Mon. Wea. Rev.*, 142(6):2198–2219, 2014.

[53] E. E. Ebert. Methods for verifying satellite precipitation estimates. In V. Levizzani, P. Bauer, and F. J. Turk, editors, *Measuring precipitation from space*, pages 345–356. Springer, Dordrecht, 2007.

[54] Ardeshir Mohammad Ebtehaj and Efi Foufoula-Georgiou. Adaptive fusion of multisensor precipitation using gaussian-scale mixtures in the wavelet domain. *Journal of Geophysical Research: Atmospheres*, 116(D22), 2011.

[55] Haralambos Feidas and Apostolos Giannakos. Classifying convective and stratiform rain using multispectral infrared meteosat second generation satellite data. *Theoretical and Applied Climatology*, 108(3-4):613–630, 2012.

[56] Luca Ferraris, Simone Gabellani, Nicola Rebora, and Antonello Provenzale. A comparison of stochastic models for spatial rainfall downscaling. *Water Resources Research*, 39(12), 2003.

[57] Ralph Ferraro and Qihang Li. Detailed analysis of the error associated with the rainfall retrieved by the NOAA/NESDIS Special Sensor Microwave/Imager algorithm 2. Rainfall over land. *J. Geophys. Res.*, 107(D23):4680, December 2002.

[58] Ralph R. Ferraro. Special sensor microwave imager derived global rainfall estimates for climatological applications. *Journal of Geophysical Research: Atmospheres*, 102(D14):16715–16735, 1997.

[59] Ralph R. Ferraro and Gerard F. Marks. The development of SSM/I rain-rate retrieval algorithms using ground-based radar measurements. *J. Atmos. Oceanic Technol.*, 12(4):755–770, August 1995.

[60] Ralph R. Ferraro, Eric A. Smith, Wesley Berg, and George J. Huffman. A screening methodology for passive microwave precipitation retrieval algorithms. *J. Atmos. Sci.*, 55(9):1583–1600, May 1998.

[61] R.R. Ferraro, C.D. Peters-Lidard, C. Hernandez, F.J. Turk, F. Aires, C. Prigent, Xin Lin, S. Boukabara, F.A. Furuzawa, K. Gopalan, K.W. Harrison, F. Karbou, Li Li, Chuntao Liu, H. Masunaga, L. Moy, S. Ringerud, G.M. Skofronick-Jackson, Yudong Tian, and Nai-Yu Wang. An evaluation of microwave land surface emissivities over the continental united states to benefit GPM-era precipitation algorithms. *Geoscience and Remote Sensing, IEEE Transactions on*, 51(1):378–398, Jan 2013.

[62] B. Fisher and Wolff D. B. Satellite sampling and retrieval errors in regional monthly rain estimates from TMI, AMSR-E, SSM/I, AMSU-B, and the TRMM PR. *J. Appl. Meteor. Climatol.*, 50:994–1023, 2011.

[63] W. T. Freeman and E. H. Adelson. The design and use of steerable filters. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(9):891–906, 1991.

[64] R. A. Fulton, J. P. Breidenbach, D. J. Seo, and D. A. Miller. The WSR-88D rainfall algorithm. *Weather and Forecasting*, 13:377–395, 1998.

[65] ÃL'tienne Gaborit, FranÃğois Anctil, Vincent Fortin, and GeneviÃĺve Pelletier. On the reliability of spatially disaggregated global ensemble rainfall forecasts. *Hydrological Processes*, 27(1):45–56, 2013.

[66] M. Gebremichael and W. Krajewski. Characterization of temporal sampling error in space-time-averaged rainfall estimates from satellites. *J. Geophys. Res.-Atmos.*, 109:D11110, 2004.

[67] Andrew Gelman, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin. *Bayesian data analysis*. CRC press, 2013.

[68] C. Gomi and Y. Kuzuha. Simulation of a daily precipitation time series using a stochastic model with filtering. *Open Journal of Modern Hydrology*, 3(4):206–213, 2013.

