UNIVERSITY OF CALIFORNIA,

IRVINE

Improving warm rainfall detection and rainfall estimation of a multiple satellite-based rainfall retrieval algorithm

DISSERTATION

submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in Civil Engineering

by

Negar Karbalaee

Dissertation Committee:

Professor Kuolin Hsu, Chair

Professor Soroosh Sorooshian

Professor Alexander Ihler

2017

DEDICATION

To

My best friend and husband Ali

For his unconditional love and support

My Mother and Father

The reason of what I become today

My Younger brother

Who I love him more than everything in this world

This dissertation would not have been possible
without their unconditional love and support

TABLE OF CONTENTS

LIST OF FIGURES vi

LIST OF TABLES xi

ACKNOWLEDGMENSTS xiii

CURRCULUM VITAE xv

ABSTRACT OF THE DISSERTATION xviii

CHAPTER 1: Introduction to current global precipitation estimation products 1

1.1. Importance of precipitation estimation 1

1.2. Precipitation estimation/measurement 1

1.3. Satellite-based precipitation estimation 2

1.4. Research Motivation 4

1.5. Objectives of this dissertation 6

CHAPTER 2: PERSIANN-CCS precipitation algorithm 8

2.1. Algorithm framework 8

2.2. Algorithm strengths and potentials for improvement 9

CHAPTER 3: Using PMW precipitation data to reduce the bias of PERSIANN-CCS estimation 11

3.1. The role of multi-satellite data in precipitation algorithms 11

3.2. Methodology 13

3.3. Model Training 16
3.4. Results and Discussion 21
3.4.1. Evaluating the cumulative distribution function during validation year 22
3.4.2. Validation over the globe 24
3.4.3. Validation of the CONUS 35
CHAPTER 4: Improving warm-clouds rainfall detection in PERSIANN-CCS 41
4.1. Introduction 41
4.2. Extending cloud top brightness temperature 41
4.3. Training the algorithm 42
4.4. Results and Discussion 45
CHAPTER 5: Improving rainfall estimation of PERSIANN-CCS using probabilistic approach 49
5.1. PERSIANN-CCS rainfall estimation method 49
5.2. Using probabilistic estimation instead of exponential regression model 51
5.3. Methodology 51
5.4. Algorithm Training 53
5.5. Results and discussions 55
CHAPETR 6: Summary, Conclusion and Future work 82
6.1. Summary and Conclusions 82
6.2. Future Work 86
APPENDIX A 87

7. References 89

LIST OF FIGURES

Figure 2.1: PERSIANN-CCS algorithm framework (Hong et al. 2004) 9

Figure 3.1: Matching rainfall from PERSIANN-CCS algorithm with the rainfall from PMW sensors using PMM 15

Figure 3.2: PERSIANN-CCS rainfall data adjustment using PMW and Probability Matching Method (PMM) 17

Figure 3.3: Frequency map for 5x5 degree subareas calculated based on the climatology data. Sample counts map for 3 summer months and 3 winter months. Blue areas have low number of samples. 19

Figure 3.4: CDF comparison between PERSIANN-CCS & PMW based on climatology data. Solid blue line represents CDF for PMW data and dashed red line represents CDF for PERSIANN-CCS data. 21

Figure 3.5: CDF comparison between PERSIANN-CCS & PMW and MA-PERSIANN-CCS during validation year 2012. Blue solid line represents CDF for PMW data, red dashed line represents CDF for PERSIANN-CCS data, and green solid line represents CDF for microwave-adjusted PERSIANN-CCS 23

Figure 3.6: Number of available concurrent data samples for each pixel for winter 2012 (December, January, February) and Summer 2012 (June, July, August) 26

Figure 3.7: Average monthly rainfall calculated based on the concurrent samples from PMW and PERSIANN-CCS data during winter 2011 (mm/month-December, January, February). (a) PERSIANN-CCS (b MA-PERSIANN-CCS (c) PMW 27

Figure 3.8: Average monthly rainfall calculated based on the concurrent samples from PMW and PERSIANN-CCS data during summer 2011 (mm/month-June, July, August). (a) PERSIANN-CCS (b MA-PERSIANN-CCS (c) PMW..... 28

Figure 3.9: Average monthly rainfall calculated based on the concurrent samples from PMW and PERSIANN-CCS data during winter 2012 (mm/month-December, January, February). (a) PERSIANN-CCS (b) MA-PERSIANN-CCS (c) PMW 31

Figure 3.10: Average monthly rainfall calculated based on the concurrent samples from PMW and PERSIANN-CCS data during summer 2012 (mm/month-June, July, August). (a) PERSIANN-CCS (b) MA- PERSIANN-CCS (c) PMW 32

Figure 3.11: Monthly average rainfall calculated based on the concurrent samples during winter 2012. (a) PERSIANN-CCS (b) MA- PERSIANN-CCS (c) PMW (d) Radar Q-2 38

Figure 3.12: Monthly average rainfall calculated based on the concurrent samples in summer 2012. (a) PERSIANN-CCS (b MA-PERSIANN-CCS (c) PMW (d) Radar Q-2..... 39

Figure 4.1: IR image and segmentation based on temperature threshold maximum equal to 253 K and 300 K. (a) IR image for September 09, 2014, UTC 17:30 (b) Segmentation based on Tb=253K for September 09, 2014, UTC 17:30 (c) Segmentation based on Tb = 300 K for September 09, 2014, UTC 17:30 42

Figure 4.2: K-means clustering method procedure..... 44

Figure 4.3: The calibrated cloud-rainfall relationships (Tb-RR) from the PERSIANN-CCS algorithm using IR and radar data. Each curve referring to a cloud cluster in the K-means classification for maximum temperature threshold 253 K..... 45

Figure 4.4: The calibrated cloud-rainfall relationships (Tb-RR) from the PERSIANN-CCS algorithm using IR and radar data. Each curve referring to a cloud cluster in the K-means classification for maximum temperature threshold 300 K..... 45

Figure 4.5: Accumulated rainfall from training data (June-September 2014). (a) PERSIANN-CCS (Tb = 253K); (b) PERSIANN-CCS (Tb = 300K); (c) MRMS radar data. (mm)..... 47

Figure 4.6: monthly average rainfall during validation period (June-September 2015). (a) PERSIANN-CCS (Tb = 253 K); (b) PERSIANN-CCS (Tb = 300 K); (c) MRMS radar data. (mm)..... 48

Figure 5.1: Dispersion in the relation Tb (IR)-RR, including rain/no-rain and positive values for one cluster..... 49

Figure 5.2: Redistribution of rainfall and cloud top temperature from one cloud cluster (black); The red line separates the zero rates and the positive rates in the redistributed data; the actual rainfall associated with brightness temperatures are shown in grey..... 50

Figure 5.3: Probabilistic precipitation identification and estimation..... 53

Figure 5.4: Calculating POP values for maximum HSS during training period for identifying the POP threshold..... 55

Figure 5.5: Time series of binary analysis of probabilistic model and PERSIANN-CCS during training period (June-September 2014). Blue dashed line represent values for PERSIANN-CCS and red solid line represents values for probabilistic model..... 58

Figure 5.6: Time series of relative bias, correlation coefficient, and RMSE of Probabilistic model and PERSIANN-CCS during training period (June to September 2014). Blue dashed line represent values for PERSIANN-CCS and red solid line represents values for Probabilistic model..... 59

Figure 5.7: Difference between POD, FAR, and HSS values of probabilistic model and PERSIANN-CCS during training period (June-September 2014)..... 60

Figure 5.8: Difference between relative bias, correlation coefficient, and RMSE values of probabilistic model and PERSIANN-CCS during training period (June-September 2014)..... 61

Figure 5.9: Total accumulated rainfall during training period (mm)..... 62

Figure 5.10: Time series of binary analysis of probabilistic model and PERSIANN-CCS model during validation period (July 2015). Blue dashed line represents values for PERSIANN-CCS and red solid line represents values for probabilistic model..... 65

Figure 5.11: Time series of binary analysis of probabilistic model and PERSIANN-CCS during training period (August 2015). Blue dashed line represents values for PERSIANN-CCS and red solid line represents values for probabilistic model..... 66

Figure 5.12: Time series of binary analysis of probabilistic model and PERSIANN-CCS model during validation period (September 2015). Blue dashed line represents values for PERSIANN-CCS and red solid line represents values for probabilistic model. 67

Figure 5.13: Time series of relative bias, correlation coefficient, and RMSE of Probabilistic model and PERSIANN-CCS during validation period (July 2015). Blue dashed line represents values for PERSIANN-CCS and red solid line represents values for Probabilistic model. 68

Figure 5.14: Time series of relative bias, correlation coefficient, and RMSE of Probabilistic model and PERSIANN-CCS during validation period (August 2015). Blue dashed line represents values for PERSIANN-CCS and red solid line represents values for Probabilistic model. 69

Figure 5.15: Time series of relative bias, correlation coefficient, and RMSE of Probabilistic model and PERSIANN-CCS during validation period (September 2015). Blue dashed line

represents values for PERSIANN-CCS and red solid line represents values for Probabilistic model..... 70

Figure 5.16: Difference between POD, FAR, and HSS values of probabilistic model and PERSIANN-CCS during validation period (July 2015). 71

Figure 5.17: Difference between POD, FAR, and HSS values of probabilistic model and PERSIANN-CCS during validation period (August 2015). 72

Figure 5.18: Difference between POD, FAR, and HSS values of probabilistic model and PERSIANN-CCS during validation period (September 2015)..... 73

Figure 5.19: Difference between relative bias, correlation coefficient, and RMSE values of probabilistic model and PERSIANN-CCS during validation period (July 2015). 74

Figure 5.20: Difference between relative bias, correlation coefficient, and RMSE values of probabilistic model and PERSIANN-CCS during validation period (August 2015). 75

Figure 5.21: Difference between relative bias, correlation coefficient, and RMSE values of probabilistic model and PERSIANN-CCS during validation period (September 2015)..... 76

Figure 5.22: Total accumulated rainfall during validation period (July 2015-mm) 77

Figure 5.23: Total accumulated rainfall during validation period (August 2015-mm) 78

Figure 5.24: Total accumulated rainfall during validation period (September 2015-mm)..... 79

Figure 5.25: CDF comparison between PERSIANN-CCS & PMW and probabilistic model during validation year 2015. Blue solid line represents CDF for PMW data, red dashed line represents CDF for PERSIANN-CCS data, and green solid line represents CDF for probabilistic model..... 81

LIST OF TABLES

Table 3-1: Statistical parameters used for global validation based on average monthly rainfall. Parameters are calculated for 8 zones comparing with PMW data (December, January, and February 2011) 33

Table 3-2: Statistical parameters used for global validation based average monthly rainfall. Parameters are calculated for 8 zones comparing with PMW data (June, July, and August 2011) 33

Table 3-3: Statistical parameters used for global validation based on average monthly rainfall. Parameters are calculated for 8 zones comparing with PMW data (December, January, and February 2012 34

Table 3-4: Statistical parameters used for global validation based on average monthly rainfall. Parameters are calculated for 8 zones comparing with PMW (June, July, and August 2012) 35

Table 3-5: Comparison between PERSIANN-CCS and MA-PERSIANN-CCS with PMW satellite data over CONUS during winter and summer 2012. Parameters are calculated based on average monthly rainfall. 40

Table 3-6: Comparison between PERSIANN-CCS, MA-PERSIANN-CCS, and PMW with Q2 ground based radar over CONUS during winter and summer 2012. Parameters are calculated based on average monthly rainfall. 40

Table 4-1: Statistical evaluation during training period (June to September 2014). The numbers are average of relative bias, correlation coefficient and RMSE over the entire training data 47

Table 4-2: Statistical evaluation during validation period (June to September 2015). The numbers are average of relative bias, correlation coefficient and RMSE over the entire validation data 48

Table 5-1: Statistical evaluation of probabilistic model and PERSIANN-CCS using MRMS data during training period. Numbers are calculated based on averaging over entire month. 57

Table 5-2: Statistical evaluation of probabilistic model and PERSIANN-CCS using MRMS data during validation period. Numbers are calculated based on averaging over entire month..... 80

ACKNOWLEDGMENSTS

I started graduate school at UCI since 2011. Now that I am finalizing my work and research from past five years I can recall the love and support of many people that I could not go through this journey without them.

I am so thankful to my wonderful advisor Professor Kuolin Hsu for his motivation and patience in mentoring me during my PhD. He is one of the most generous persons in sharing his knowledge and I learned a lot from him during our collaboration.

I cannot thank enough from my co-advisor Professor Soroosh Sorooshian for his advice, motivation, and inspiration throughout the years. His dedication and love to what his doing as well as winning various awards because of his excellent works and worldwide contributions in science were the greatest inspiration for me to set the quality as a priority for my works.

I would like to appreciate my committee member Professor Alex Ihler for his valuable advice on my research.

I appreciate Dan Brathwaite the CHRS IT specialist for providing me various types of data that I used for my research. His effort in preparing suitable data for our research is very valuable.

Being a part of CHRS family was one of the greatest experiences I had in my life. I have many great friends at CHRS that I would like to thank each of them and wish them success in their life.

My deepest gratitude goes to my wonderful parents Soraya and Alireza. I cannot thank enough of them because of their unconditional love and support. They always did their best to let me achieve my goals and ambitions. My younger brother Erfan who always cheered me up with his great sense of humor when I was feeling alone and stressed in the US because of being far from my lovely family. He is the most precious gift from god to me and I am so thankful to him. I am

so thankful to my husband Ali who motivated and supported me through this path and was with me in all ups and downs and I am so grateful to have him in my life.

The primary financial support for this study was provided by NASA Precipitation Measurement Mission (grant NNX10AK07G), NASA MEaSURES (grant NNX13AM12G), and NASA MIRO (grant NNX15AQ09A).

CURRCULUM VITAE

Negar Karbalaee

Email: nkarbala@uci.edu

Education

University of California, Irvine (UCI) 2012-2017

PhD in Hydrology and Water Resources Management (Civil Engineering Department)

Advisors: Professors Kuolin Hsu and Soroosh Sorooshian

University of California Irvine (UCI) 2011-2013

Master of Science in Civil and Environmental Engineering

Amirkabir University of Technology (Tehran Polytechnic) 2005-2010

Bachelor of Science in Civil and Environmental Engineering

Professional Appointments

Graduate Research Assistant 2012-2017

Center for Hydrometeorology and Remote Sensing (CHRS)

University of California, Irvine

Teacher Assistant 2012-2017

University of California, Irvine

Visitor Researcher (Internship) 2016

NSSL - NOAA (National Severe Storms Laboratory –National Oceanic Atmospheric Administration)

National Weather Center (NWC) – Norman, OK

Research Experience

Center for Hydrometeorology and Remote Sensing (CHRS) 2012-2017

University of California Irvine

- Precipitation estimation from remotely sensed data for hydrological applications.
- Developing a data mining framework to incorporate remote sensing and other sources of rainfall data to improve rainfall estimation.
- Application of statistical analysis and machine learning techniques in precipitation estimation based on the collected information from radar and satellites.
- Validating precipitation estimation models using ground truth information to improve algorithms performance.
- Evaluation of Salton Sea surface water change in the past three decades using HEC-RAS, HEC- HMS and ArcGIS and investigate the potential impacts of climate change.

- Trend analysis of climatological indices to study the climate change impacts on extreme events such as severe storms.
- Hydrologic modeling of the Leaf River basin located in Mississippi to perform rainfall-runoff analysis and flood forecasting using different rainfall products.

NOAA-NSSL (National Oceanic and Atmospheric Administration - National Severe Storm Laboratory) 2016

National Weather Center

- Using multi-radar multi-sensor data to validate the satellite precipitation data.
- Developing probabilistic based precipitation algorithm using infrared data from satellites.
- Preparing a proposal on probabilistic rainfall estimation from GOES-R for water resources management and hydrological modeling purposes (Accepted and funded by NOAA staring July 2017).

Teaching Experience

University of California, Irvine 2012-2016

- Hydrology (Fall 2016)
- Fluid Mechanics (Fall 2012)
- Statics (Winter 2013, 2014, and 2016 and Fall 2014)
- Structural Analysis (Winter 2012)

Responsibilities: Holding discussion sessions, lecture classes, and office hours, solving problems, designing homework, exams, preparing required materials.

Publications and Poster Presentations

- **Karbalaee, N.**, K. Hsu, S. Sorooshian, and D. Braithwaite (2017), Bias adjustment of infrared-based rainfall estimation using passive microwave satellite rainfall data, J. Geophys. Res. Atmos., 122,

- **Karbalaee N.**, Sellars S., Tao Y., Ihler A., Gao X., Hsu K., Sorooshain S., "Machine Learning in Hydrometeorology", *Forth workshop on understanding climate change from data. Boulder, Colorado,2014*.

- **Karbalaee N.**, Hsu K., Sorooshian S.", Improving PERSIANN-CCS rainfall data using PMW rainfall data", *American Geophysical Union, 15-19 December, 2014. San Francisco, 2014*.

Oral Conference Presentations

- **Karbalaee N.**, Hsu K., Sorooshian S., "Improving warm rain estimation in the PERSIANN-CCS satellite-based rainfall algorithm", *American Geophysical Union, 14-18 December, 2015. San Francisco, 2015*.

- **Karbalaee N.**, Hsu K., Sorooshian S., Kirstetter P., Hong Y., "Improving PERSIANN-CCS rain estimation using probabilistic approach and multi-sensor information", *American Geophysical Union, 12-16 December, 2016. San Francisco, 2016*.

- **Karbalaee N.**, Hsu K., Sorooshian S., "Improving PERSIANN-CCS rainfall estimates using Probability Matching Method (PMM*)", American Meteorological Society, January 2017. Seattle, Washington.*

Professional Services

- AGU Student Volunteer Service (2015)
- Journal reviewer
  IEEE Transaction on Geoscience and Remote Sensing
  Journal of Hydrometeorology

Honor/Awards

- National Weather Center (NWC) internship fellowship. (2016)
- Department of Civil and Environmental Engineering Travel Grant, University of California Irvine. Travel to American Geophysical Union (AGU) Conference, San Francisco. (2014,2015, and 2016)
- American Geophysical Union student conference registration award, San Francisco. (2015)
- Association of Graduate Students (AGS) Travel Grant, University of California Irvine. Travel to American Geophysical Union (AGU) Conference, San Francisco. (2014)

Software

- **Engineering Software:** ArcGIS, HEC-RAS (1D & 2D), HEC-HMS, ENVI, LSPC
- **Programming:** Matlab, Python, R, Excel Macro (VB)

Professional Membership

- American Geophysical Union (AGU)
- American Meteorological Society (AMS)
- American Society of Civil Engineers (ASCE)

ABSTRACT OF THE DISSERTATION

Improving warm rain detection and rainfall estimation

of a multiple satellite-based rainfall retrieval algorithm

By

Negar Karbalaee
Doctor of Philosophy in Civil Engineering
University of California, Irvine, 2017
Professor Kuolin Hsu, Chair

Precipitation as an essential component of the hydrologic cycle has a great importance to be measured accurately due to various applications such as hydrologic modeling, extreme weather analysis, and water resources management. Among different methods, meteorological satellites are one of the instruments that are widely used for precipitation estimation in fine spatial and temporal resolution. Precipitation Estimation from Remotely Sensed Imagery using Artificial Neural Network Cloud Classification System (PERSIANN-CCS) uses infrared (IR) data from Geostationary Earth Orbit (GEO) satellites to retrieve precipitation based on relationship between clout top temperature (Tb) and rainfall rate (RR) using a neural network technique. The complexity of Tb-RR relationship for estimating precipitation causes uncertainty in PERSIANN-CCS rainfall product. This research is focused on improving PERSIANN-CCS rainfall retrieval using several approaches:

1) Bias adjustment of PERSIANN-CCS rainfall estimates using PMW satellite rainfall data: Using multi satellite data can enhance the quality of rainfall estimation considerably; in

this research we have combined the rainfall data from PERSIANN-CCS and PMW rainfall to enhance the bias of PERSIANN-CCS precipitation estimates. The results showed improvement of rainfall estimation during summer and winter time.

2) Increasing the rainfall detection by including warm clouds rainfall: PERSIANN-CCS currently cannot detect rainfall from clouds with temperature warmer than 253 K. This study explores the impacts of increasing the temperature threshold on precipitation estimation. The results show that increasing the threshold level can improve the PERSIANN-CCS rainfall detection.

3) Generating a probabilistic framework for precipitation retrieval: The current version of PERSIANN-CCS retrieves precipitation based on the exponential function fitted to Tb-RR. The major assumption behind this relationship is that the heavier rainfalls are associated with colder clouds which cause underestimation of warmer clouds and overestimation of colder clouds rainfall. The probabilistic approach uses the corresponding sample relationship between cloud temperature and rainfall rate. The model is evaluated during a full summer season which showed improvement in both detection and estimation of rainfall in compare with the current PERSIANN-CCS algorithm.

CHAPTER 1: Introduction to current global precipitation estimation products

1.1. Importance of precipitation estimation

Wherever there is life there is a need for water. Many parts of the world are facing water problems caused by human and environmental impacts on water resources. Increasing population rates are leading to an increased demand for water. Also, water pollution is a huge problem that is caused by different factors such as a weak sewage system and discharging industrial waste that contaminate water resources. Climate change and variability have altered weather patterns around the world, resulting in changing precipitation patterns. Some regions such as California in the past few years received less precipitation as compared with the normal precipitation amounts, resulting in severe drought and water shortages. Analysis of precipitation observation also show that the frequency and amount of heavy precipitation has increased over most of the United States in the past several decades, resulting in increased flooding and human loss and infrastructure damages. Recognizing the importance of precipitation as a key component of hydrologic cycle and most important input for modeling of the rainfall-runoff process, this dissertation focuses on improving the precipitation estimation from satellites. Knowing the intensity and location of the precipitation can improve the water resources planning and management.