[69] D.I.F. Grimes. An ensemble approach to uncertainty estimation for satellite-based rainfall estimates. In S. Sorooshian, K. L. Hsu, E. Coppola, B. Tomassetti, M. Verdecchia, and G. Visconti, editors, *Hydrological Modeling and the Water Cycle*, pages 145–162. Springer, Berlin, 2008.

[70] Norman C. Grody. Classification of snow cover and precipitation using the special sensor microwave imager. *Journal of Geophysical Research: Atmospheres*, 96(D4):7423–7435, 1991.

[71] Rohit Gupta, V. Venugopal, and Efi Foufoula-Georgiou. A methodology for merging multisensor precipitation estimates based on expectation-maximization and scale-recursive estimation. *Journal of Geophysical Research: Atmospheres*, 111(D2), 2006.

[72] V. Gupta and T. Over. A space-time theory of mesoscale rainfall using random cascades. *J. Geophys. Res.*, 101:26319–26331, 1997.

[73] V. Gupta and E. Waymire. A statistical analysis of mesoscale rainfall as a random cascade. *J. Appl. Meteorol.*, 32(2):251–267, 1993.

[74] E. Habib, A. Henschke, and R. F. Adler. Evaluation of TMPA satellite-based research and real-time rainfall estimates during six tropical-related heavy rainfall events over louisiana, USA. *Atmospheric Research*, 94:373–388, 2009.

[75] Emad Habib, Alemseged Tamiru Haile, Yudong Tian, and Robert J. Joyce. Evaluation of the high-resolution CMORPH satellite rainfall product using dense rain gauge observations and radar-based estimates. *J. Hydrometeor*, 13(6):1784–1798, 2012.

[76] D. K. Hall, V. V. Salomonson, and G. A. Riggs. MODIS/Terra Snow Cover Daily L3 Global 0.05Deg CMG. Version 5, 2006.

[77] Thomas M. Hamill. Interpretation of rank histograms for verifying ensemble forecasts. *Mon. Wea. Rev.*, 129(3):550–560, 2001.

[78] K. A. Hilburn and F. J. Wentz. Mitigating the impact of RADCAL beacon contamination on F15 SSM/I ocean retrievals. *Geophys. Res. Lett.*, 35(18):L18806, September 2008.

[79] Kyle Hilburn. Including temperature effects in the F15 RADCAL correction. Technical Report 051209, Remote Sensing Systems, 2009.

[80] Yang Hong, Kuo-Lin Hsu, Soroosh Sorooshian, and Xiaogang Gao. Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *J. Appl. Meteor.*, 43(12):1834–1853, December 2004.

[81] F. Hossain and E. N. Anagnostou. A two-dimensional satellite rainfall error model. *IEEE Transactions of Geoscience and Remote Sensing*, 44(6):1511–1522, 2006.

[82] Arthur Y. Hou, Ramesh K. Kakar, Steven Neeck, Ardeshir A. Azarbarzin, Christian D. Kummerow, Masahiro Kojima, Riko Oki, Kenji Nakamura, and Toshio Iguchi. The global precipitation measurement mission. *Bull. Amer. Meteor. Soc.*, 95(5):701–722, 2013.

[83] K. Hsu and S. Sorooshian. Satellite-based precipitation measurement using PERSIANN system. In S. Sorooshian, K. L. Hsu, E. Coppola, B. Tomassetti, M. Verdecchia, and G. Visconti, editors, *Hydrological Modeling and the Water Cycle*, pages 27–48. Springer-Verlag, 2008.

[84] George J. Huffman, David T. Bolvin, Eric J. Nelkin, David B. Wolff, Robert F. Adler, Guojun Gu, Yang Hong, Kenneth P. Bowman, and Erich F. Stocker. The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeor*, 8(1):38–55, February 2007.

[85] Behnam Jafarpour, VivekK. Goyal, DennisB. McLaughlin, and WilliamT. Freeman. Compressed history matching: Exploiting transform-domain sparsity for regularization of nonlinear dynamic data integration problems. *Mathematical Geosciences*, 42(1):1–27, 2010.

[86] Behnam Jafarpour, Dennis B McLaughlin, et al. Reservoir characterization with the discrete cosine transform. *SPE Journal*, 14(01):182–201, 2009.

[87] R. J. Joyce, P. Xie, Y. Yarosh, J. E. Janowiak, and P. A. Arkin. CMORPH: A "morphing" approach for high resolution precipitation product generation. In M. Gebremichael and F. Hossain, editors, *Satellite Rainfall Applications for Surface Hydrology*, pages 23–37. Springer, 2010.

[88] Robert Joyce and Phillip A. Arkin. Improved estimates of tropical and subtropical precipitation using the goes precipitation index. *J. Atmos. Oceanic Technol.*, 14(5):997–1011, October 1997.

[89] Robert J. Joyce, John E. Janowiak, Phillip A. Arkin, and Pingping Xie. CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeor*, 5(3):487–503, June 2004.

[90] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Journal of Fluids Engineering*, 82(1):35–45, 1960.

[91] Dmitri Kavetski, George Kuczera, and Stewart W. Franks. Bayesian analysis of input uncertainty in hydrological modeling: 1. theory. *Water Resources Research*, 42(3), 2006.

[92] C. Kidd. Satellite rainfall climatology: a review. *International Journal of Climatology*, 21(9):1041–1066, 2001.

[93] C. Kidd and V. Levizzani. Status of satellite precipitation retrievals. *Hydrology and Earth System Sciences*, 15(4):1109–1116, 2011.

[94] C. Kidd, F. J. Tapiador, V. Sanderson, and D. Kniveton. The university of birmingham global rainfall algorithms. In V. Levizzani, P. Bauer, and F. J. Turk, editors, *Measuring precipitation from space*, pages 255–267. Springer, 2007.

[95] Chris Kidd, Dominic R. Kniveton, Martin C. Todd, and Tim J. Bellerby. Satellite rainfall estimation using combined passive microwave and infrared algorithms. *J. Hydrometeor*, 4(6):1088–1104, December 2003.

[96] Chris Kidd, Vincenzo Levizzani, and Peter Bauer. A review of satellite meteorology and climatology at the start of the twenty-first century. *Progress in Physical Geography*, 33(4):474–489, 2009.

[97] Chris Kidd, Vincenzo Levizzani, and Sante Laviola. Quantitative precipitation estimation from earth observation satellites. In *Rainfall: State of the Science*, pages 127–158. American Geophysical Union, 2013.

[98] Y. J. Kim, S. Kimball, J. Glassy, and K. C. McDonald. MEaSUREs global record of daily landscape freeze/thaw status, version 2,, 2010, Updated 2012.

[99] Youngwook Kim, J.S. Kimball, K.C. McDonald, and J. Glassy. Developing a global data record of daily landscape freeze/thaw status using satellite passive microwave remote sensing. *Geoscience and Remote Sensing, IEEE Transactions on*, 49(3):949–960, 2011.

[100] Michael D King and Reynold Greenstone, editors. *EOS reference handbook: a guide to NASA's Earth Science Enterprise and the Earth Observing System*. EOS Project Science Office, NASA/Goddard Space Flight Center, Greenbelt, MD, 1999.

[101] Pierre-Emmanuel Kirstetter, Y. Hong, J. J. Gourley, M. Schwaller, W. Petersen, and J. Zhang. Comparison of TRMM 2a25 products, version 6 and version 7, with NOAA/NSSL ground radar-based national mosaic QPE. *J. Hydrometeor*, 14(2):661–669, 2012.

[102] Pierre-Emmanuel Kirstetter, Nicolas Viltard, and Marielle Gosset. An error model for instantaneous satellite rainfall estimates: evaluation of BRAIN-TMI over west africa. *Quarterly Journal of the Royal Meteorological Society*, 139(673):894–911, 2013.

[103] T. N. Krishnamurti, Sajani Surendran, D. W. Shin, Ricardo J. Correa-Torres, T. S. V. Vijaya Kumar, Eric Williford, Chris Kummerow, Robert F. Adler, Joanne Simpson, Ramesh Kakar, William S. Olson, and F. Joseph Turk. Real-time multianalysis-multimodel superensemble forecasts of precipitation using TRMM and SSM/I products. *Mon. Wea. Rev.*, 129(12):2861–2883, December 2001.