1.2. Precipitation estimation/measurement

There are a number of approaches and instruments in use for measuring precipitation. Rain gauges are one of the primary methods in measuring the rain depth. Even though, rain gauges provide a direct measurement of precipitation reaching the ground, they contain uncertainties due to different factors (Yilmaz et al. 2005). The sparse distribution of rain gauges does not provide a

good representation for areal precipitation especially over mountainous regions since most gauges are located in lower elevation sites. Other factors such as wind, evaporation, and recording errors can increase the bias of gauge measurement (Groisman and Legates 1994, Legates and DeLiberty 1993).

Ground-based radars are available and provide in high spatial and temporal resolution for different hydrological applications. Radars measure precipitation from the backscattered signal from raindrops. Radar rainfall estimates are not available over mountains due to beam blockage and typically have biases due to the uncertainty between rainfall and radar reflectivity (Morin et al. 2003 and Kirstetter et al. 2010). In addition, radar data are not available over oceans.

The limitations of rain gauges and ground-based radars pave the way for a new generation of remote sensing instruments for global rainfall measurement with high spatial and temporal accuracy. The history of measuring precipitation from meteorological satellites goes back to 1960 (Barret 2001). Meteorological satellites carry different sensors that can measure different climatological parameters from the space. Global Precipitation Measurement (GPM) mission is one major improvement in measuring rainfall and snowfall from space to improve our understanding from Earth's water and energy cycle (Hou et al. 2014).

Although many efforts have been made to improve the satellite data quality so far, the mission of enhancing the quality of precipitation from satellites will be continuing.

1.3. Satellite-based precipitation estimation

Geostationary Earth Orbit (GEO) and Low Earth Orbit (LEO) are the two main types of satellite instruments being used for hydrometeorological applications. GEO satellites is an earth orbiting satellite, placed into an orbit with approximately 36,000 km altitude that rotates in the same direction that earth rotates to maintain a fixed position relative to a point on the ground. GEO

satellites carry sensors with spectral wavelengths from visible (VIS) to longwave infrared (IR) which are relevant to albedo and cloud top temperature with spatial resolution of 1-4 km and temporal resolution of 15-30 minutes covering up to 60° latitudes (Sorooshian et al. 2000).

LEO satellites are located between 300 and 800 km with less sampling frequency compare to GEO satellites. Although having lower spatial and temporal sampling frequency and resolution, Passive Microwave (PMW) sensors from LEO satellites can measure thermal emissions which are attenuated by raindrops (Marzano et al. 2004; Tapiador 2008; Behrangi et al. 2009).

In the past several decades GEO and LEO satellites have been frequently used for precipitation analysis (Ebert and Manton 1998). Griffith et al. (1978) used IR and VIS data to estimate convective rainfall using radar data as a reference. Wu et al. (1985) implemented a pattern recognition technique to develop a model for identifying rainfall based on textural and radiance features from IR and VIS images. Scofield (1987) estimated the half hourly rainfall from convective clouds based on information such as cloud top temperature, overshooting tops, and saturated environment from GOES satellites. Adler and Negri (1988) used IR data to estimate both tropical convective and stratiform precipitation. The minimum of brightness temperature was assigned to the convective center and rainfall estimation was completed based on mode temperature of thunderstorm anvils. Arkin and Xie (1994) used the relationship between fractional coverage and rainfall to estimate precipitation over Japanese Island. Vicente at al. (1998) uses GOES-IR data to compute real time precipitation using power law regression that has many applications for flash fool forecasting, numerical modeling, and operational hydrology. Weng et al. (1994) showed the precipitation estimation analysis over land and ocean by using Special Sensor Microwave Imager (SSM/I). Ferraro and Marks (1995) used radar data over the

United States, Japan, and the United Kingdom to retrieve precipitation based on both linear and non-linear relationship using SSM/I.

1.4. Research Motivation

Precipitation estimation with appropriate time and space resolution is crucial for different hydrometeorological applications such as flash flood warning, climate extreme analysis, and weather forecasting. Sparse gauge distribution over land and radar beam blockage due to mountain regions as well as unavailability of rain gauges and radars over ocean makes precipitation data inaccessible over different regions especially over water bodies. Some events such as hurricanes form over the large water bodies and bring heavy rain to land depending on location and strength which can be trackable with remote sensing technology. Therefore, satellite remotely sensed data provide feasibility of measuring different climatological parameters such as precipitation in high spatial and temporal resolution over land and oceans. GEO based satellite precipitation estimation based on VIS/IR imageries contains uncertainty due to the complication of retrieving rainfall from cloud top temperature and albedo. PMW sensors from polar orbiting satellites can detect hydrometeor information in clouds and collect information about water and ice content of the clouds formations based on emissions of cloud water which is more correlated with rainfall estimation. While PMW rainfall algorithms are more reliable measurements, the high spatial and temporal resolution of GEO based algorithms (15 min/4 km) is highly preferred compared to PMW resolution (3-6 h/15 km) for instantaneous and near real time precipitation analysis. Building upon the success of TRMM which operated from November 1997 to April of 2015, the Global Precipitation Measurement (GPM) satellite was launched in February 2014. Integrated Multi-SatellitE Retrievals for GPM (IMERG) is a concept based on using the

available LEO-PMW satellite data and adding GEO-IR based precipitation data to provide global precipitation estimation at near real time (Huffman et al 2015a, Liu 2015).

The research motivation leading to this dissertation is improving precipitation estimation using PERSIANN-CCS algorithm that can be beneficial for the IMERG algorithm and other operational use. In specific, Integrated Multi-satellitE Retrievals for GPM (IMERG) uses all LEO-satellites rainfall samples and then filling the gaps using GEO-based precipitation algorithm called Precipitation Estimation using Remotely Sensed Imagery Artificial Neural Network Cloud Classification System (PERSIANN-CCS) to generate a precipitation map at fine spatial and temporal resolution. While the high spatial and temporal resolution of CCS that makes it a very popular algorithm for operational use, it has some limitations. PERSIANN-CCS computes precipitation based on IR information from GOES satellites based on the relationship between cloud top temperature and rainfall rate relationship assuming that the rainfall intensities are greater for deeper cloud formations. This assumption is associated with convective clouds but cannot be applied for other types of clouds such cirrus, stratiform, and orographic. The current version of CCS also assumes a threshold level for cloud brightness temperature that assumes zero rainfall for clouds with warmer than certain threshold resulting in missing rainfall from warm clouds. In addition, in order to link a cloud top brightness temperature (Tb) and Rainfall Rate (RR), the PERSIANN-CCS algorithm uses an exponential function to fit a Tb-RR data based on the training data. A key assumption for using this exponential function is that the heavier rainfall should be associated with colder clouds that cause losing the actual relationship between IR and RR data. The results reported in this dissertation have been obtained from the research focused on overcoming some of the above listed limitations of the PERSIANN-CCS algorithm that can be beneficial for IMERG and other operational use.

1.5. Objectives of this dissertation

Global Precipitation Mission (GPM) is a joint collaboration between NASA and JAXA to enhance scientific understandings of the Earth's water and energy cycle as well as producing near real-time data for a variety of meteorological applications such as climate prediction, extreme weather analysis, soil moisture, and flood forecasting.

One key dataset from the NASA GPM program is the Integrated Multi-satellitE Retrievals for GPM (IMERG). IMERG integrates both low earth orbital (LEO) and geostationary earth orbital (GEO) satellite information and surface precipitation gauge analysis to provide global precipitation (Huffman et al., 2015a,b). The satellite-based estimation of IMERG is a merged retrieval using algorithms developed from three groups, including (1) TRMM Multi-satellite Precipitation Analysis (TMPA) from NASA Godard Flight Center, (2) Climate Prediction Center Morphing with Kalman Filter (CMORPH-KF) from NOAA CPC, and (3) microwave-calibrated PERSIANN-CCS from the University of California, Irvine, Center for Hydrometeorology and Remote Sensing (CHRS) (Huffman et al., 2007; Joyce and Xie 2011; Hong et al. 2004). IMERG has been developed to provide better quality and shorter time latency for global precipitation monitoring at 0.1^{o}x0.1^{o} and half-hourly samples. The time latency of the data products is around 4 hours and 12 hours from observation time for the "Early" and "Late" near real-time multi-satellite products. The "final" satellite-gauge merged product is available around 2-months after observation time (Huffman et al., 2015a,b). The aim of this research is to accomplish some improvements in PERSIANN-CCS rainfall algorithm component of IMERG, focusing on the following aspects:

- *Using multi- satellite data and combining GEO-based and LEO-based satellite estimate to improve precipitation derived from PERSIANN-CCS algorithm over the globe. How much improvement is achieved by combining PERSIANN-CCS and PMW satellite data over different regions and seasons knowing that PMW sensors retrieve precipitation more accurately?*

- *PERSIANN-CCS algorithm estimates precipitation up to certain brightness temperature threshold. This causes that PERSIANN-CCS to miss rainfall from clouds with temperature higher than 253 K. This issue is investigated to see how much precipitation can PERSIANN-CCS capture if we raise the temperature threshold to higher levels.*

- *PERSIANN-CCS algorithm retrieves precipitation based on the exponential relationship between Tb-RR (cloud top temperature and rainfall intensity). The rational for using the exponential relationship is that the heavier rainfalls rates are associated with colder clouds. Two potential issues arise: 1) falsely estimate precipitation from non-precipitating cold clouds (cirrus clouds and) 2) missing rainfall from warm clouds (Behrangi et al. 2010a). In this research the degree of improvement which can be achieved for implementing a probabilistic rainfall estimation model based on the actual Tb-RR relationship is explored.*

CHAPTER 2: PERSIANN-CCS precipitation algorithm

2.1.Algorithm framework

Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN) was developed employing the Artificial Neural Network (ANN) machine learning modeling framework and using IR images to fit the pixel brightness temperature and its neighbor temperature textures, in terms of means and standard deviations, to calculate pixel rain rates at 0.25x0.25 degree spatial resolution (Hsu et al. 1997). PERSIANN rainfall estimate is continuously adjusted using PMW rainfall data (Sorooshian et al. 2000). PERSIANN Cloud Classification System (PERSIANN-CCS) is a transition for pixel based rainfall estimates of PERSIANN to patch based estimation for improving the precipitation retrievals as well as improving the spatial resolution to 0.04x0.04 degree. PERSIANN-CCS is an object-based algorithm that uses cloud coverage areas under defined IR temperature thresholds to estimate rainfall. The algorithm uses the longwave IR channel (10.7 μm) to classify cloud images based on the IR brightness temperature segmentation and then extract features for the cloud patches at three levels (220 K, 235 K, and 253 K). The rain estimation from PERSIANN-CCS is based on calibrating the cloud top temperature and rainfall (Tb-R) relationships for the classified cloud groups, initially using gauge corrected radar hourly rainfall data from the Next-Generation Weather Radar (NEXRAD) network when the algorithm was first developed over the CONUS (Hong et al. 2004; Hsu et al. 2007). The current operational PERSIANN-CCS algorithm was trained using one full year of (MWCOMB) rainfall data, obtained from the NOAA Climate Prediction Center (CPC) for a near global 60N to 60S and 180W to 180E coverage, with separately trained 24 overlapping subareas to account for latitude and terrain variability over the globe. Figure 2.1 demonstrates the PERSIANN-CCS algorithm structure.

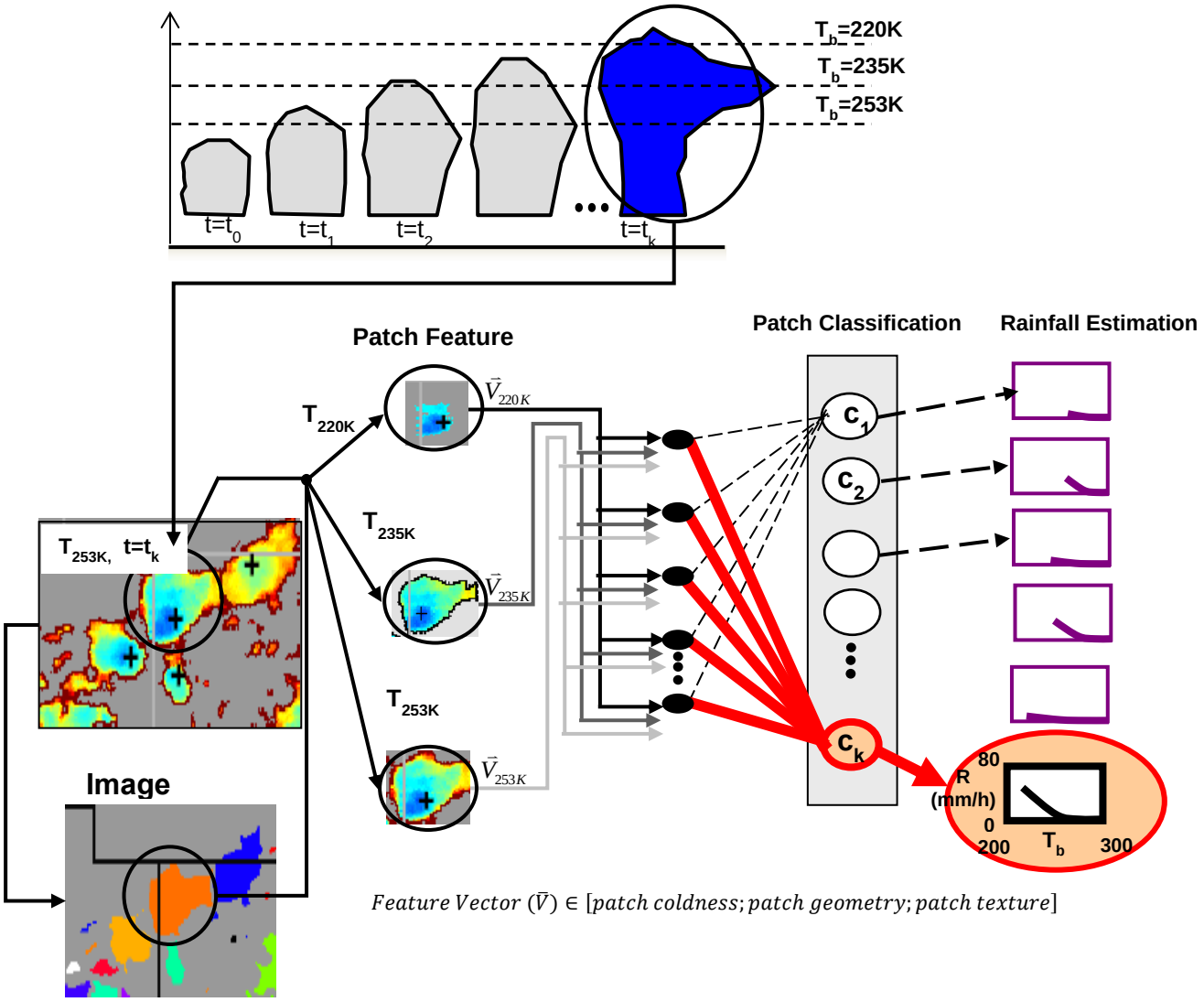


Figure 2.1: PERSIANN-CCS algorithm framework (Hong et al. 2004)

2.2.Algorithm strengths and potentials for improvement

PERSIANN-CCS retrieves precipitation from IR data at spatial resolution 0.04x0.04 degree and temporal resolution 30 min. Although IR sensors on board of the GEO satellites provide high spatial and temporal observation, the cloud top brightness temperature is not always related to microphysical properties of clouds and cloud morphology to estimate rainfall intensity.

PERSIANN-CCS algorithm estimates precipitation based on the exponential relationship between Tb-RR assuming that the heavier rainfalls are associated with colder clouds at higher altitudes. This assumption is more valid for convective clouds and less valid for other types of clouds such as cirrus, stratiform, and orographic clouds causing underestimation and overestimation depending on region and season. Improving PERSIANN-CCS can be considered from different prospective:

1) Enhancing the rainfall estimation from PERSIANN-CCS by combining the algorithm with other satellites data and benefit from multi-satellite product instead of using a single channel.
2) PERSIANN-CCS captures rainfall from the clouds with the brightness temperature up to 253-K. Increasing the temperature threshold can enhance the ability of PERSIANN-CCS to capture the rainfall from warm clouds.
3) The pixel based rainfall estimation for PERSIANN-CCS is based on fitting a nonlinear exponential function to the redistributed pixels for each cloud groups. The improvement in rainfall retrieval from PERSIANN-CCS can be achieved by fitting a function on the actual Tb-RR distribution.

CHAPTER 3: Using PMW precipitation data to reduce the bias of PERSIANN-CCS estimation

3.1. The role of multi-satellite data in precipitation algorithms

As mentioned in the previous chapters, GEO satellites equipped with IR and VIS channels and LEO satellites with PMW sensors are two major satellites used for precipitation estimation. Despite the high spatial and temporal resolution of IR and VIS sensors, they measure cloud top brightness temperature and cloud albedo which are less physically related to rain rate. PMW sensors has less frequent sampling which is not sufficient to capture precipitation variability, but offer more accurate rainfall estimation in compare with GEO satellites. To take advantage of the ability of both types of satellites in measuring precipitation, different approaches have been used in order to combine rainfall data from different sensors to retrieve precipitation more accurately at a fine spatial and temporal resolution (Levizzani et al. 1996; Turk et al. 2003). Effectively integrating estimations from LEO-PMW and GEO-VIS/IR information provides an option for improving precipitation at high spatial and temporal resolution. Using multi-spectral precipitation analysis was the research topic in the past two decades and a lot of improvements have made through this concept. Adler et al. (1993) used IR data from geosynchronous satellite combined with microwave sensors on board the Defense Meteorological Satellite Program (DMSP) to calculate the mean monthly rainfall. Another technique used to overcome the PMW poor sampling was using coincident microwave and infrared observation. For model calibration, Kummerow et al. (1995) used infrared temperature threshold that corresponds to an area equal to the raining area in a coincident observed microwave image. The mean conditional rainfall rate was assigned to pixels colder than the certain threshold in infrared images. This approach was compared with fixed threshold equal to 235-K for monthly accumulated rainfall.

More recently, Huffman et al. (2001) used Threshold-Matched Precipitation Index (TMPI) to estimate precipitation from 3-hourly IR data histograms from GEO satellites to fill the gaps for LEO satellites data for estimating global daily precipitation. Also, Probability Matching Method (PMM) was used to estimate the rainfall based on the distribution of rainfall from microwave sensors and IR brightness temperature (Turk et al 2000).Microwave Infrared Rainfall Algorithm (MIRA) (Todd et al. 2001) uses PMW and IR data for rainfall estimation assuming that PMW sensors have more accurate estimates of instantaneous rainfall. The Self-Calibrating Multivariate Precipitation Retrieval (SCaMPR) algorithm estimates rainfall at a fine temporal resolution using PMW and GEO satellites. SCaMPR uses Special Microwave Sensor Imager (SSM/I) data to distinguish between rain/no-rain pixels, and then uses Geostationary Operational Environmental Satellites (GOES) data to calibrate the relationship between Tb-RR via linear regression for the precipitating pixels (Kuligowski 2002 & 2009). Kidd (2003) used the histogram matching technique between PMW rainfall data and IR cloud-top temperature to estimate rainfall over Africa during a four month calibration period. Hong et al. (2005) used Tropical Rainfall Measurement Mission Microwave Imager (TMI) for bias adjustment of PERSIANN algorithm using coincident rainfall data derived from these two dataset. In addition to using IR and PMW rainfall data combined to enhance the quality of precipitation, some researches have done to take advantage of high spatial and temporal resolution of IR data and track the cloud motions and evolution to estimate the rainfall based on the available PMW rainfall data for one time step. The Climate Prediction Center morphing method (CMORPH) uses motion advection vectors from dynamic GEO-IR images to fill the temporal gaps between two available PMW rainfall estimates (Joyce et al. 2004). Bellerby et al. (2009) used single-band IR data on board of GOES satellites to compute the cloud motions and growth using Lagrangian Model (LMODEL) and estimated

the precipitation from PMW data using a Kalman filter method. Behrangi et al. (2010b) developed Rain Estimation using Forward-Adjusted advection of Microwave Estimates (REFAME) to use GEO-IR data for generating cloud motion vectors and estimate precipitation based on the previous PMW overpass for the time PMW rainfall estimates are not available.

Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA) combined precipitation estimates from multiple satellites, as well as gauges where feasible to generate rainfall data with a 0.25x0.25 degree resolution every 3 hours. In continue of improving the rainfall accuracy at fine spatial and temporal resolution using multi-satellite data, GPM is the latest improvement toward this approach. As discussed earlier in more detail, IMERG uses PERSIANN-CCS to fill the gaps for global rainfall coverage when PMW observations are not available. In this chapter we explain the methodology and results of bias adjusting PERSIANN-CCS using PMW data which leads to IMERG data quality enhancement.