[104] T. Kubota, S. Shige, H. Hashizume, K. Aonashi, N. Takahashi, S. Seto, Y.N. Takayabu, T. Ushio, K. Nakagawa, K. Iwanami, M. Kachi, and K. Okamoto. Global precipitation map using satellite-borne microwave radiometers by the GSMaP project: Production and validation. *Geoscience and Remote Sensing, IEEE Transactions on*, 45(7):2259 –2275, july 2007.

[105] Robert J. Kuligowski and Ana P. Barros. Blending multiresolution satellite data with application to the initialization of an orographic precipitation model. *J. Appl. Meteor.*, 40(9):1592–1606, September 2001.

[106] P. Kumar and E. Foufoula-Georgiou. A multicomponent decomposition of spatial rainfall fields: 1. segregation of large and small scale features using wavelet transforms. *Water Resour. Res.*, 29(8):2515–2532, 1993.

[107] Praveen Kumar and Efi Foufoula-Georgiou. Characterizing multiscale variability of zero intermittency in spatial rainfall. *J. Appl. Meteor.*, 33(12):1516–1525, 1994.

[108] Christian Kummerow, William Barnes, Toshiaki Kozu, James Shiue, and Joanne Simpson. The tropical rainfall measuring mission (trmm) sensor package. *J. Atmos. Oceanic Technol.*, 15(3):809–817, June 1998.

[109] Hyun-Han Kwon, Upmanu Lall, and Abedalrazq F. Khalil. Stochastic simulation model for nonstationary time series using an autoregressive wavelet decomposition: Applications to rainfall and temperature. *Water Resour. Res.*, 43(5):W05407, 2007.

[110] Etienne Leblois and Jean-Dominique Creutin. Space-time simulation of intermittent rainfall with prescribed advection field: Adaptation of the turning band method. *Water Resources Research*, 49(6):3375–3387, 2013.

[111] B. G. Lee, R. T. Chin, and D. W. Martin. Automated rain-rate classification of satellite images using statistical pattern recognition. *IEEE Transactions of Geoscience and Remote Sensing*, pages 315–324, 1985.

[112] I. M. Lensky and V. Levizzani. Estimation of precipitation from space-based platforms. In S. Michaelides, editor, *Precipitation: Advances in Measurement, Estimation and Prediction*, pages 195–217. Springer-Verlag, Berlin, 2008.

[113] S. Lespinats, B. Fertil, P. Villemain, and J. Herault. RankVisu: Mapping from the neighborhood network. *Neurocomputing*, 72(13âĂŞ15):2964 – 2978, 2009. Hybrid Learning Machines (HAIS 2007) / Recent Developments in Natural Computation (ICNC 2007).

[114] S. Lespinats, M. Verleysen, A Giron, and B. Fertil. DD-HDS: A method for visualization and exploration of high-dimensional data. *Neural Networks, IEEE Transactions on*, 18(5):1265–1279, Sept 2007.

[115] Q. Li, R. Ferraro, and N. Grody. Global distribution of SSM/I rainfall retrieval error. In *Ninth Conference on Satellite Meteorology, American Meteorological Society*, 1998.

[116] Ying Lin and Kenneth E. Mitchell. The NCEP Stage II/IV hourly precipitation analyses: development and applications. In *19th Conf. on Hydrology, American Meteorological Society, San Diego, CA*, page Paper 1.2, 2005.

[117] Changhai Liu, Kyoko Ikeda, Gregory Thompson, Roy Rasmussen, and Jimy Dudhia. High-resolution simulations of wintertime precipitation in the colorado headwaters region: Sensitivity to physics parameterizations. *Monthly Weather Review*, 139(11):3533–3553, 2011.

[118] Z. Liu, K. Tsukada, K. Hanasaki, Y. K. Ho, and Y. P. Dai. Image fusion by using steerable pyramid. *Pattern Recognition Letters*, 22:929–939, 2001.

[119] S. Lovejoy and D. Schertzer. Multifractals, universality classes, satellite and radar measurements of clouds and rain. *J. Geophys. Res.*, 95:2021–2034, 1990.

[120] Viviana Maggioni, Mathew R. P. Sapiano, Robert F. Adler, Yudong Tian, and George J. Huffman. An error model for uncertainty quantification in high-time-resolution precipitation products. *J. Hydrometeor*, 15(3):1274–1292, March 2014.

[121] D. Marsan, D. Schertzer, and S. Lovejoy. Casual space-time multifractal processes. predictability and forecasting of rainfall fields. *J. Geophys. Res.*, 26:333–346, 1996.

[122] F. S. Marzano, D. Cimini, and F. J. Turk. Multivariate probability matching for microwave infrared combined rainfall algorithm (micra). In V. Levizzani, P. Bauer, and F. J. Turk, editors, *Measuring precipitation from space*, pages 269–279. Springer, 2007.

[123] I. Mason. A model for assessment of weather forecasts. *Aust. Meteorol. Mag.*, 30:291–303, 1982.

[124] F. J. Massey. The kolmogorov-smirnov test for goodness of fit. *Journal of the American Statistical Association*, 46(253):68–788, 1951.

[125] Jiri Matousek. On the distortion required for embedding finite metric spaces into normed spaces. *Israel Journal of Mathematics*, 93(1):333–344, 1996.

[126] Peter S. Maybeck. *Stochastic models, estimation, and control*, volume 3. Academic Press, 1982.

[127] Jeffrey R. McCollum, Witold F. Krajewski, Ralph R. Ferraro, and Mamoudou B. Ba. Evaluation of biases of satellite rainfall estimation algorithms over the continental United States. *J. Appl. Meteor.*, 41(11):1065–1080, November 2002.

[128] D. McLaughlin, Y. Zhou, D. Entekhabi, and V. Chatdarong. Computational issues for large-scale land surface data assimilation problems. *J. Hydrometeor.*, 7:494–510, 2006.

[129] Dennis McLaughlin. An integrated approach to hydrologic data assimilation: interpolation, smoothing, and filtering. *Advances in Water Resources*, 25(8-12):1275 – 1286, 2002.

[130] Dennis McLaughlin. Data assimilation. In Eni G. Njoku, editor, *Encyclopedia of Remote Sensing*, pages 131–134. Springer, 2014.

[131] Merab Menabde, Alan Seed, Daniel Harris, and Geoff Austin. Self-similar random fields and rainfall simulation. *Journal of Geophysical Research: Atmospheres*, 102(D12):13509–13515, 1997.

[132] S. W. Miller, P. A. Arkin, and R. Joyce. A combined microwave/infrared rain rate algorithm. *International Journal of Remote Sensing*, 22(17):3285–3307, 2001.

[133] Hamid Moradkhani, Caleb M. DeChant, and Soroosh Sorooshian. Evolution of ensemble data assimilation for uncertainty quantification using the particle filter-markov chain monte carlo method. *Water Resources Research*, 48(12), 2012.

[134] M. L. Morrissey and J. S. Green. Uncertainty analysis of satellite rainfall algorithms over the tropical pacific. *J. Geo Res*, 103:19569–19576, 1998.

[135] Tarek A. El Moselhy and Youssef M. Marzouk. Bayesian inference with optimal maps. *Journal of Computational Physics*, 231(23):7815 – 7850, 2012.

[136] W.A. Muller, C. Appenzeller, F. J. Doblas-Reyes, and M. A. Liniger. A debiased ranked probability skill score to evaluate probabilistic ensemble forecasts with small ensemble sizes. *J. Climate*, 18:1513–1523, 2005.

[137] A. H. Murphy. Forecast verification: its complexity and dimensionality. *Mon. Wea. Rev.*, 119:1590–1601, 1991.

[138] Mohammad Reza Najafi, Hamid Moradkhani, and Thomas C. Piechota. Ensemble streamflow prediction: Climate signal weighting methods vs. climate forecast system reanalysis. *Journal of Hydrology*, 442-443(0):105 – 116, 2012.

[139] Deborah K. Nykanen and Daniel Harris. Orographic influences on the multiscale statistical properties of precipitation. *Journal of Geophysical Research: Atmospheres*, 108(D8), 2003.