3.2. Methodology

This study explores the use of LEO satellite microwave precipitation data for bias adjustment of GEO-based satellite estimates from the PERSIANN-CCS algorithm. The PMW precipitation dataset (MWCOMB), obtained from the NOAA Climate Prediction Center (CPC), was used in this study. MWCOMB is a blended precipitation data from multiple sensors and orbits, such as DMSP SSM/I, NOAA AMSU-B, and TRMM TMI (Ferraro et al. 1994, Kummerow et al. 2001, and Weng et al. 2003; Huffman et al., 2007). The Probability Matching Method (PMM) is used to match PERSIANN-CCS precipitation estimation toward MWCOMB data from LEO satellites. PMM looks at the relationship between two variables based on their marginal distribution without considering their joint distribution (Ciach et al. 1997). PMM was developed in 1987

(Calherios and Zawadzki; 1986) and has since been extended to meteorological applications Piman et al. (2007). Krajewski et al. (1991) and Rosenfeld et al. (1994) used PMM to improve the radar rainfall estimation using gauge measurements by assuming that the radar observed reflectivity has the same probability of occurrence as the gauge measured rain intensity. In this study the rainfall distribution estimated from PERSIANN-CCS is processed to match LEO-PMW rainfall data for their concurrent samples. Figure 3.1 explains how the precipitation from PERSIANN-CCS can be adjusted with PMW rainfall estimates based on the Cumulative Distributed Function (CDF) computed from the Probability Distribution Function (PDF). Figure 3.1.a describes the PDFs for PMW (solid line) and PERSIANN-CCS (dashed line) calculated based on the concurrent samples of climatology data showing the probability of rainfall rate. Figure 3.1.b is the CDF function calculated based on the PDFs of PMW and PERSIANN-CCS data. This figure explains how the bias adjustment is done using the PMM method. As the arrow shows on the right figure the rainfall rate from PERSIANN-CCS will be replaced by the PMW estimate for the same CDF value. Equations 1, 2, and 3 explain how the PDF and CDF are calculated.

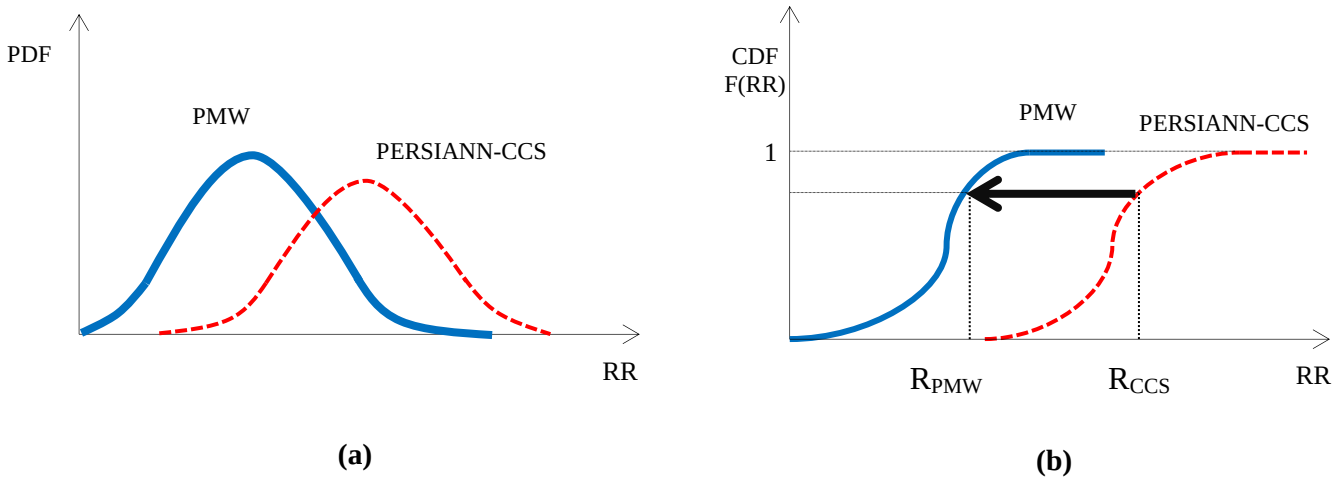


Figure 3.1: *Matching rainfall from PERSIANN-CCS algorithm with the rainfall from PMW sensors using PMM*

$$\int_0^{R_{CCS(i)}} p(R_{CCS})dR_{CCS} = \int_0^{R_{PMW(i)}} p(R_{PMW})\, dR_{PMW} \quad (1)$$

$$F(R_{CCS}) = \int_0^{R_{CCS(i)}} p(R_{CCS})\, dR_{CCS} \quad (2)$$

$$F(R_{PMW}) = \int_0^{R_{PMW(i)}} p(R_{PMW})\, dR_{PMW} \quad (3)$$

In equations (1), (2), and (3) R_{CCS} represents the precipitation from PERSIANN-CCS and R_{PMW} is the precipitation from LEO PMW estimates; $p(R_{CCS})$ and $p(R_{PMW})$ are the probability density functions of R_{CCS} and R_{PMW}. The PDFs of R_{CCS} and R_{PMW} are calculated from concurrent samples of historical data assuming that precipitation estimates R_{PMW} and R_{CCS} follow the same distribution each year.

3.3.Model Training

For model calibration we have used 4 years of climatology data for training (2008-2001).

PMM can also be estimated from PDFs calculated based on more recent data within a fixed time period window and dynamically adjusted over time when new data are available. Bias adjustment of PERSIANN-CCS can also be processed based on this approach of using both climatology and recent data adjustment. In this study, PMM based on climatology data only is presented. The concurrent data samples were collected each month separately for a four year period (2008-2011). Precipitation estimation from PERSIANN-CCS is at a 0.04^{O}x0.04º spatial resolution every 0.5 hour, while PMW precipitation from MWCOMB is at 0.25ºx0.25º. The Cumulative Distribution Functions (CDFs) of both datasets are calculated from the concurrent samples at the same spatial resolution. Therefore, PERSIANN-CCS is regridded to 0.25ºx0.25º spatial resolution. Meanwhile, to cover the regional and seasonal variability, the PDFs are calculated for each 5ºx5º subarea of the global coverage and each month separately to include a sufficient number of samples for model calibration. A lookup table is used for rainfall adjustment for each month and subarea. Figure 3.2 explains the PMM details discussed above briefly in one chart.

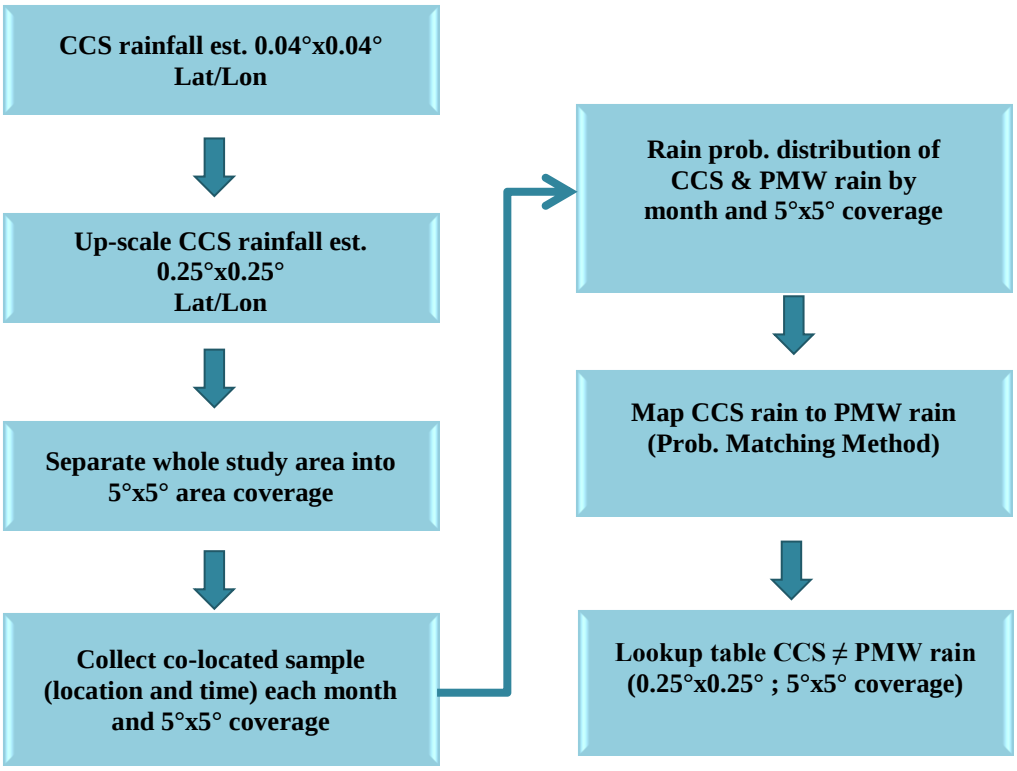


Figure 3.2: PERSIANN-CCS rainfall data adjustment using PMW and Probability Matching Method (PMM)

Figure 3.3 demonstrates the frequency map for each 5x5 degree subareas for 3 winter months and 3 summer months during the calibration period based on climatology data. During January, February, and December, which is the winter season over the northern hemisphere, regions located between 30N to 60N, do not have enough concurrent samples over land due to data flagging. PMW data exhibit high uncertainty when representing rain pixels over snow and ice surfaces, with flags for those pixels that are possible ice or cold ground during the cold seasons. Also, some areas in the Southern Hemisphere such as Australia have fewer concurrent samples compared with other regions in summer and winter due to some inconsistent GEO satellite coverage used to generate PERSIANN-CCS. For the bins with lower sample counts, the CDFs for PMW and CCS contain uncertainty, which means more data are required for CDF generation.

Figure 3.4 displays the empirical CDFs calculated based on climatology data during the calibration period. Subareas I to V are selected from different regions to compare the CDFs. Subareas I, II, and V are located in high latitude areas, showing that the rainfall distribution from PERSIANN-CCS and PMW is significantly different in these areas during winter and summer. Subareas III and IV are selected from tropical regions. The CDF comparison for subareas III and IV implies that the rainfall distribution for PMW and PERSIANN-CCS are more similar for tropical and near tropical regions; especially during January and February there is no significant difference between PERSIANN-CCS and PMW CDFs. Based on CDF comparison calculated from 4 years of climatology data we expect more adjustment for regions located in high latitude areas and less adjustment for regions located in tropical regions during summer and winter.

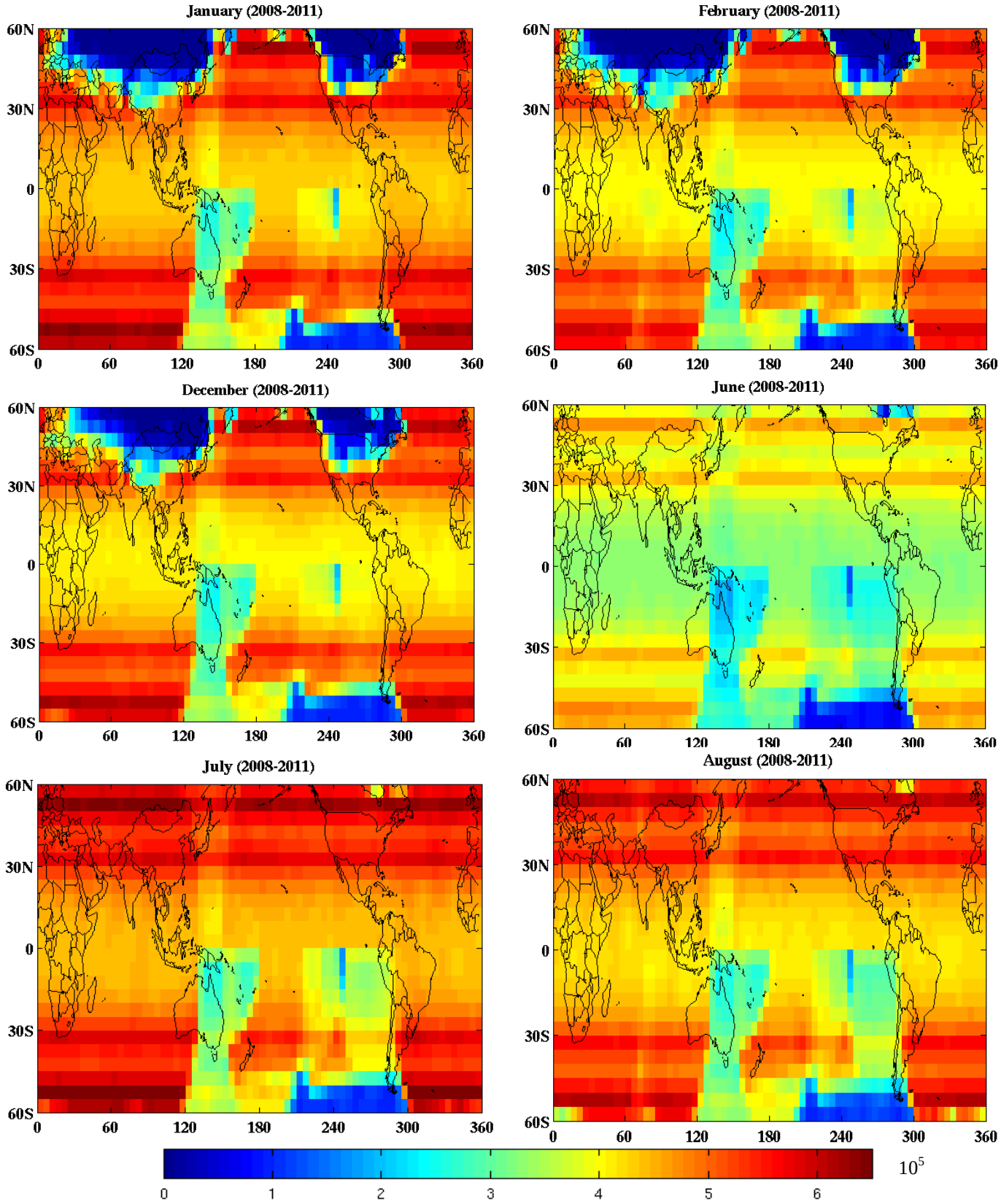
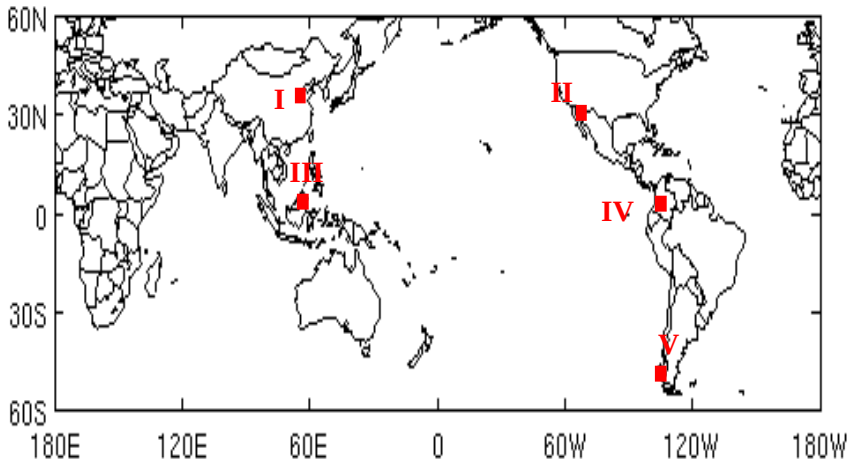
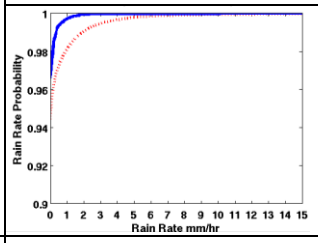
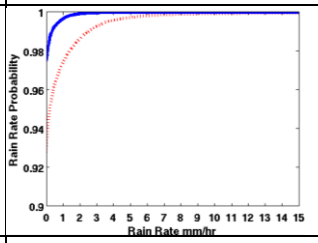
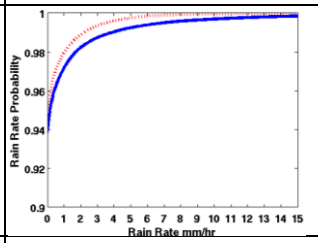
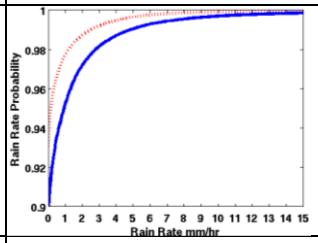
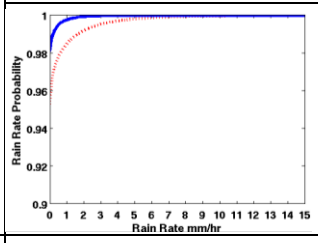
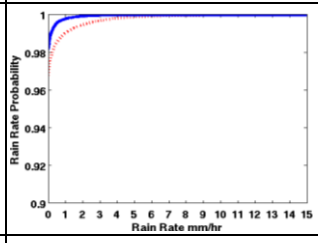
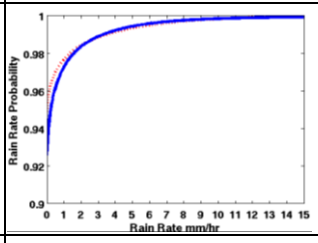
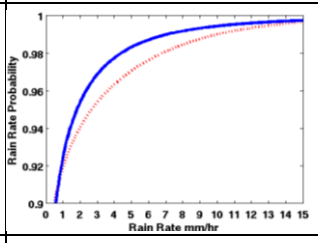
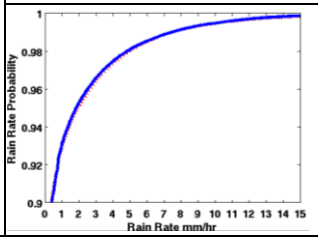
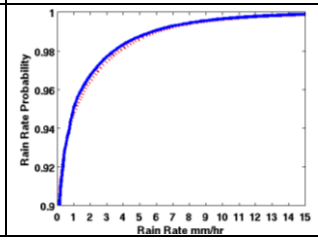
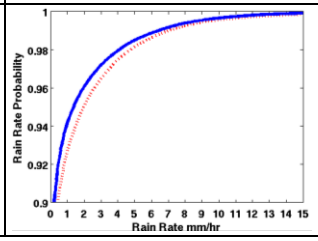
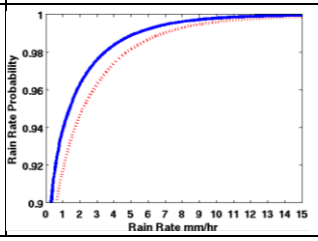


Figure 3.3: Frequency map for 5x5 degree subareas calculated based on the climatology data. Sample counts map for 3 summer months and 3 winter months. Blue areas have low number of samples.



	January	February	June	July
I				
II				
III				

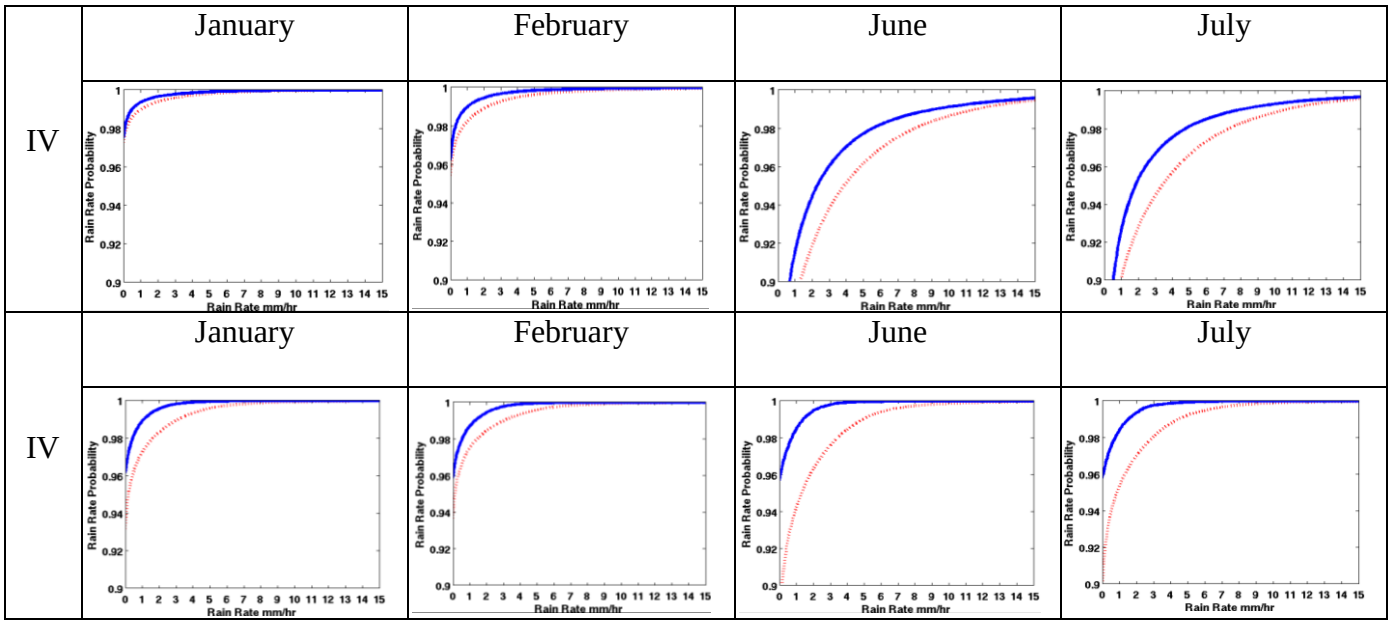


Figure 3.4: CDF comparison between PERSIANN-CCS & PMW based on climatology data. Solid blue line represents CDF for PMW data and dashed red line represents CDF for PERSIANN-CCS data.

3.4. Results and Discussion

After training the model based on the PERSIANN-CCS and PMW data from 2008 through 2011, the look up tables for each month contain PERSIANN-CCS and corresponding PMW rainfall values for bias adjustment that can be applied to PERSIANN-CCS data during any year for validation. In this section we discuss the validations that are done to evaluate the model performance. In section 3.4.1 we explain the CDF comparison between PERSIANN-CCS, PMW, and Microwave Adjusted PERSIANN-CCS (here after MA-PERSIANN-CCS) for validation year 2012 during winter and summer at the subareas I to V that was discussed in Figure 3.4. The global validation is performed using concurrent PMW, PERSIANN-CCS, MA-PERSIANN-CCS to explore how the bias adjustment improves the rainfall estimates from PERSIANN-CCS. The results for global validation during summer and winter 2012 are discussed in section 3.4.2. Also, the PMM model is evaluated over the CONUS using ground based radar in 2012. In section

3.4.3, we discuss the results regarding evaluation of PERSIANN-CCS and MA-PERSIANN-CCS against radar Q2 over the CONUS during summer and winter 2012. The lookup table for each subarea is calculated based on the concurrent samples of PMW and PERSIANN-CCS for the entire month. So, the global validation and CONUS validation is done over the monthly data to see the improvement in monthly scale.

3.4.1. Evaluating the cumulative distribution function during validation year

Figure 3.5 shows the empirical CDFs for subareas I to V during validation year 2012 for summer and winter. Subarea I is selected from high latitude areas in northern hemisphere. As we discussed earlier in Figure 3.4, PERSIANN-CCS overestimates rainfall during winter and underestimate rainfall during summer for bin I. In figure 3.5, Bin I also has the same pattern during summer and winter, while the CDF for MA-PERSIANN-CCS has shifted toward PMW CDF for summer and winter. Subarea II selected from mid-latitude in northern hemisphere. During January and February 2012, PERSIANN-CCS overestimates the rainfall for Bin II while MA-PERSIANN-CCS displays a similar rainfall distribution to PMW during these two months.

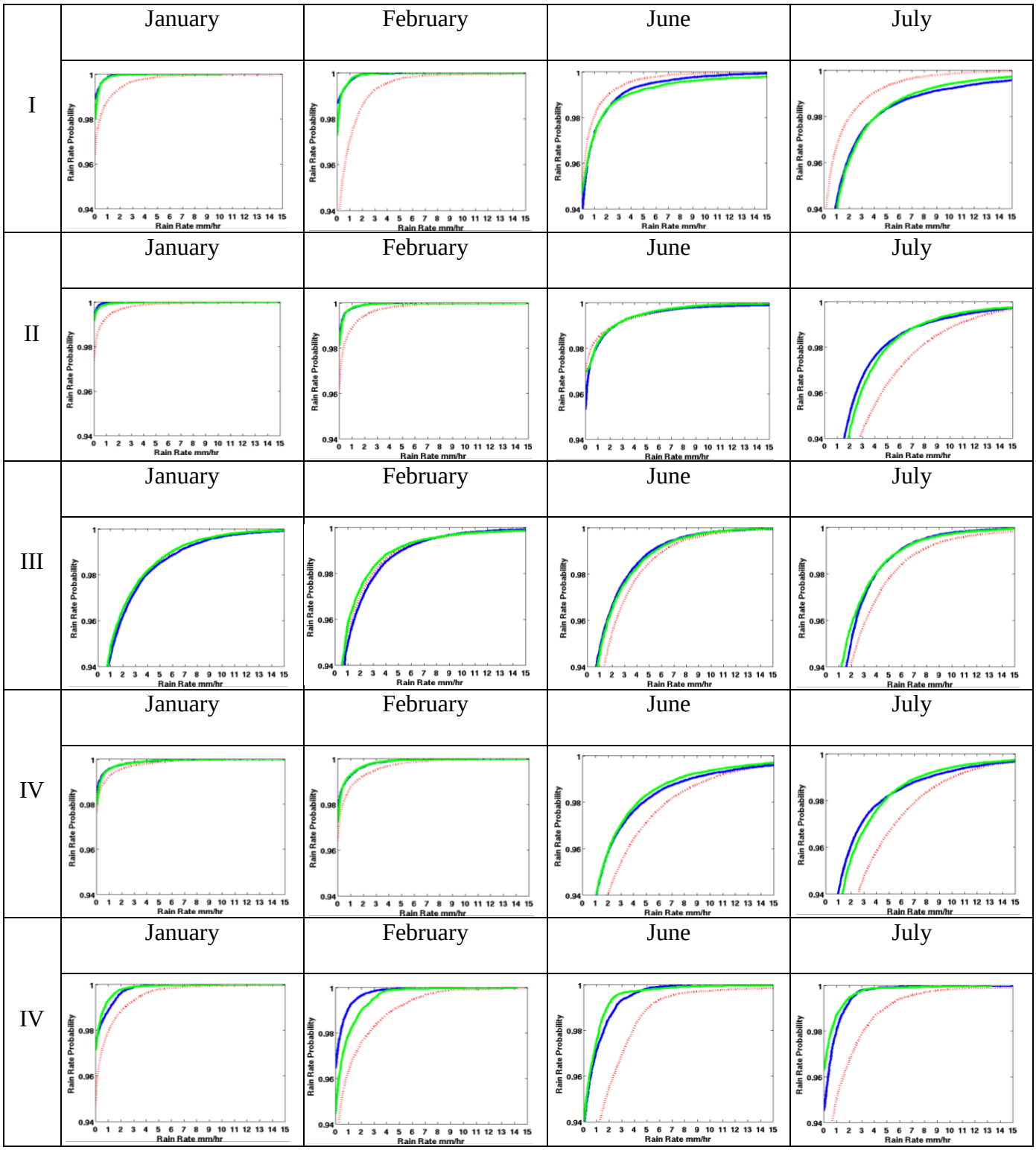


Figure 3.5: CDF comparison between PERSIANN-CCS & PMW and MA-PERSIANN-CCS during validation year 2012. Blue solid line represents CDF for PMW data, red dashed line represents CDF for PERSIANN-CCS data, and green solid line represents CDF for microwave-adjusted PERSIANN-CCS

Also, as was expected from CDF comparison during calibration years for June, PMW and PERSIANN-CCS CDFs are close to each other and no significant adjustment happened during validation year, so all three CDFs for Bin II are very close to each other during June 2012. During July 2012 PERSIANN-CCS overestimates rainfall for Bin II, while the CDF for MA-PERSIANN-CCS is well matched with the PMW CDF. For evaluation of the PMM model over a tropical region, subareas III and IV are chosen over tropical area. During January and February 2012, PERSIANN-CCS, MA-PERSIANN-CCS, and PMW CDFs are very close to each other and no significant adjustment is observed during winter season for tropical regions. Bins III and IV show an overestimation of rainfall from PERSIANN-CCS during June and July 2012, while the rainfall distributions are very similar to PMW after PMM application for these Bins. Subarea V is selected from high latitude regions in the southern hemisphere. As discussed earlier in Figure 3.4, PERSIANN-CCS overestimates rainfall for Bin V during summer and winter. The evaluation of CDFs for validation year 2012 shows improvement in the rainfall distribution of Bin V for both winter and summer. In summary, the PMM approach uses a rainfall distribution based on the concurrent samples of PMW and PERSIANN-CCS to match the rainfall based on rainfall probability from two different datasets. As was expected, the rainfall distribution of MA-PERSIANN-CCS is similar to PMW rainfall distribution, with more significant adjustment over high latitude regions. Regions located in tropical areas show more adjustment during summer time than winter time, since the PMW and PERSIANN-CCS CDFs are very similar during summer for this area.

3.4.2. Validation over the globe

The validation is done at global scale using PMW data to see how PMM can improve the PERSIANN-CCS rainfall estimate toward PMW data in different regions during warm and cold

seasons. Figure 3.6 displays the concurrent sample counts for each pixel during 3 winter months and 3 summer months during validation year 2012. During winter time (January, February and December) regions between 60N to 40N, 0 to 120E and 120W to 60W have low sample counts due to PMW data uncertainty. Also, some areas between 0 to 60S and 120E to 60W have lower sample counts because of unavailability of GEO satellite data for rainfall estimation. Since this method is calibrated based on the CDF of climatology data, providing a sufficient number of samples for each year, validation and calibration, is essential. For statistical validation the area was divided to 8 longitudinal regions to make sure the study areas have enough samples and frozen lands can be excluded from other regions with sufficient sample counts. To evaluate the consistency of results one year of calibration and one year of validation is tested using statistical parameters. Figure 3.7 shows the average monthly rainfall map during winter 2011 over the globe. Some areas are highlighted with black circles to visualize the improvement. Circles A and B in Figure 3.7 display some areas where PERISIANN-CCS overestimates of rainfall, while MA-PERSIANN-CCS estimates are better matched with PMW data. Figure 3.8 is shows the average monthly rainfall during summer 2011 which is included in calibration period. PERSIANN-CCS overestimates rainfall in area C and underestimates it in area D during summer 2011. Figure 3.8.a displays MA-PERSIANN-CCS average monthly rainfall which gives rainfall estimates close to PMW as is highlighted in areas C and D.

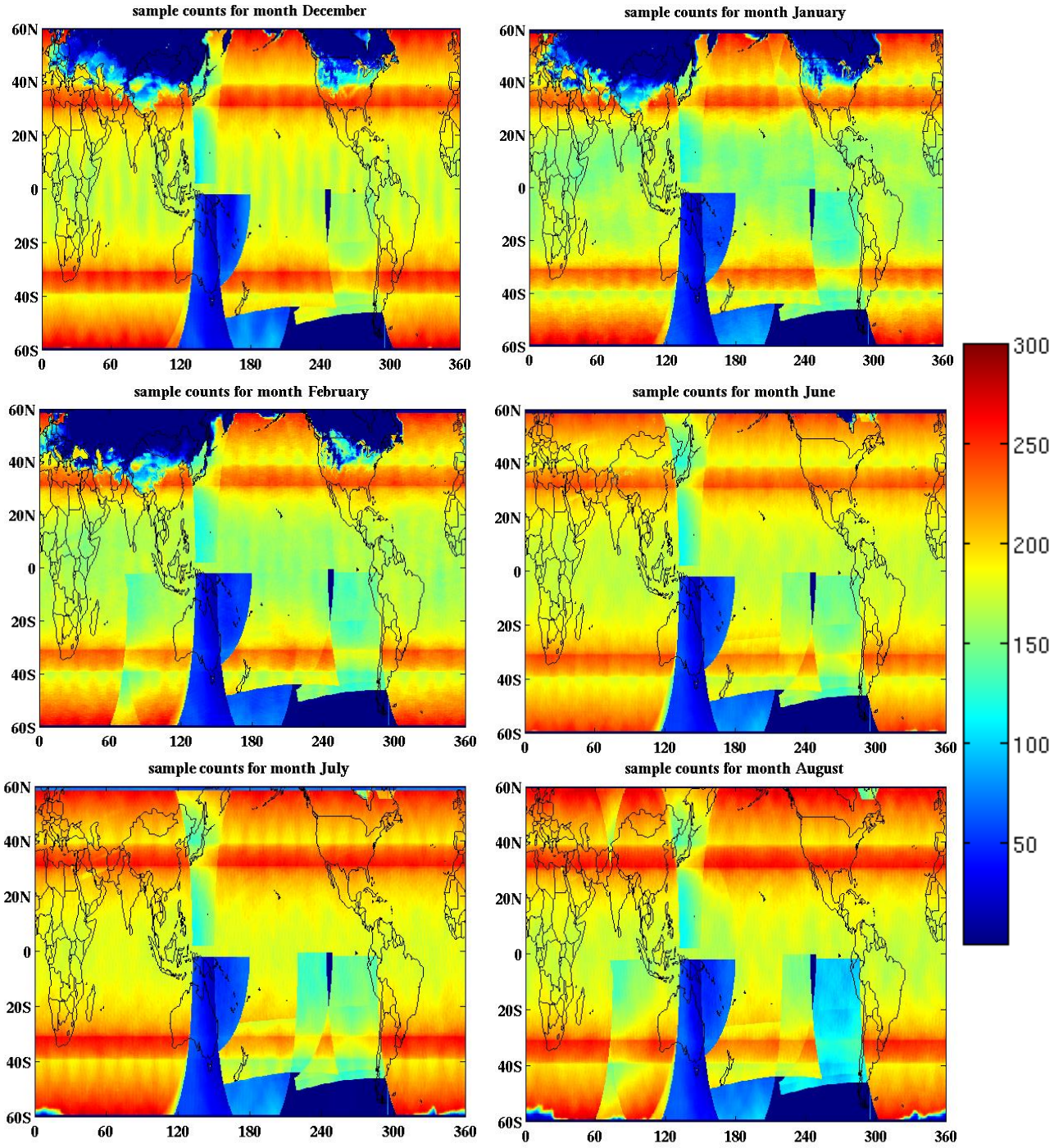


Figure 3.6: Number of available concurrent data samples for each pixel for winter 2012 (December, January, February) and Summer 2012 (June, July, August)

Figure a

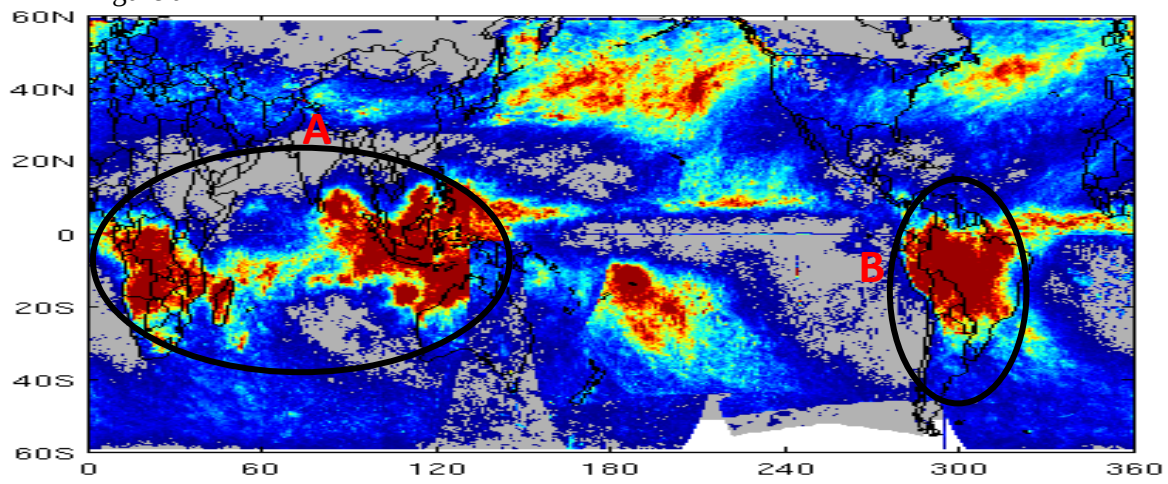


Figure b

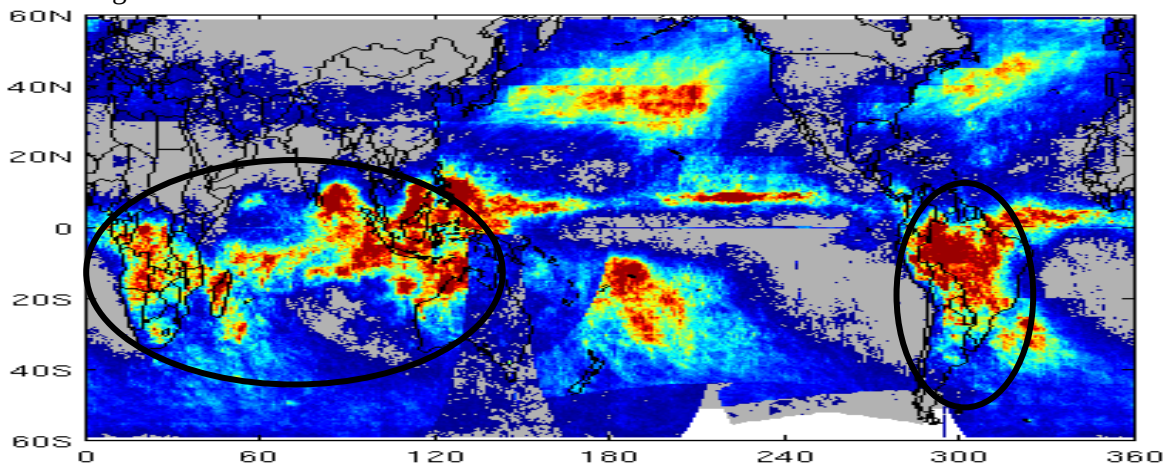


Figure c

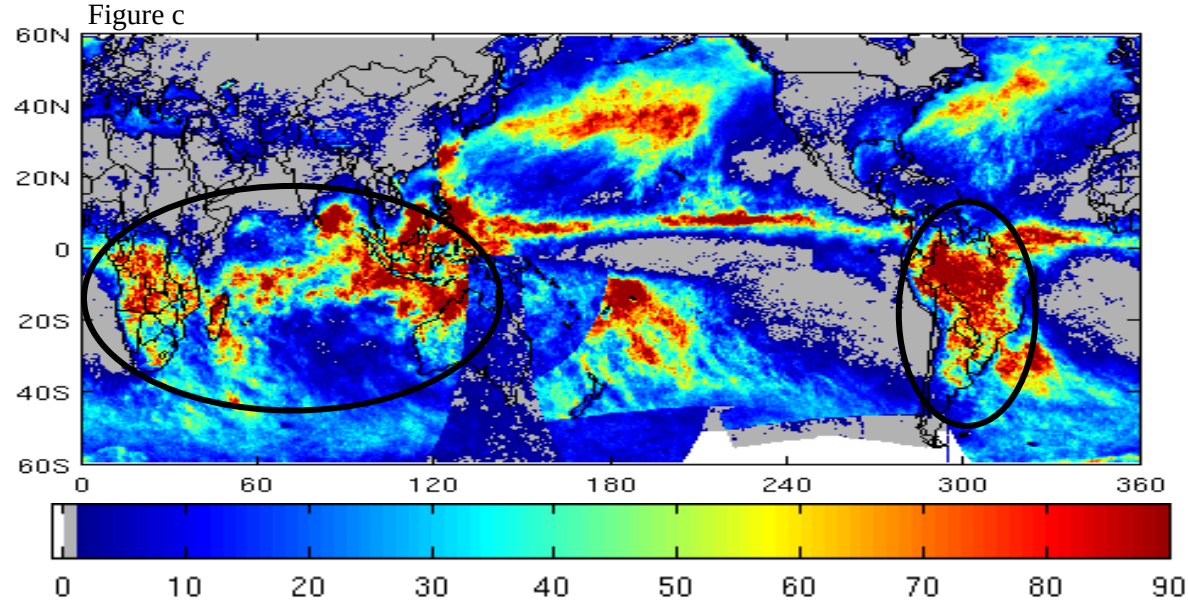


Figure 3.7: Average monthly rainfall calculated based on the concurrent samples from PMW and PERSIANN-CCS data during winter 2011 (mm/month-December, January, February). (a) PERSIANN-CCS (b MA-PERSIANN-CCS (c) PMW

Figure a

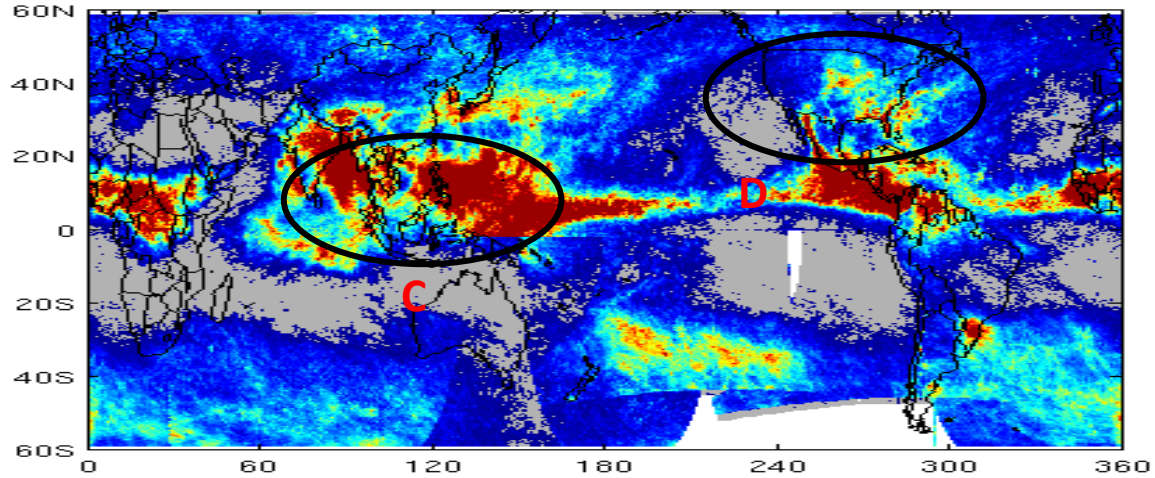


Figure b

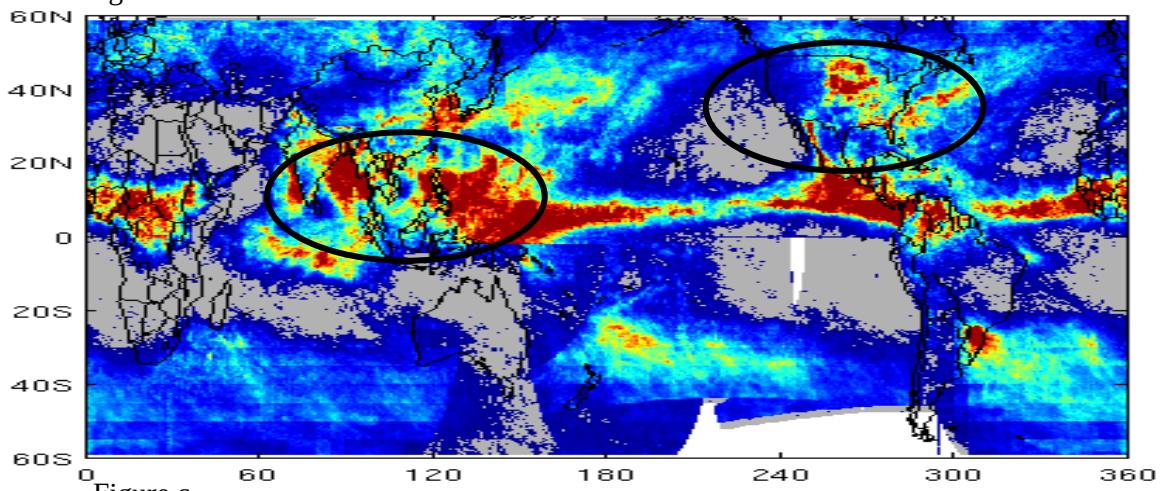


Figure c

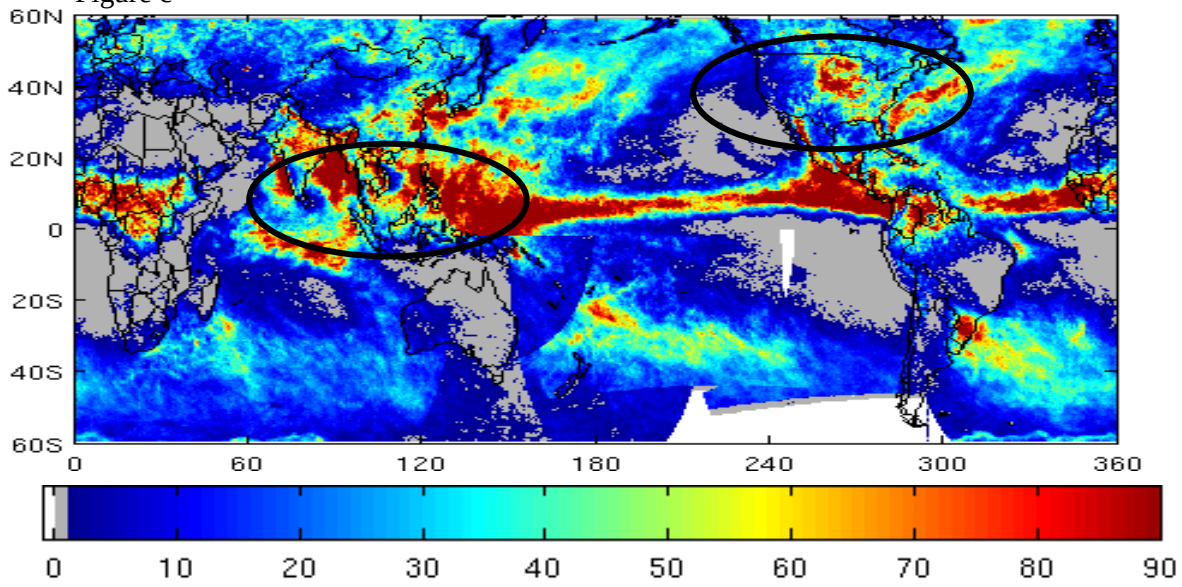


Figure 3.8: Average monthly rainfall calculated based on the concurrent samples from PMW and PERSIANN-CCS data during summer 2011 (mm/month-June, July, August). (a) PERSIANN-CCS (b MA-PERSIANN-CCS (c) PMW