[140] F. Oriani, J. Straubhaar, P. Renard, and G. Mariethoz. Simulation of rainfall time-series from different climatic regions using the direct sampling technique. *Hydrology and Earth System Sciences Discussions*, 11(3):3213–3247, 2014.

[141] Athanasios Paschalis, Peter Molnar, Simone Fatichi, and Paolo Burlando. A stochastic model for high-resolution space-time precipitation simulation. *Water Resour. Res.*, 49(12):8400–8417, 2013.

[142] A. Pathirana and S. Herath. Multifractal modelling and simulation of rain fields exhibiting spatial heterogeneity. *Hydrology and Earth System Sciences*, 6(4):695–708, 2002.

[143] G.G.S. Pegram and A.N. Clothier. High resolution space-time modelling of rainfall: the "string of beads" model. *Journal of Hydrology*, 241:26 – 41, 2001.

[144] Grant W. Petty and Witold F. Krajewski. Satellite estimation of precipitation over land. *Hydrological Sciences Journal*, 41(4):433–451, 1996.

[145] J. Portilla and E. P. Simoncelli. A parametric texture model based on joint statistics of complex wavelet coefficients. *Int. J. of Comp. Vision*, 40(1):49–71, 2000.

[146] Anand Rajaraman and Jeffrey David Ullman. *Mining of massive datasets*. Cambridge University Press, 2011.

[147] Benjamin Renard, Dmitri Kavetski, Etienne Leblois, Mark Thyer, George Kuczera, and Stewart W. Franks. Toward a reliable decomposition of predictive uncertainty in hydrological modeling: Characterizing rainfall errors using conditional simulation. *Water Resources Research*, 47(11), 2011.

[148] C. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer-Verlag, New York, 2nd edition, 2004.

[149] N. M. Roberts and H. W. Lean. Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Weather Rev.*, 136:78–97, 2008.

[150] RÃl'my Roca, Philippe Chambon, Isabelle Jobard, Pierre-Emmanuel Kirstetter, Marielle Gosset, and Jean Claude Berges. Comparing satellite and surface rainfall products over west africa at meteorologically relevant scales during the AMMA campaign using error estimates. *J. Appl. Meteor. Climatol.*, 49(4):715–731, 2009.

[151] S.T. Roweis and L.K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326, 2000. cited By (since 1996)5003.

[152] J.W. Sammon. A nonlinear mapping for data structure analysis. *Computers, IEEE Transactions on*, C-18(5):401–409, May 1969.

[153] M. R. P. Sapiano and P. A. Arkin. An intercomparison and validation of high-resolution satellite precipitation estimates with 3-hourly gauge data. *J. Hydrometeor*, 10(1):149–166, 2009.

[154] Marc Schleiss, Sabine Chamoun, and Alexis Berne. Stochastic simulation of intermittent rainfall using the concept of dry drift. *Water Resources Research*, 50(3):2329–2349, 2014.

[155] David M. Schultz, W. James Steenburgh, R. Jeffrey Trapp, John Horel, David E. Kingsmill, Lawrence B. Dunn, W. David Rust, Linda Cheng, Aaron Bansemer, Justin Cox, John Daugherty, David P. Jorgensen, Jose Meitin, Les Showell, Bradley F. Smull, Keli Tarp, and Marilu Trainor. Understanding utah winter storms: The intermountain precipitation experiment. *Bull. Amer. Meteor. Soc.*, 83(2):189–210, 2002.

[156] Y. Shen, A. Xiong, Y. Wang, and P. Xie. Performance of high-resolution satellite precipitation products over china. *J. Geophys. Res.*, 115:D02114, 2010.

[157] M. Sivapalan and Eric F. Wood. A multidimensional model of nonstationary space-time rainfall at the catchment scale. *Water Resources Research*, 23(7):1289–1299, 1987.

[158] Dimitrios Stampoulis and Emmanouil N. Anagnostou. Evaluation of global satellite rainfall products over continental europe. *J. Hydrometeor*, 13(2):588–603, 2012.