We also validate performance on 2012, both summer and winter to evaluate PMM approach. Figure 3.9 shows the monthly average precipitation during winter 2012. The areas marked A and B are selected to illustrate a comparison between PERSIANN-CCS, MA-PERSIANN-CCS, and the PMW data. PERSIANN-CCS overestimates rainfall in regions A and B while MA-PERSIANN-CCS shows better correspondence with PMW rainfall data. Figure 3.10 displays monthly average rainfall during summer 2012. Regions C and D which were also selected in Figure 3.8 are selected in Figure 3.10 as well. PERSIANN-CCS overestimates rainfall over region C and underestimates it over region D, but MA-PERSIANN-CCS shows improved estimation in compared with PMW data.

In summary, based on visual comparison, PMM improves the rainfall estimation over different parts of the globe, and the results appear consistent during validation and calibration year over both summer and winter. A more significant adjustment is observed over high latitude regions, while tropical regions have less significant adjustment.

To quantify the PERSIANN-CCS, MA-PERSIANN-CCS, and PMW evaluation, the statistical parameters such as relative bias, correlation coefficient, and Root Mean Squared Error (RMSE) are calculated for 8 latitudinal zones over the globe in average monthly scale (see appendix A for the equations). Tables 3.1 and 3.2 display statistical parameters calculated for 8 zones during winter and summer 2011 respectively. To make sure that during evaluation we only consider regions with reliable data counts and quality, we concentrate over the southern hemisphere during winter and the northern hemisphere during summer. According to Table 3.1, zones located in southern hemisphere have improved relative bias, correlation coefficient, and RMSE after using bias adjustment. Also, MA-PERSIANN-CCS has better relative bias, correlation

coefficient, and RMSE than PERSIANN-CCS in comparison with PMW data during summer over the southern hemisphere (see table 3.2).

Figure a

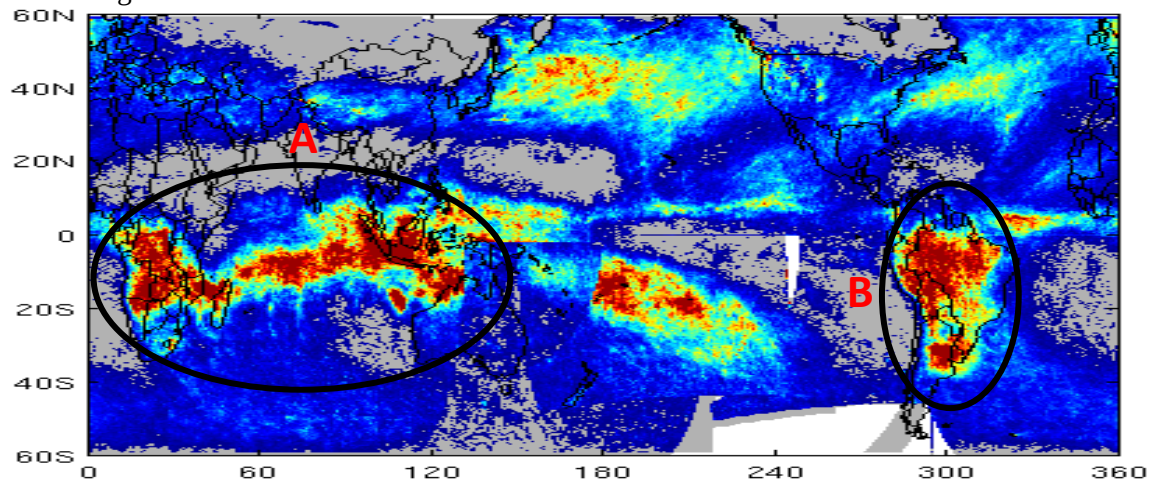


Figure b

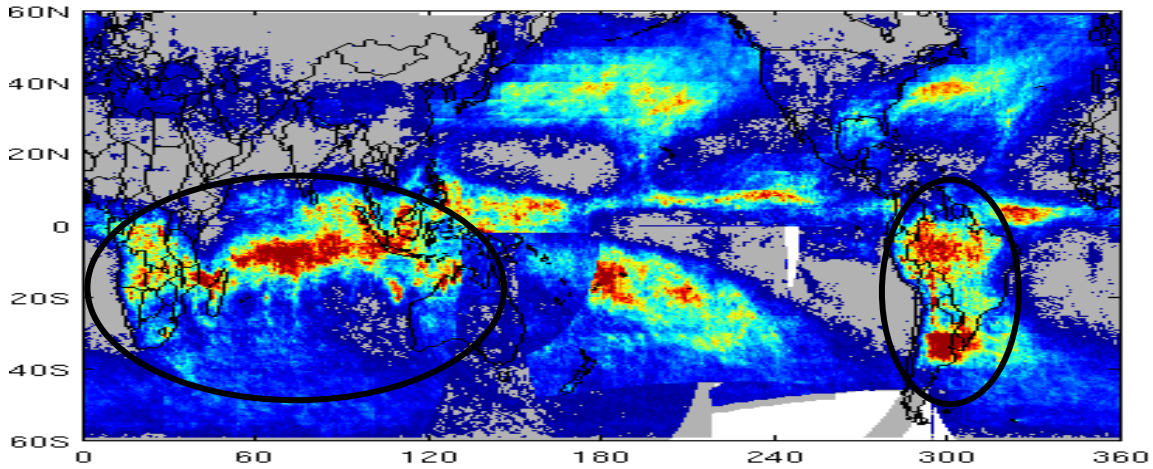


Figure c

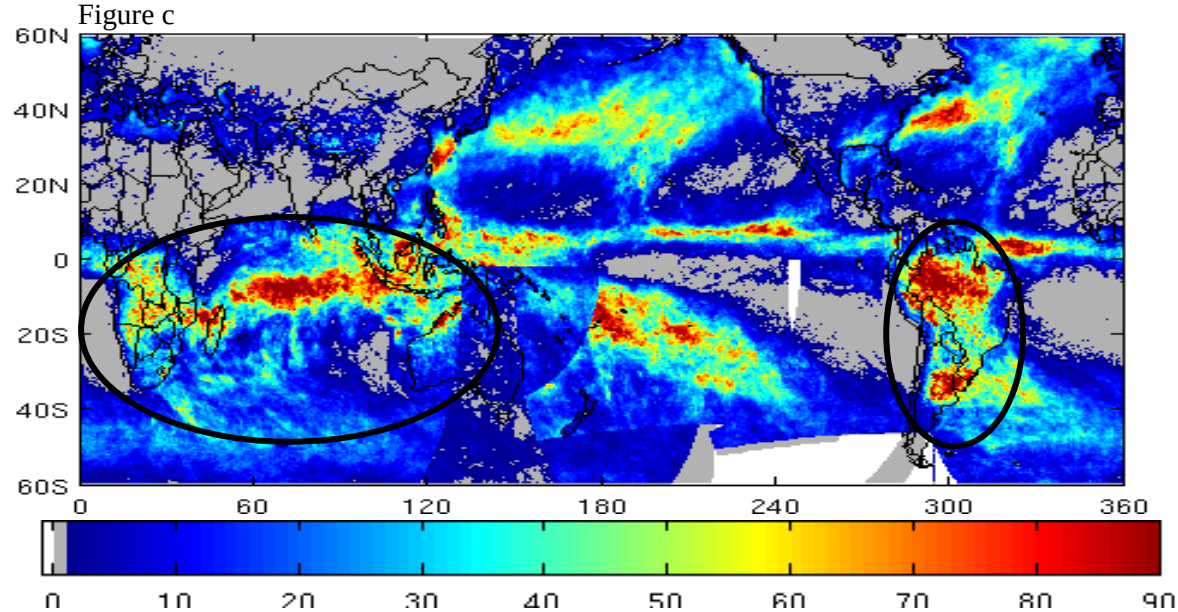


Figure 3.9: Average monthly rainfall calculated based on the concurrent samples from PMW and PERSIANN-CCS data during winter 2012 (mm/month-December, January, February). (a) PERSIANN-CCS (b) MA-PERSIANN-CCS (c) PMW

Figure a

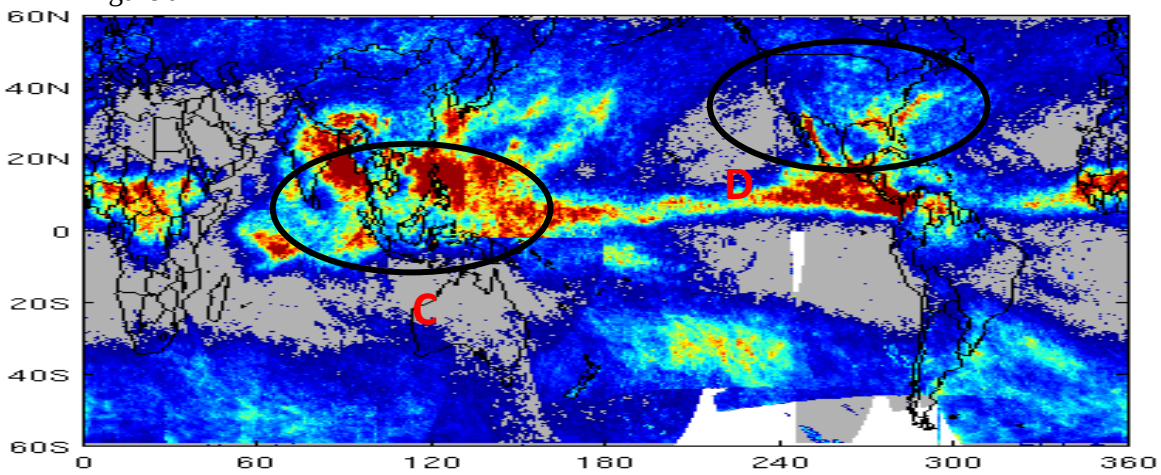


Figure b

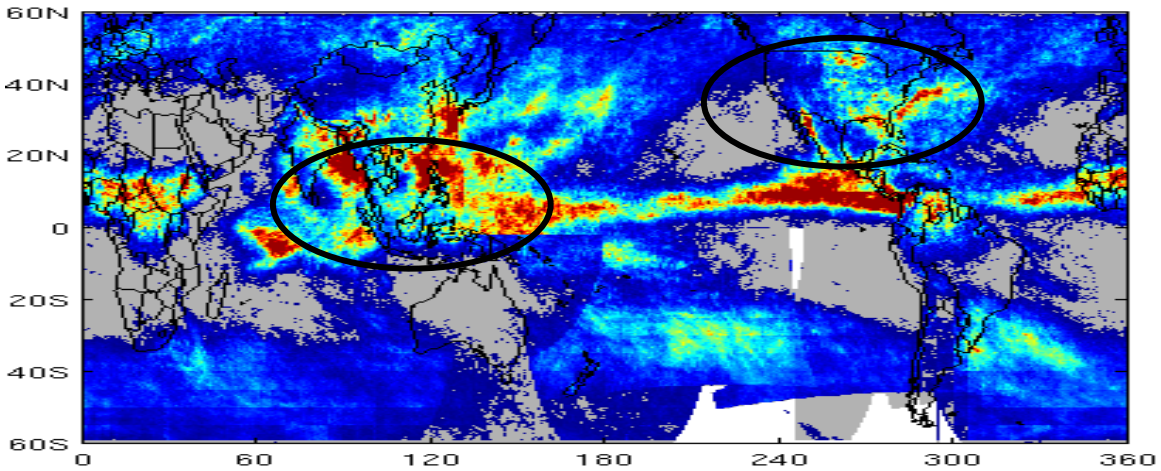


Figure c

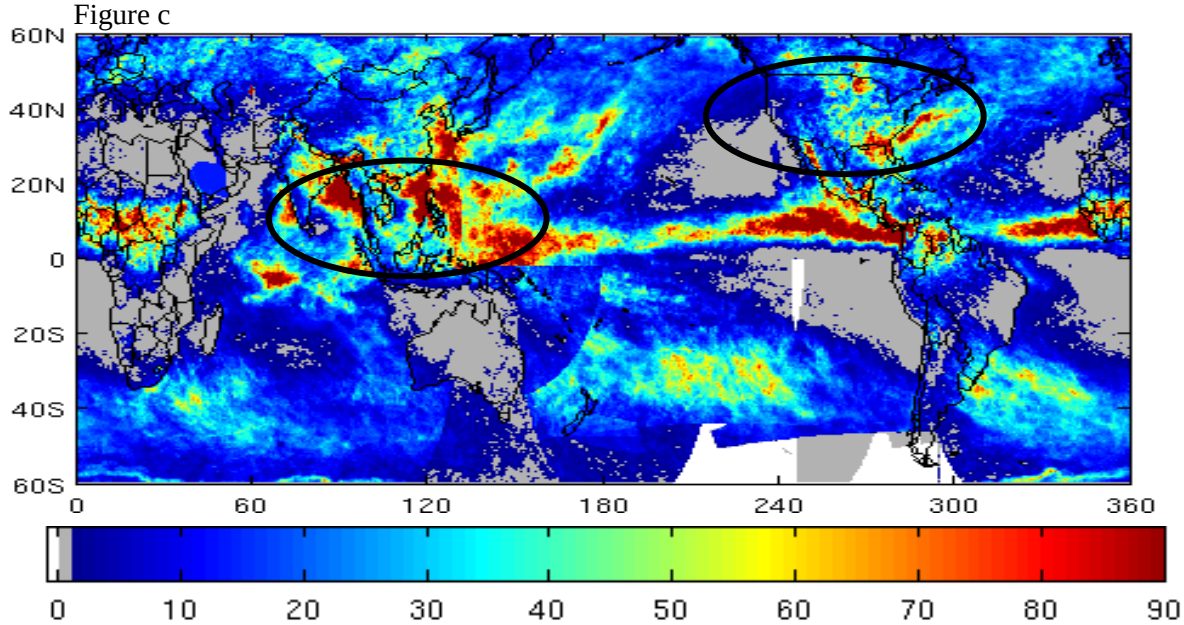


Figure 3.10: Average monthly rainfall calculated based on the concurrent samples from PMW and PERSIANN-CCS data during summer 2012 (mm/month-June, July, August). (a) PERSIANN-CCS (b) MA- PERSIANN-CCS (c) PMW

Table 3-1: Statistical parameters used for global validation based on average monthly rainfall. Parameters are calculated for 8 zones comparing with PMW data (December, January, and February 2011)

	Winter-2011					
	PERSIANN-CCS			MA-PERSIANN-CCS		
zone	Relative Bias	Corr.	RMSE (mm/month)	Relative Bias	Corr.	RMSE (mm/month)
60N-45N	0.96	0.69	17.87	-0.11	0.91	6.38
45N-30N	0.37	0.78	16.81	-0.06	0.93	8.75
30N-15N	-0.29	0.46	13.74	-0.30	0.70	10.99
15N-0	-0.05	0.86	17.51	-0.08	0.93	11.88
0-15S	0.28	0.91	22.64	-0.05	0.94	11.66
15S-30S	0.1	0.88	17.57	-0.08	0.94	9.79
30S-45S	-0.52	0.71	16.91	-0.31	0.87	11.15
45S-60S	-0.62	0.66	13.08	-0.51	0.8	11

Table 3-2: Statistical parameters used for global validation based average monthly rainfall. Parameters are calculated for 8 zones comparing with PMW data (June, July, and August 2011)

	Summer-2011					
	PERSIANN-CCS			MA-PERSIANN-CCS		
zone	Relative Bias	Corr.	RMSE (mm/month)	Relative Bias	Corr.	RMSE (mm/month)
60N-45N	-0.50	0.43	18.54	-0.23	0.66	12.57
45N-30N	-0.23	0.79	15.67	-0.09	0.91	9.99
30N-15N	0.11	0.91	14.89	-0.07	0.94	11
15N-0	0.02	0.86	22.72	-0.06	0.93	15.54
0-15S	-0.19	0.93	7.90	-0.17	0.95	6.78
15S-30S	-0.39	0.84	9.43	-0.23	0.90	7.40
30S-45S	-0.07	0.66	12.27	-0.1	0.87	7.27
45S-60S	0.72	0.55	13.76	-0.07	0.76	6.06

The statistical parameters for comparing PERSIANN-CCS and MA-PERSIANN-CCS with PMW data are calculated for winter and summer 2012 as well. Table 3.3 displays the comparison over 8 latitudinal zones during winter 2012. The results imply relative bias, correlation coefficient, and RMSE for MA-PERSIANN-CCS are significantly improved over the southern hemisphere during winter time. Also, statistical calculations during summer 2012 are shown in Table 3.4 which, shows improvement by MA-PESIANN-CCS compared with PERSIANN-CCS during summer 2012 over northern hemisphere.