[159] M. Steiner, T. L. Bell, Y. Zhang, and E. F. Wood. Comparison of two methods for estimating the sampling-related uncertainty of satellite rainfall averages based on a large radar dataset. *J. Climate*, 16:3759–3778, 2003.

[160] David J Stensrud. *Parameterization schemes: keys to understanding numerical weather prediction models*. Cambridge University Press, 2007.

[161] Graeme L. Stephens and Christian D. Kummerow. The remote sensing of clouds and precipitation from space: A review. *J. Atmos. Sci.*, 64(11):3742–3765, November 2007.

[162] M. T. Stoelinga and et al. Improvement of microphysical parameterization through observational verification experiment. *Bull. Amer. Meteor. Soc.*, 84:1807–1826, 2003.

[163] P. N. Tan, M. Steinbach, and V. Kumar. *Introduction to Data Mining*. Addison-Wesley, 2005.

[164] Ling Tang, Yudong Tian, and Xin Lin. Validation of precipitation retrievals over land from satellite-based passive microwave sensors. *Journal of Geophysical Research: Atmospheres*, 119(8):4546–4567, 2014.

[165] A. Tassa, S. Di Michele, A. Mugnai, F. S. Marzano, P. Bauer, and J. P. V. P. Baptista. Modeling uncertainties for passive microwave precipitation retrieval: evaluation of a case study. *IEEE Transactions on Geoscience and Remote Sensing*, 44:78–89, 2006.

[166] Joshua B. Tenenbaum, Vin de Silva, and John C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.

[167] C. K. Teo and D. I. F. Grimes. Stochastic modelling of rainfall from satellite data. *Journal of Hydrology*, 346:33–50, 2007.

[168] N. Thomas, B. Thies, A. Turek, J. Bendix, and A. Kokhanovsky. Operational discrimination of raining from non-raining clouds in mid-latitudes using multispectral satellite data. In S. Michaelides, editor, *Precipitation: Advances in Measurement, Estimation and Prediction*, pages 171–194. Springer-Verlag, Berlin, 2008.

[169] Yudong Tian, George J. Huffman, Robert F. Adler, Ling Tang, Mathew Sapiano, Viviana Maggioni, and Huan Wu. Modeling errors in daily precipitation measurements: Additive or multiplicative? *Geophys. Res. Lett.*, 40(10):2060–2065, 2013.

[170] Yudong Tian and Christa D. Peters-Lidard. Systematic anomalies over inland water bodies in satellite-based precipitation estimates. *Geophysical Research Letters*, 34(14), 2007.

[171] Yudong Tian and Christa D. Peters-Lidard. A global map of uncertainties in satellite-based precipitation measurements. *Geophysical Research Letters*, 37(24), 2010.

[172] Yudong Tian, Christa D. Peters-Lidard, John B. Eylander, Robert J. Joyce, George J. Huffman, Robert F. Adler, Kuo-lin Hsu, F. Joseph Turk, Matthew Garcia, and Jing Zeng. Component analysis of errors in satellite-based precipitation estimates. *Journal of Geophysical Research: Atmospheres*, 114(D24), 2009.

[173] Martin C. Todd, Chris Kidd, Dominic Kniveton, and Tim J. Bellerby. A combined satellite infrared and passive microwave technique for estimation of small-scale rainfall. *J. Atmos. Oceanic Technol.*, 18(5):742–755, May 2001.

[174] Warren S. Torgerson. Multidimensional scaling: I. theory and method. *Psychometrika*, 17(4):401–419, 1952.

[175] F. J. Turk, G. Rohaly, J. Hawkins, E. A. Smith, F. S. Marzano, A. Mugnai, and V. Levizzani. Meteorological applications of precipitation estimation from combined SSM/I, TRMM and geostationary satellite data. In P. Pampaloni and S. Paloscia, editors, *Microwave Radiometry and Remote Sensing of the Earth's Surface and Atmosphere*, pages 353–364. VSP Int. Sci. Publisher, Utrecht (The Netherlands), 2000.

[176] Fawwaz T Ulaby and David G Long. *Microwave Radar and Radiometric Remote Sensing*. University of Michigan Press, 2014.