Table 3-3: Statistical parameters used for global validation based on average monthly rainfall. Parameters are calculated for 8 zones comparing with PMW data (December, January, and February 2012

	Winter-2012					
	PERSIANN-CCS			MA-PERSIANN-CCS		
zone	Relative Bias	Corr.	RMSE (mm/month)	Relative Bias	Corr.	RMSE (mm/month)
60N-45N	0.72	0.61	14.16	-0.23	0.86	6.62
45N-30N	0.26	0.71	15.28	-0.13	0.89	9.60
30N-15N	-0.13	0.38	11.58	-0.19	0.68	8.96
15N-0	-0.06	0.83	12.12	-0.05	0.92	8.63
0-15S	0.27	0.90	18.63	0.0008	0.94	9.32
15S-30S	0.08	0.87	13.72	-0.08	0.92	8.53
30S-45S	-0.39	0.79	11.83	-0.12	0.86	9.32
45S-60S	-0.51	0.69	8.71	-0.37	0.80	6.97

Table 3-4: Statistical parameters used for global validation based on average monthly rainfall. Parameters are calculated for 8 zones comparing with PMW (June, July, and August 2012)

	Summer-2012					
	PERSIANN-CCS			MA-PERSIANN-CCS		
zone	Relative Bias	Corr.	RMSE (mm/month)	Relative Bias	Corr.	RMSE (mm/month)
60N-45N	-0.50	0.52	13.95	-0.23	0.71	9.34
45N-30N	-0.25	0.80	12.29	-0.11	0.89	8.96
30N-15N	0.06	0.91	13.24	-0.08	0.94	10
15N-0	0.05	0.87	16.55	-0.01	0.93	12.11
0-15S	-0.07	0.91	7.36	-0.04	0.93	6.79
15S-30S	-0.44	0.86	7.79	-0.28	0.88	6.49
30S-45S	-0.25	0.70	11.81	-0.29	0.85	9.55
45S-60S	0.32	0.5	10.15	-0.30	0.72	7.08

3.4.3. Validation of the CONUS

In this section, we used ground based radar data over the CONUS to evaluate satellite data precipitation estimates. Radar rainfall estimation from the next generation QPE (Q2) project, a joint initiative between NOAA/National Severe Storm Laboratory (NSSL), the Salt River Project (SRP), and the Federal Aviation Administration's Aviation Weather Research Program (Chen et. al 2012 and Zhang et al. 2011), is used in the evaluation of PERSIANN-CCS rainfall estimates before and after the adjustment. Radar Q2 uses NEXRAD and rain gauge observations for rainfall estimation at 1-km resolution every 5 minutes. At this time, NSSL is developing quantitative precipitation estimation which produces a 3D radar mosaic grid from the precipitation (Q3) over CONUS. In this study, Q2 estimation is processed to 0.25x0.25 degree resolution, 0.5-hourly, for PERSIANN-CCS rainfall evaluation after and before adjustment

during validation year 2012. Figure 3.11 shows the average monthly rainfall during the winter season over the US based on the concurrent samples of PERSIANN-CCS, PMW, and radar Q2 for 2012 (December, January, and February). As shown in the figure, PERSIANN-CCS overestimates the rainfall over western areas and underestimates over southern regions during the winter season of 2012 over the CONUS, while MA-PERSIANN-CCS shows closer estimates to PMW. Comparing average monthly rainfall during winter from PERSIANN-CCS, MA-PERSIANN-CCS, and PMW with radar Q2 rainfall data, it can be seen that PERSIANN-CCS and MA-PERSIANN-CCS underestimate the rainfall events over the southern regions of the U.S. MA-PERSIANN-CCS does not show improved rainfall estimation as compared with radar Q2 because PMW and radar Q2 rain estimation are not well-matched during the winter season. The dissimilarity between PMW and radar Q2 is a result of the performance of PMW sensors over high latitude regions during the winter time. PMW sensors not only have lower sample counts but also fail to catch rainfall during the cold season in high latitudes (Behrangi et al. 2014). Figure 3.12 shows the monthly average precipitation during the summer season 2012 (June, July, and August) over the CONUS for PERSIANN-CCS, MA-PERSIANN-CCS, PMW and Radar Q2. According to Figure 3.12, PERSIANN-CCS underestimates the rainfall over the U.S. particularly over the northeast and eastern parts compared with PMW and Radar Q2 during summer season 2012. Figure 3.12-b displays the MA-PERSIANN-CCS which has rainfall values more similar with PMW and radar Q2. For evaluating the performance of PERSIANN-CCS and MA-PERSIANN-CCS in comparison with PMW and radar Q2, statistical parameters such as relative bias, correlation coefficient and, root mean square error (RMSE) are calculated for winter and summer 2012 (Tables 3.5 and 3.6). Table 3.5 gives the statistical results for PERSIANN-CCS and MA-PERSIANN-CCS compared to PMW and Table 3.6 shows the

statistical performance of PERSIANN-CCS, MA-PERSIANN-CCS, and PMW against radar data as ground truth. During winter 2012, relative bias, correlation coefficient and RMSE between MA-PERSIANN-CCS and PMW shows improvement over those from PERSIANN-CCS and PMW. This implies that implementation of climatology data over the validation year 2012 has shifted the PERSIANN-CCS rainfall estimation toward PMW rainfall data. However, MA-PERSIANN-CCS doesn't show any improvement in terms of rainfall estimation for winter time compared with radar Q2. The statistical parameters show that the relative bias value of -0.37 between PERSIANN-CCS and radar data decreased to -0.69 (more underestimation) between MA-PERSIANN-CCS and radar Q2. The reason for this is due to inconsistencies between PMW and radar Q2 data. Statistical validation results for summer are also exhibited in Tables 3.5 and 3.6. As expected, all the statistics for MA-PERSIANN-CCS show improvement from PERSIANN-CCS compared with PMW. When compared to radar Q2, MA-PERSIANN-CCS shows improvement in relative bias by 53% during the summer season with no significant changes in the other two statistics.

In summary we have provided a systematic approach for reducing the bias between the GEO-based PERSIANN-CCS and PMW estimates. The next step towards further improvement will be focused on increasing the population of PMW samples over high latitude and cold seasons. The use of CLOUDSAT and ground-based radars covering the higher latitudes are currently under study.

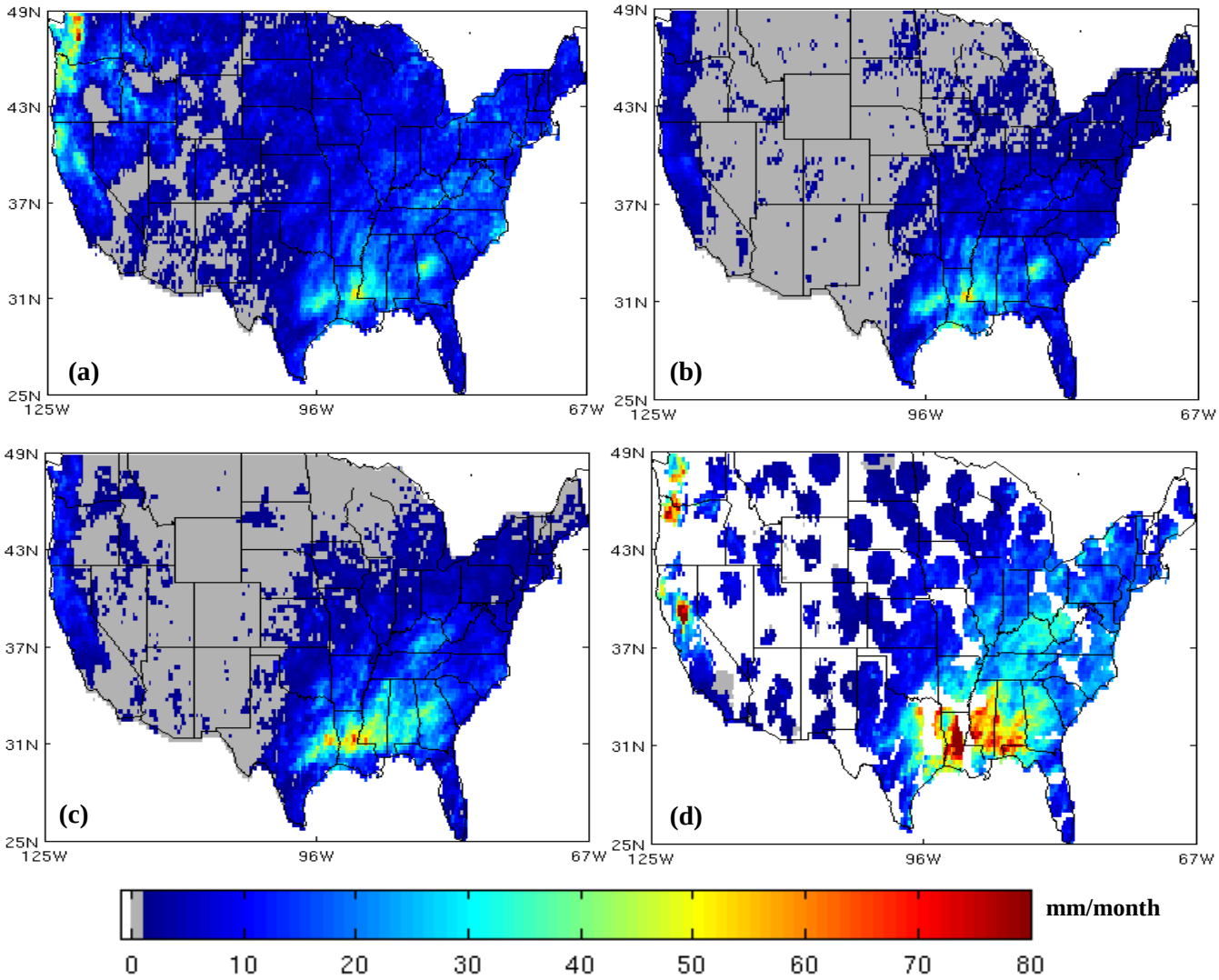


Figure 3.11: Monthly average rainfall calculated based on the concurrent samples during winter 2012. (a) PERSIANN-CCS (b) MA- PERSIANN-CCS (c) PMW (d) Radar Q-2

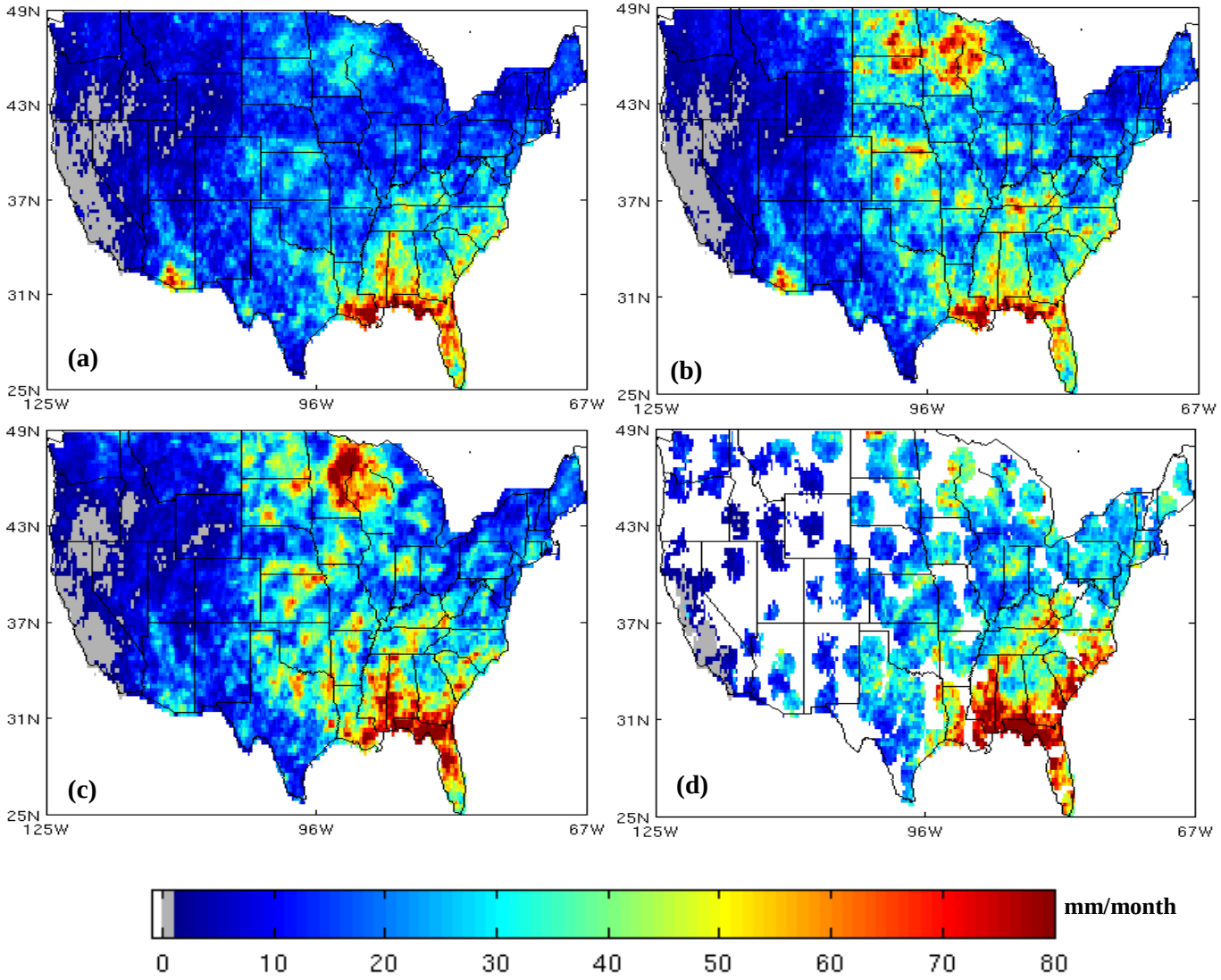


Figure 3.12: Monthly average rainfall calculated based on the concurrent samples in summer 2012. (a) PERSIANN-CCS (b MA-PERSIANN-CCS (c) PMW (d) Radar Q-2

Table 3-5: Comparison between PERSIANN-CCS and MA-PERSIANN-CCS with PMW satellite data over CONUS during winter and summer 2012. Parameters are calculated based on average monthly rainfall.

	CONUS year 2012					
	PERSIANN-CCS			MA-PERSIANN-CCS		
	Relative Bias	Corr.	RMSE (mm/month)	Relative Bias	Corr.	RMSE (mm/month)
Winter	0.59	0.61	8.35	-0.23	0.86	4.86
Summer	-0.25	0.8	13.51	-0.07	0.84	10.93

Table 3-6: Comparison between PERSIANN-CCS, MA-PERSIANN-CCS, and PMW with Q2 ground based radar over CONUS during winter and summer 2012. Parameters are calculated based on average monthly rainfall.

	CONUS year 2012								
	PERSIANN-CCS			MA-PERSIANN-CCS			PMW		
	Relative Bias	Corr.	RMSE (mm/month)	Relative Bias	Corr.	RMSE (mm/month)	Relative Bias	Corr.	RMSE (mm/month)
Winter	-0.37	0.67	13.78	-0.69	0.7	16.83	-0.6	0.76	14.86
Summer	-0.3	0.82	14.41	-0.14	0.72	14.75	-0.07	0.8	12.95

CHAPTER 4: Improving warm-clouds rainfall detection in PERSIANN-CCS

4.1. Introduction

The PERSIANN-CCS algorithm relies on cloud brightness temperature from an IR channel for precipitation estimation, which increases the uncertainty of rainfall analysis. Also, PERSIANN-CCS can estimate rainfall from clouds with certain top temperature which misses the rainfall from warmer clouds. Some previous studies have been done to improve PERSIANN-CCS precipitation estimates. Behrangi et al. (2010a) combines infrared and visible channels to enhance the ability of PERSIANN-CCS to capture warm cloud rainfall and distinguish thin cirrus clouds without rain. Mahrooghi et al. (2013) used lightning data as an extra feature in addition to the extracted features from IR data for PERSIANN CCS. Using co-located lightning data with the segmented cloud patches and categorizing clouds to electrified patches and non-electrified patches improved the rainfall identification during winter and fall and also enhanced the quantitative estimation during all the seasons in terms of bias, correlation coefficient, and RMSE. Also Mahrooghi et al (2012) used wavelet and selected features (WSF), median merging (MM), and selected curve-fitting (SCF) techniques to improve the PERSIANN-CCS precipitation estimation. In this chapter, we discuss the impacts of raising the cloud temperature threshold to warmer temperature in order to capture the rainfall from warm clouds as well.

4.2.Extending cloud top brightness temperature

PERSIANN-CCS estimates the rainfall based on three temperature thresholds (220, 235, and 253K). In this research we extended the threshold up to 300 K to include the warm clouds as well. Figure 4.1 displays the infrared image for one time step and the corresponding segmentations for maximum temperature threshold equal to 253 K and 300 K. From Figure 4.1,

we see that by increasing the temperature threshold, we obtained more cloud patches which result in more potential for detecting precipitation. The details of training the PERSIANN-CCS algorithm will be explained in the next section.

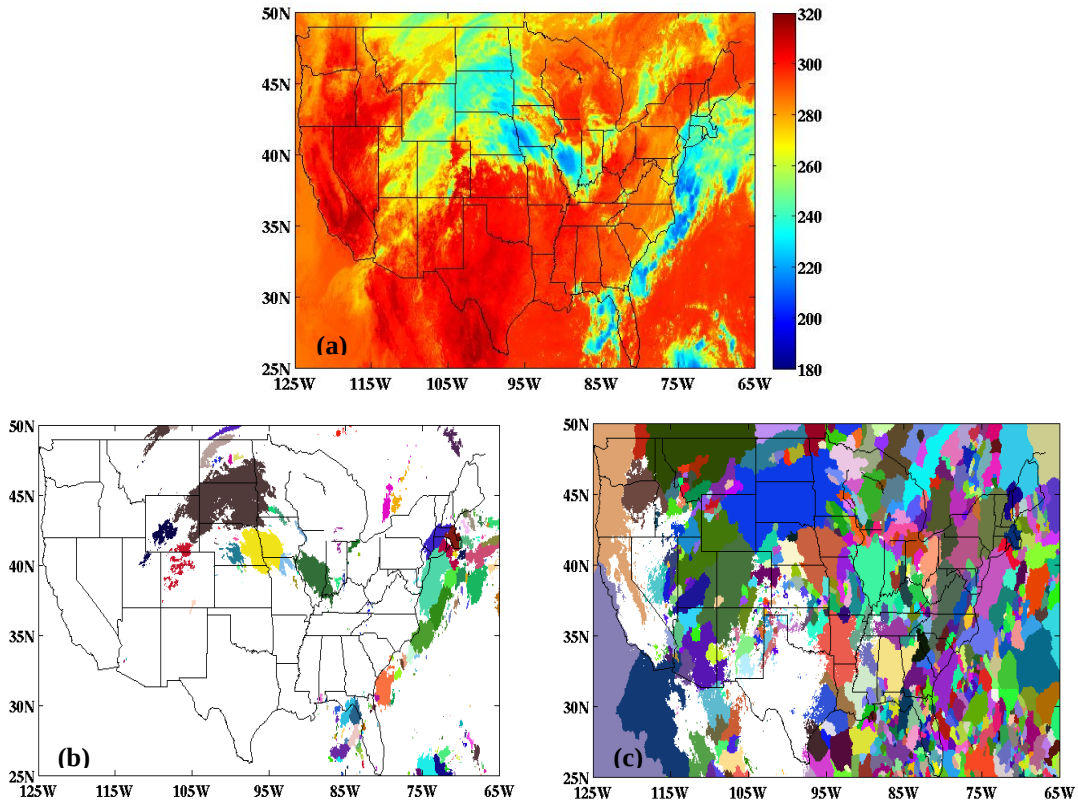


Figure 4.1: IR image and segmentation based on temperature threshold maximum equal to 253 K and 300 K. (a) IR image for September 09, 2014, UTC 17:30 (b) Segmentation based on Tb=253K for September 09, 2014, UTC 17:30 (c) Segmentation based on Tb = 300 K for September 09, 2014, UTC 17:30

4.3. Training the algorithm

For training the algorithm, we have used IR data from GOES satellite and corresponding precipitation data from MRMS over the CONUS. The training period is selected from June to September 2014 since the warm cloud rainfall is more common during summer season. The first part of the training is cloud segmentation. The temperature threshold for the cloud segmentation

is raised up to 300 K. After the segmentation, PERSIANN-CCS extracts coldness, geometry, and texture features from cloud patches. We also define similar features to PERSIANN-CCS and add a few more for the new threshold level. The original PERSIANN-CCS algorithm uses 12 different features from cloud patches. In this dissertation research we have increased the number of features to 22 to cover cloud patches up to 300 K. Given defined features, PERSIANN-CCS uses these features to classify cloud patches to a number of groups based on a neural network technique. We used k-means method to cluster the cloud patches into different groups. K-means clustering is a classical unsupervised learning methods that groups a data set by following iterative process 1) Assuming a value for k, the number of cluster, choose cluster centroids randomly; 2) Compute the Euclidean distance from the centroids and assign each element to the closest centroid 3) Re-compute the centroids based on the assignments from the previous step and repeat step 2 and 3 until clusters converge. Figure 4.2 shows the K-means clustering method procedure. In this study we selected k = 200 clusters for cloud classification. After selecting the number of groups that we want to divide the cloud patches into, we use the k-means algorithm to cluster cloud patches based on the defined features. K-means algorithm classifies cloud patches based on minimizing the squared error function defined in equation 1:

$$z_i = argmin \|x_i - \mu_c\|^2 \qquad \forall_i \quad (1)$$

$$\forall_c \qquad \mu_c = \frac{1}{N_c}\sum_{i \in S_c} x_i \qquad S_c = \{i : z_i = c\}, N_c = |S_c|$$

where:

μ_c is the mean of all assigned data to each cluster

x_i is/are the features of data ith

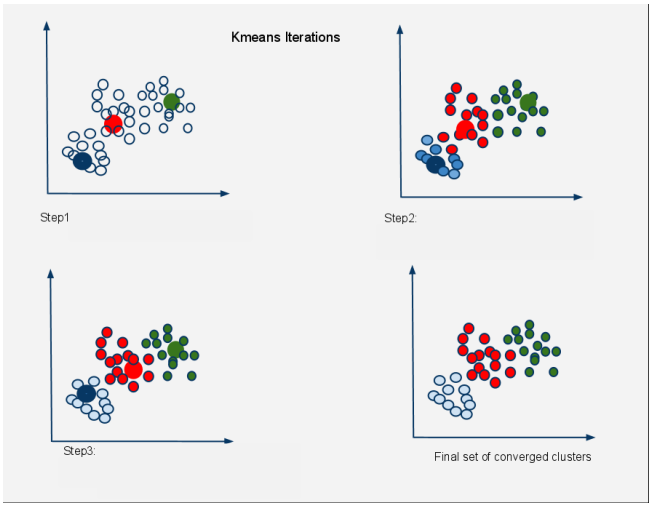


Figure 4.2: K-means clustering method procedure

http://humble-developer.blogspot.com/2011/01/kmeans-clustering-algorithm-part-1.html

After classifying clouds to 200 groups, we estimate pixels rainfall based on the cloud patches and corresponding rainfall rate in each group. In order to define a relationship between rainfall and cloud top temperature we assume that colder clouds are associated with heavier precipitation so we redistribute the Tb-RR relationship and assign heavier rainfalls to colder clouds. The redistributed pixels are fitted with a nonlinear exponential function defined in equation 2 for each patch group using least square regression method. For cluster j, the Tb-RR relationship is specified as:

$$R^j = v_6^j + v_2^j . exp\left[v_3^j . (T_b + v_4^j)^{v_5^j}\right] \qquad (2)$$

Figure 4.3 demonstrates the Tb-RR relationship for 200 clusters based on the maximum cloud temperature threshold equal to(Tb)max = 253 K and Figure 4.4 shows the Tb-RR relationship calibrated based on the maximum cloud temperature threshold equal to 300 K.

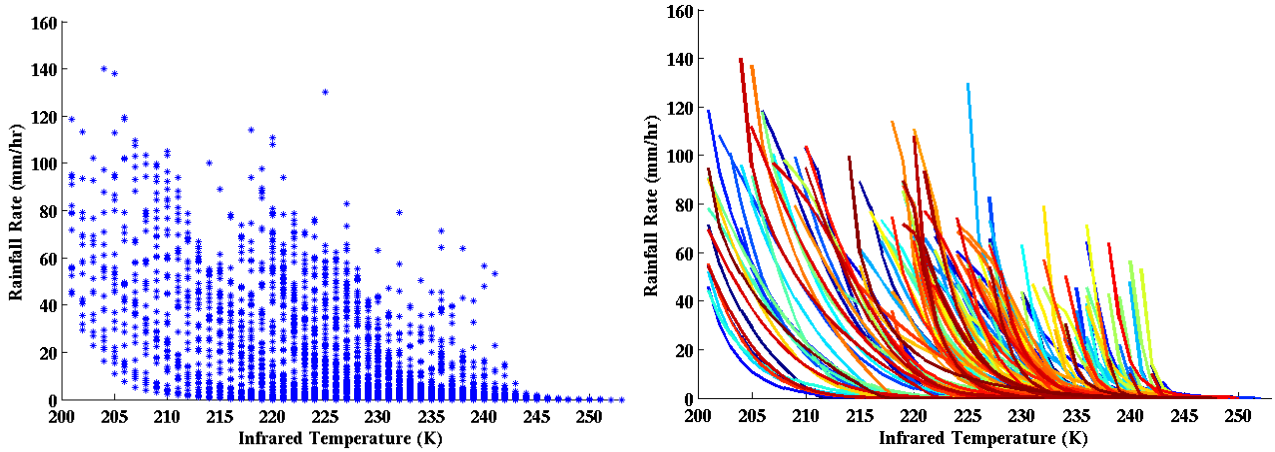


Figure 4.3: The calibrated cloud-rainfall relationships (Tb-RR) from the PERSIANN-CCS algorithm using IR and radar data. Each curve referring to a cloud cluster in the K-means classification for maximum temperature threshold 253 K.

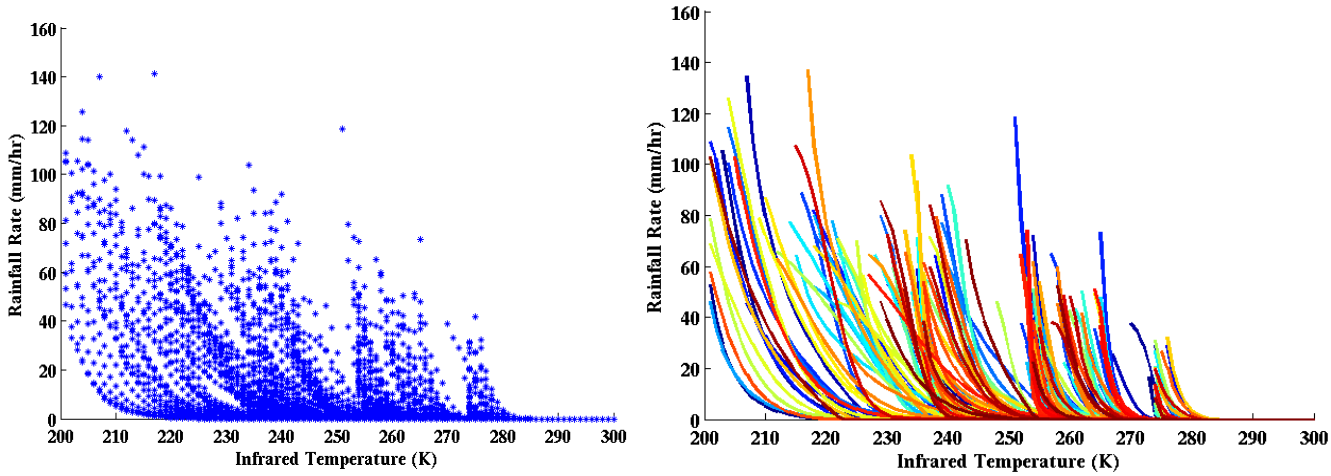


Figure 4.4: The calibrated cloud-rainfall relationships (Tb-RR) from the PERSIANN-CCS algorithm using IR and radar data. Each curve referring to a cloud cluster in the K-means classification for maximum temperature threshold 300 K

4.4.Results and Discussion

For model evaluation, we validate the PERSIANN-CCS algorithm during training and testing period against MRMS data. The training data is consists of 4 samples per day from June to September 2014. Figure 4.5 shows the total accumulated rainfall over the US during training

period from PERSIANN-CCS $(Tb)_{max}$ = 253 K and $(Tb)_{max}$ = 300 K) and MRMS data. Based on statistical parameters displayed in Table 4.1, increasing the brightness temperature threshold improves rainfall estimation and decreases the relative bias from -0.27 to -0.10. No significant change is observed for correlation and RMSE.

The impact of extending the cloud top brightness temperature threshold in the PERSIANN-CCS algorithm is also evaluated during the validation period, from June to September 2015, over the US. Figure 4.6 shows average monthly rainfall during the validation period (June-September 2015). Based on Table 4.2, the statistical parameters calculated for monthly average rainfall during the validation period show that the average relative bias improves for PERSIANN-CCS after increasing the cloud top brightness temperature threshold, while average correlation coefficient and average RMSE do not change significantly.

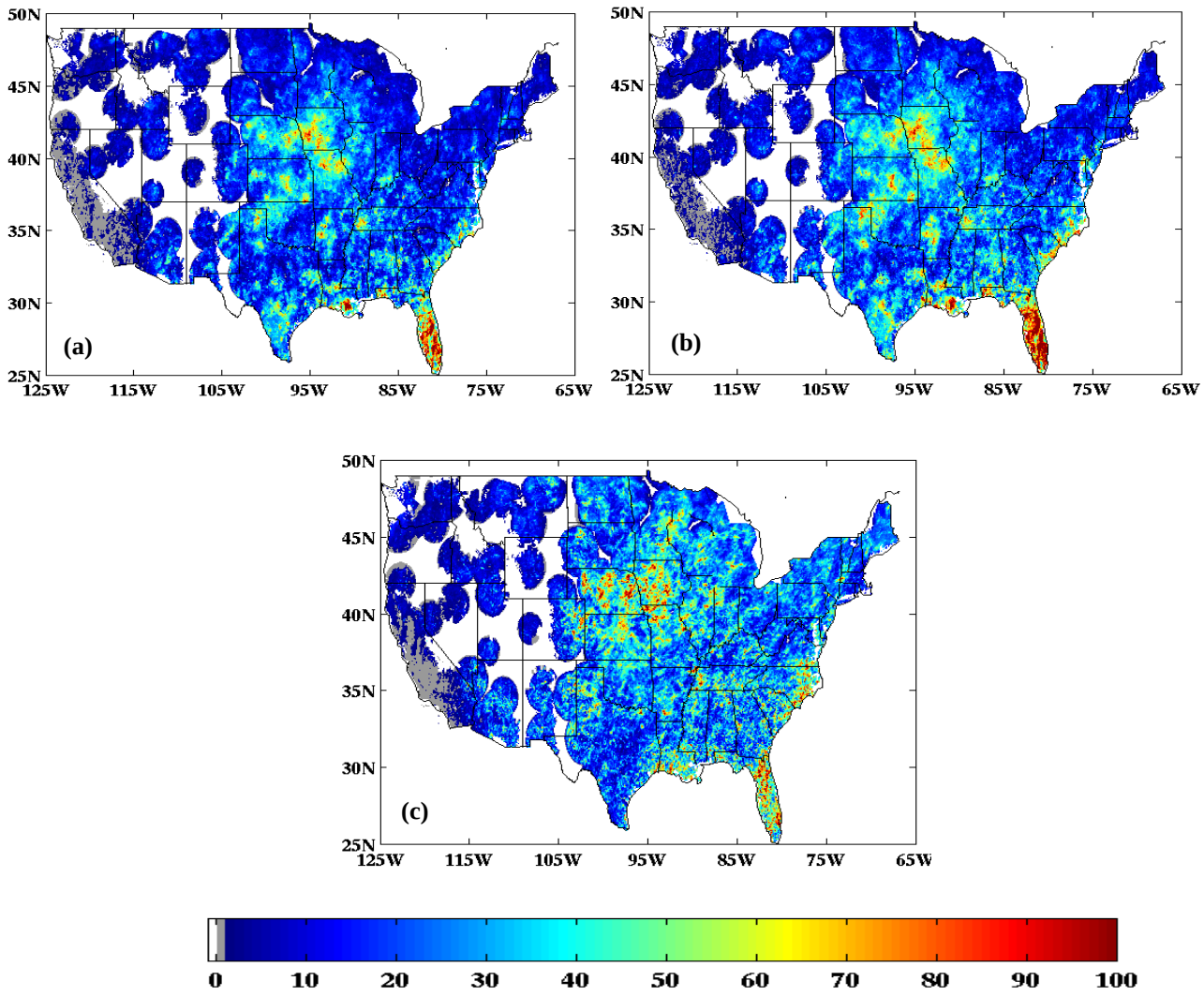


Figure 4.5: Accumulated rainfall from training data (June-September 2014). (a) PERSIANN-CCS (Tb = 253K); (b) PERSIANN-CCS (Tb = 300K); (c) MRMS radar data. (mm)

Table 4-1: Statistical evaluation during training period (June to September 2014). The numbers are average of relative bias, correlation coefficient and RMSE over the entire training data

	Relative Bias	Correlation	RMSE (mm/hr)
Tb_{max} = 253K	-0.27	0.24	1.50
Tb_{max} = 300K	-0.1	0.24	1.53

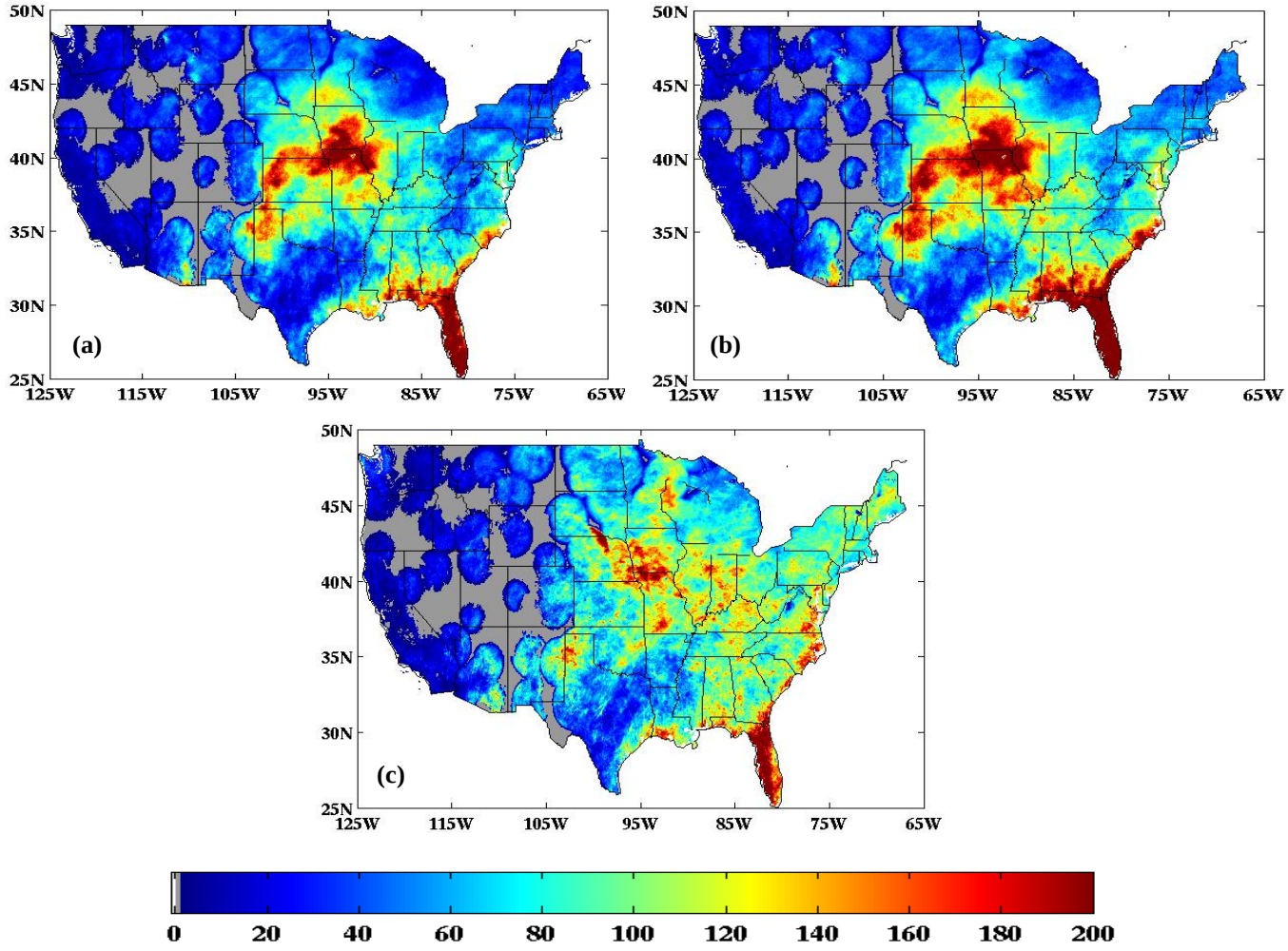


Figure 4.6: monthly average rainfall during validation period (June-September 2015). (a) PERSIANN-CCS (Tb = 253 K); (b) PERSIANN-CCS (Tb = 300 K); (c) MRMS radar data. (mm)

Table 4-2: Statistical evaluation during validation period (June to September 2015). The numbers are average of relative bias, correlation coefficient and RMSE over the entire validation data

	Relative Bias	Correlation	RMSE (mm/hr)
Tb_{max} = 253K	-0.15	0.23	1.43
Tb_{max} = 300K	0.06	0.23	1.48

CHAPTER 5: Improving rainfall estimation of PERSIANN-CCS using probabilistic approach

5.1. PERSIANN-CCS rainfall estimation method

As discussed in Chapter 2, the pixel based rainfall estimation for PERSIANN-CCS is based on fitting a nonlinear exponential function to redistributed precipitation rate (RR) and brightness temperatures (Tb) for each cloud group. The redistribution of data is based on assuming that the heavier rainfalls are associated with colder clouds. This assumption causes missing rainfall from warm clouds and false alarms from non-precipitating cold clouds, since we assign the rainfall to cold clouds. Figure 5.1 shows the MRMS rainfall rate as a function of the corresponding cloud brightness temperature from the IR channel. The distribution of rainfall shifts to higher rates with colder clouds. However, they are many clouds with warmer temperature that also have positive rain rate.

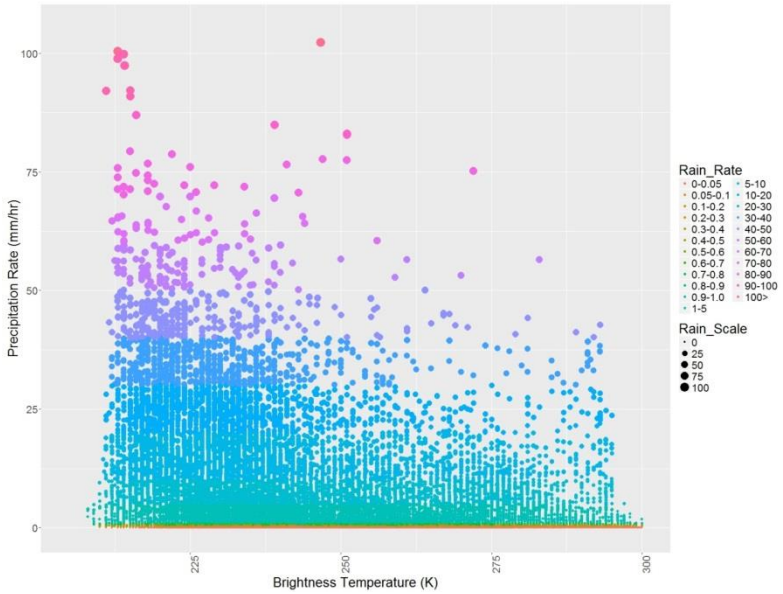


Figure 5.1: Dispersion in the relation Tb (IR)-RR, including rain/no-rain and positive values for one cluster

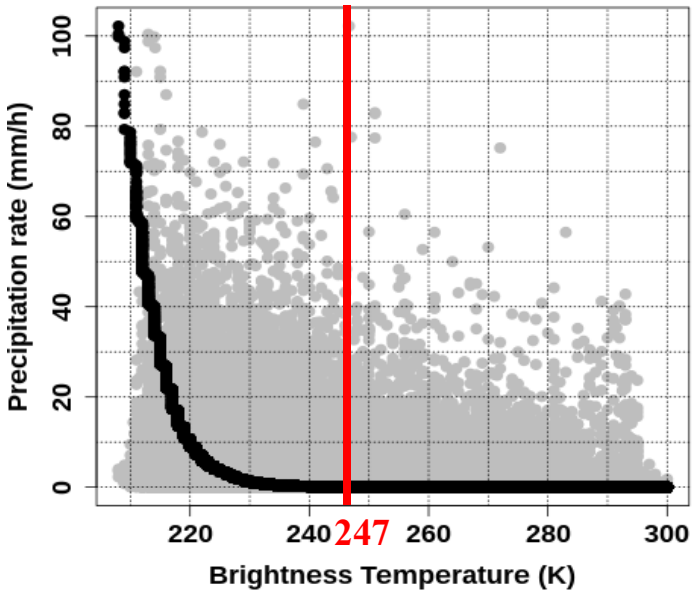


Figure 5.2: Redistribution of rainfall and cloud top temperature from one cloud cluster (black); The red line separates the zero rates and the positive rates in the redistributed data; the actual rainfall associated with brightness temperatures are shown in grey.

Figure 5.2 displays the Tb-RR relationship after redistributing the precipitation rates and brightness temperature to fit the nonlinear exponential function for the same cluster shown in Figure 5.1. After the redistribution all the rainy pixels are assigned to the coldest brightness temperatures and all the zero rain rates are assigned to brightness temperatures greater than 247 K. This brightness temperature threshold depends on the proportion of zero values in the rainfall rate dataset. In the retrieval, any clouds with temperature warmer than 247 K will be assigned a zero rain rate. Although the cloud temperature threshold has been extended to 300K, the redistribution of data assigns the positive rainfall rates to colder clouds. For this reason we cannot fully improve the rainfall detection from PERSIANN-CCS with the current rain retrieval method. The brightness temperature versus rainfall rate relation needs to be redesigned. In the next section, we introduce the new method for rainfall estimation based on the actual Tb-RR relationship.

5.2. Using probabilistic estimation instead of exponential regression model

In this section we introduce the new method that is used to retrieve precipitation from IR data. Kirstetter et al. (2015) used a probabilistic precipitation approach to estimate rainfall using ground-based radar observation. The surface reflectivity factor Z is related to a distribution of true precipitation rates that accounts for influential meteorological and observational factors. A similar approach is used here to improve the precipitation retrieval.

5.3. Methodology

The structure of this algorithm is very similar to what we have used to retrieve warm cloud precipitation from PERSIANN-CCS, except that instead of using a nonlinear exponential function, a probabilistic model of the Tb-RR relation is used for precipitation estimation. In this model we use the cloud top temperature from the infrared channel and the corresponding rainfall data from MRMS data. In the first step we perform cloud segmentation using the Incremental Threshold Technique (ITT) to define cloud patches at five different temperature thresholds (220, 235, 253, 273, and 300 K). Then we extract geometric, coldness, and textural features from cloud patches and classify cloud patches to a number of clusters using a K-means clustering method. Each cluster contains cloud top brightness temperature and the corresponding rainfall rate from ground based radar. For each cluster we use a two stage probabilistic model based on the actual relationship of Tb-RR. The empirical probability of precipitation is defined to be the portion of positive rainfall values relative to portion of zero rainfall conditioned on brightness temperature. For positive precipitation rate values we fit a lognormal distribution conditioned to brightness temperature, from which we extract various quantiles. Figure 5.3 displays various conditional quantiles of precipitation rates and the probability of precipitation (POP) for each cloud

temperature for the cluster shown in Figure 5.1. Based on the analyses we have done for precipitation rate estimation, the best value for precipitation rate would be the expected value of precipitation multiply by probability of precipitation conditioned on the brightness temperature. As the Figure 5.3 on the right shows, the probability of rainfall is higher for colder clouds and it decreases by increasing cloud temperature. In this model we use two stages for precipitation retrieval. In the first stage the precipitation value is estimated as the expected value of the probabilistic model for different temperatures multiply by POP, and in the second stage false alarms are removed based on the probability of precipitation (POP). Equation 1 defines the model for estimating precipitation:

$$RR = \frac{\sum rr}{N} = \frac{\sum rr^{+}}{N} = \frac{\sum rr^{+}}{N} \times \frac{N^{+}}{N^{+}} = mean(+rr) \times POP \qquad (1)$$

Where RR is the final rainfall retrieved for each pixel based training the probabilistic model ; rr is the rainfall values for each brightness temperature; N stands for the total number of pixels either with rainfall or zero rainfall for each brightness temperature; N^{+} is the total number of pixels with rainfall value greater than zero for each brightness temperature; and POP is the probability of precipitation.