[177] Juan B. Valdes, Ignacio Rodriguez-Iturbe, and Vijay K. Gupta. Approximations of temporal rainfall from a multidimensional model. *Water Resources Research*, 21(8):1259–1270, 1985.

[178] Laurens van der Maaten, Eric Postma, and Jaap van den Herik. Dimensionality reduction: A comparative review. Technical Report TiCC TR 2009, Tilburg University, 2009.

[179] G. A. Vicente, R. A. Scofield, and W. P. Menzel. The operational GOES infrared rainfall estimation technique. *Bulletin of the American Meteorological Society*, 79:1883–1898, 1998.

[180] Daniel Vila, Cecilia Hernandez, Ralph Ferraro, and Hilawe Semunegus. The performance of hydrological monthly products using SSM/I-SSMI/S sensors. *J. Hydrometeor*, 14(1):266–274, July 2012.

[181] Gabriele Villarini, Witold F. Krajewski, Grzegorz J. Ciach, and Dale L. Zimmerman. Product-error-driven generator of probable rainfall conditioned on WSR-88D precipitation estimates. *Water Resour. Res.*, 45(1):W01404, 2009.

[182] Gabriele Villarini, Witold F. Krajewski, and James A. Smith. New paradigm for statistical validation of satellite precipitation estimates: Application to a large sample of the TMPA 0.25-deg 3-hourly estimates over oklahoma. *Journal of Geophysical Research: Atmospheres*, 114(D12), 2009.

[183] Jasper A. Vrugt and Mojtaba Sadegh. Toward diagnostic model calibration and evaluation: Approximate bayesian computation. *Water Resources Research*, 49(7):4335–4345, 2013.

[184] H. S. Wheater, V. S. Isham, D. R. Cox, R. E. Chandler, A. Kakou, P. J. Northrop, L. Oh, C. Onof, and I. Rodriguez-Iturbe. Spatial-temporal rainfall fields: modelling and statistical aspects. *Hydrology and Earth System Sciences*, 4(4):581–601, 2000.

[185] Christopher K. Wikle and L. Mark Berliner. A bayesian tutorial for data assimilation. *Physica D: Nonlinear Phenomena*, 230(1-2):1 – 16, 2007. Data Assimilation.

[186] D. S. Wilks. The minimum spanning tree histogram as a verification tool for multidimensional ensemble forecasts. *Mon. Wea. Rev.*, 132:1329–1340, 2004.

[187] R. Wojcik, D. McLaughlin, A. Konings, and D. Entekhabi. Conditioning stochastic rainfall replicates on remote sensing data. *IEEE Transactions of Geoscience and Remote Sensing*, 47(8):2436–2449, 2009.

[188] Rafal Wojcik, Dennis McLaughlin, Seyed Hamed Alemohammad, and Dara Entekhabi. Ensemble-based characterization of uncertain environmental features. *Advances in Water Resources*, 70(0):36 – 50, 2014.

[189] David B. Wolff and Brad L. Fisher. Assessing the relative performance of microwave-based satellite rain-rate retrievals using TRMM ground validation data. *J. Appl. Meteor. Climatol.*, 48(6):1069–1099, June 2009.

[190] P. Xie, P. A. Arkin, and J. E. Janowiak. CMAP: The CPC merged analysis of precipitation. In V. Levizzani, P. Bauer, and F. J. Turk, editors, *Measuring precipitation from space*, pages 319–328. Springer, 2007.

[191] Pingping Xie and Phillip A. Arkin. Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.*, 78(11):2539–2558, November 1997.

[192] Pingping Xie, John E. Janowiak, Phillip A. Arkin, Robert Adler, Arnold Gruber, Ralph Ferraro, George J. Huffman, and Scott Curtis. GPCP pentad precipitation analyses: An experimental dataset based on gauge observations and satellite estimates. *J. Climate*, 16(13):2197–2214, 2003.

[193] Zhi-Yong Yin, Xueqin Zhang, Xiaodong Liu, Mike Colella, and Xiaoling Chen. An assessment of the biases of satellite rainfall estimates over the Tibetan plateau and correction methods based on topographic analysis. *J. Hydrometeor*, 9(3):301–326, June 2008.