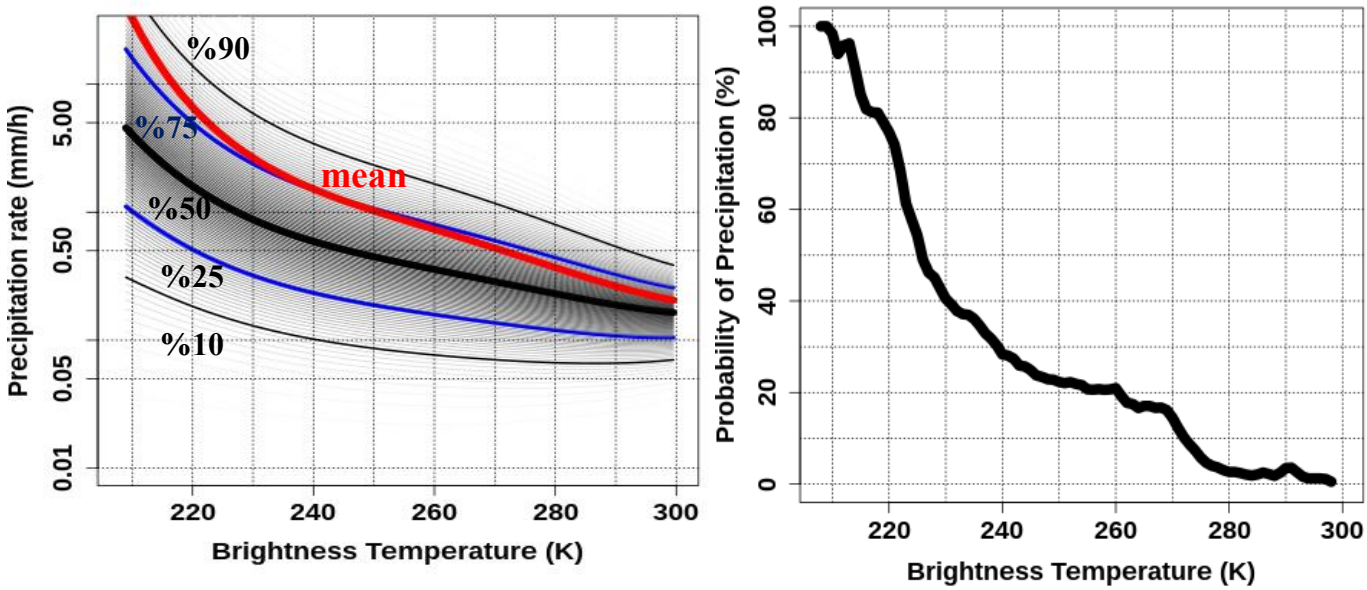


Figure 5.3: Probabilistic precipitation identification and estimation

5.4.Algorithm Training

The algorithm was trained over the CONUS. The data we used are infrared data from the GOES satellite and the corresponding MRMS rainfall data from June to September 2014. As discussed in the methodology, we used IR data for cloud segmentation and defined cloud patches. Then we extract different features from cloud patches to classify the patches into 200 groups using K-means clustering method. The result of the clustering method is 200 groups of IR data and corresponding rain rate values. For each cluster, a model of the probability of precipitation and a conditional lognormal distribution of the rain are fitted to the data. The conditional expected value of the distribution is then used for the precipitation retrieval for each cloud temperature. As displayed in Figure 5.3, while the probability of precipitation is lower for warmer clouds, the rainy pixels are very diverse ranging from cold clouds to warm clouds indicating that the portion of rainfall associated to warm clouds is not negotiable. The retrieval would then assign positive

although small precipitation rates everywhere. In order to remove the possible false alarms from the model, we define a threshold for the probability of precipitation based on maximizing the Heidke Skill Score (HSS) relative to MRMS for all the time steps in training data. Figure 5.4 demonstrates the process for removing false alarm based on training data. During the training period for each data sample we find the probability of precipitation threshold that maximizes the HSS. In the retrieval, the threshold is used to replace by zero value the rainy pixels with probability of precipitation less than the threshold. Figure 5.4 (bottom) shows the histogram of the probability of precipitation thresholds corresponding to maximum HSS values for all the training data. As we can see in the histogram the probability of precipitation corresponding to maximum HSS for most of the samples ranges between 30% and 50%. The threshold that we set for running this algorithm is 35% based on the histogram. In the retrieval, this threshold is used to replace by zero value the rainy pixels with probability of precipitation less than threshold.

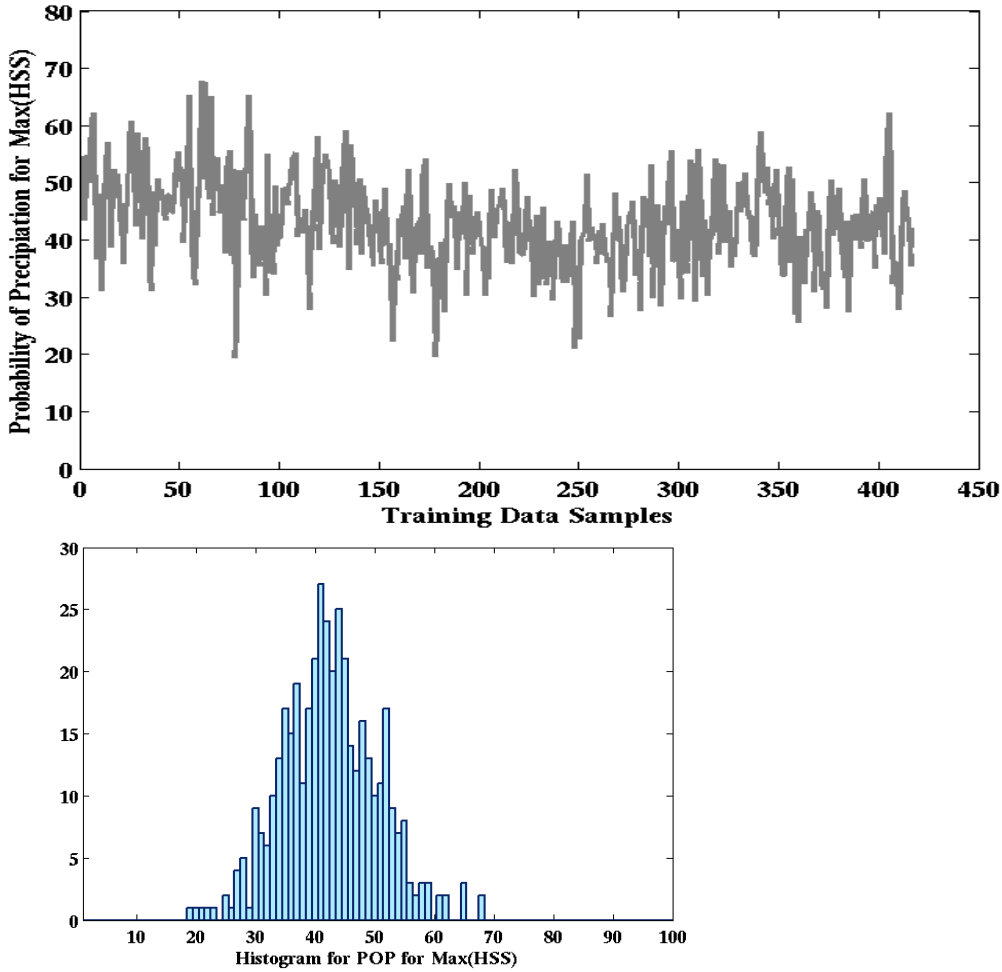


Figure 5.4: Calculating POP values for maximum HSS during training period for identifying the POP threshold

5.5.Results and discussions

To verify the model we chose an additional year not covered in the training period. The time between June to September 2015 was chosen as a validation period and the precipitation retrieved from the new model was evaluated against PERSIANN-CCS considering MRMS data as a ground truth over the CONUS during validation and calibration period to make sure the results are consistent. Figure 5.5 displays the time series of binary analysis for training data (see appendix A for the equations). To enhance the visualization of the time series analysis, the difference between binary parameters calculated for probabilistic model and PERSIANN-CCS

are showed in Figure 5.7 (values for Probabilistic model subtracted from PERSIANN-CCS parameter values for all the parameters). According to Figure 5.7 Probability Of Detection (POD) difference is positive in all time steps meaning that the POD is higher in probabilistic model than PERSIANN-CCS in all the samples. The difference amounts for False Alarm Ratio are positive for majority of time steps but the difference rate is lower than POD difference rate meaning that the FAR does not increase as much as POD increases. Also we have calculated the Heidke Skill Score (HSS) that is a measurement for the model performance in terms of POD and FAR. According to Figure 5.7 the HSS difference amount is positive for almost all the time steps indicating that the total performance of the probabilistic model in terms of rain detection is better than PERSIANN-CCS. Also, the model was evaluated based on calculating relative bias, correlation coefficient, and RMSE during training period (see appendix A for the equations). Figure 5.6 shows the relative bias, correlation coefficient, and RMSE for each time step during training period. In Figure 5.8 we have plotted the difference amounts for the parameters calculated in Figure 5.6. As the Figure 5.8 displays, the relative bias difference is negative for majority of time steps meaning that the relative bias value is lower for probabilistic model. Also the correlation coefficient difference is positive cases and RMSE is negative for almost all the time steps indicating the probabilistic model outperforms PERSIANN-CCS algorithm. In addition to time series analysis, we evaluated the precipitation measured from the new model based on the monthly total rainfall during validation and calibration period. Image 5.9 shows the total accumulated rainfall during training period estimated from Probabilistic model, PERSIANN-CCS, and MRMS data. The figure shows that the new model can capture some portion of rainfall that PERSIANN-CCS is not capturing over the North and South East of the US. As the results show in Table 5.1, the statistical parameters are calculated to evaluate the total

performance of the probabilistic model and compare it with PERSIANN-CCS against MRMS data during training period. As is displayed in the table, the statistical parameters are improved for the probabilistic model comparing to PERSIANN-CCS. The average relative bias, average correlation coefficient, and average RMSE are improved by 67%, 36%, and 18% respectively. In the next section we evaluate the model further to verify that the results during the training and testing periods are consistent.

Table 5-1: Statistical evaluation of probabilistic model and PERSIANN-CCS using MRMS data during training period. Numbers are calculated based on averaging over entire month.

	CCS			Probabilistic model		
	Relative Bias	Corr.	RMSE (mm/hr)	Relative Bias	Corr.	RMSE (mm/hr)
Training data	-0.29	0.25	1.48	-0.1	0.34	1.22

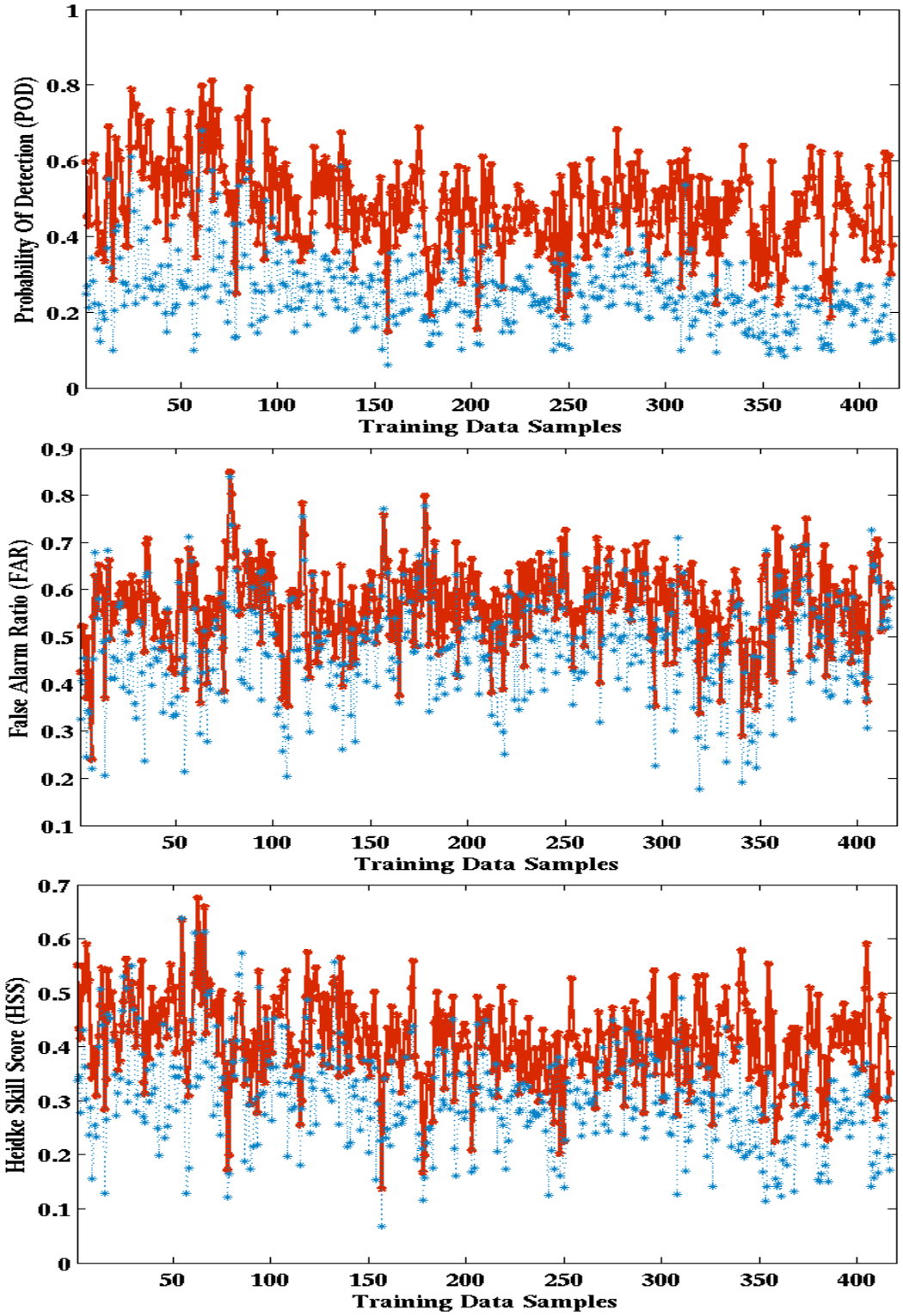


Figure 5.5: Time series of binary analysis of probabilistic model and PERSIANN-CCS during training period (June-September 2014). Blue dashed line represent values for PERSIANN-CCS and red solid line represents values for probabilistic model.

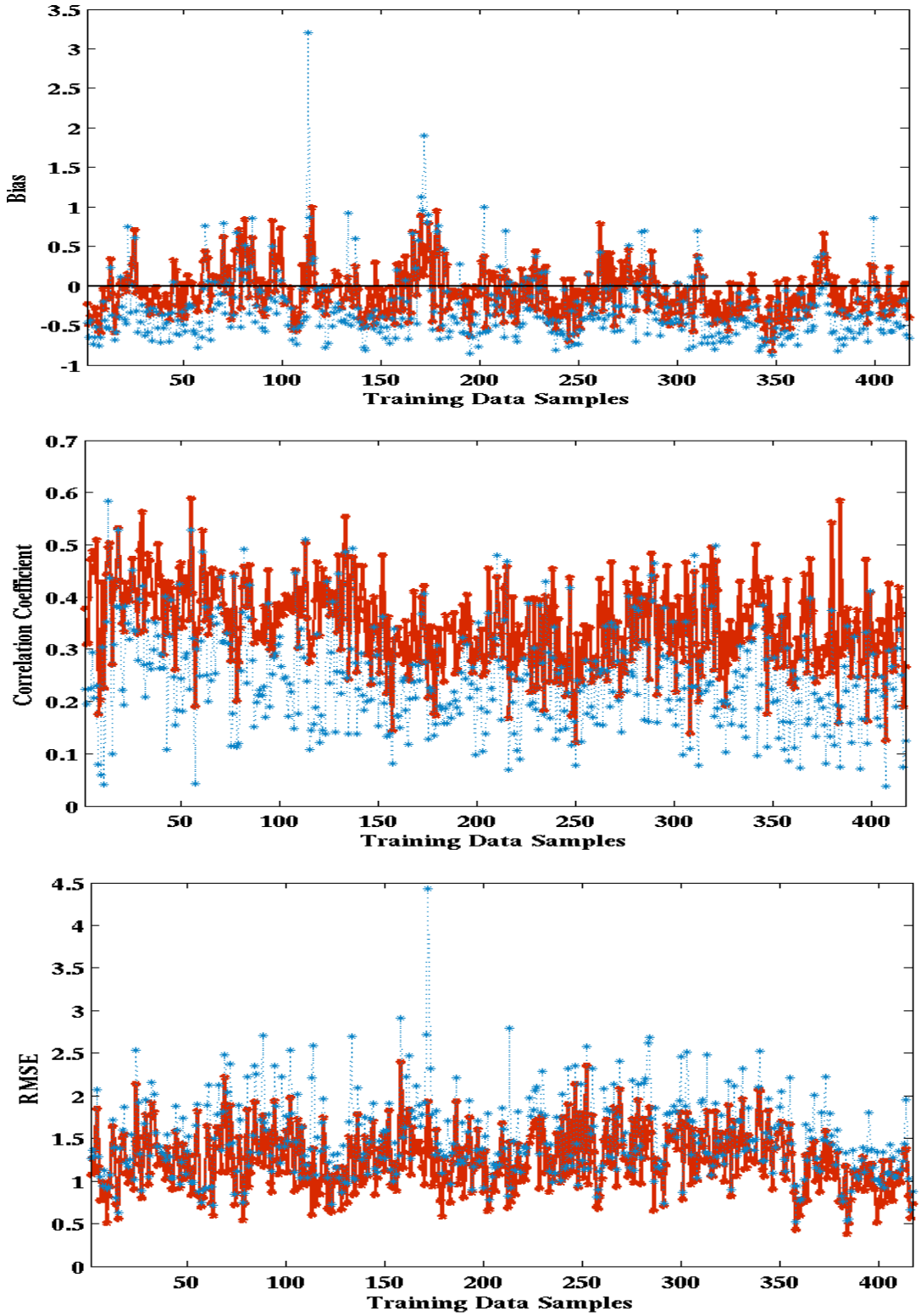


Figure 5.6: Time series of relative bias, correlation coefficient, and RMSE of Probabilistic model and PERSIANN-CCS during training period (June to September 2014). Blue dashed line represent values for PERSIANN-CCS and red solid line represents values for Probabilistic model.

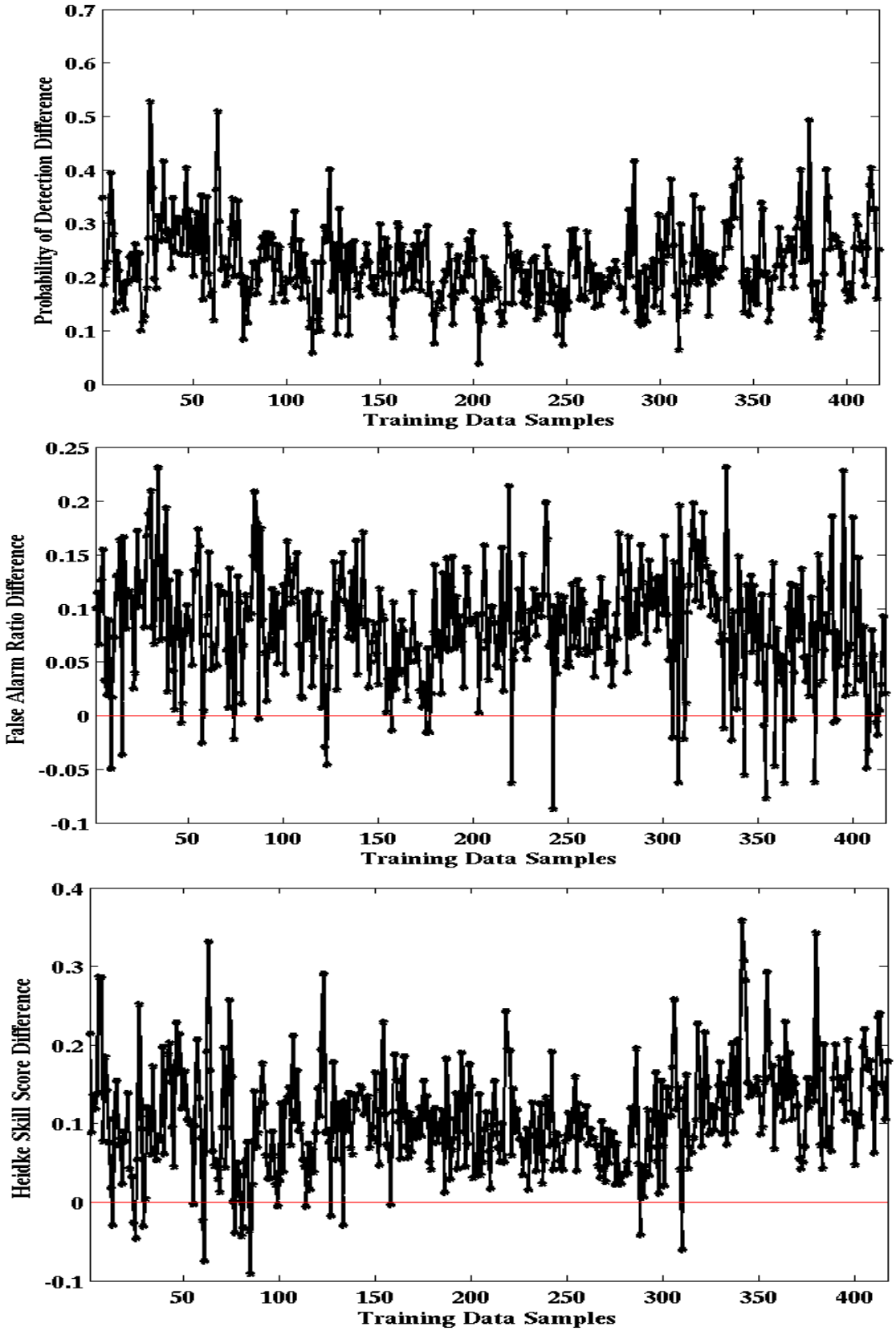


Figure 5.7: Difference between POD, FAR, and HSS values of probabilistic model and PERSIANN-CCS during training period (June-September 2014).

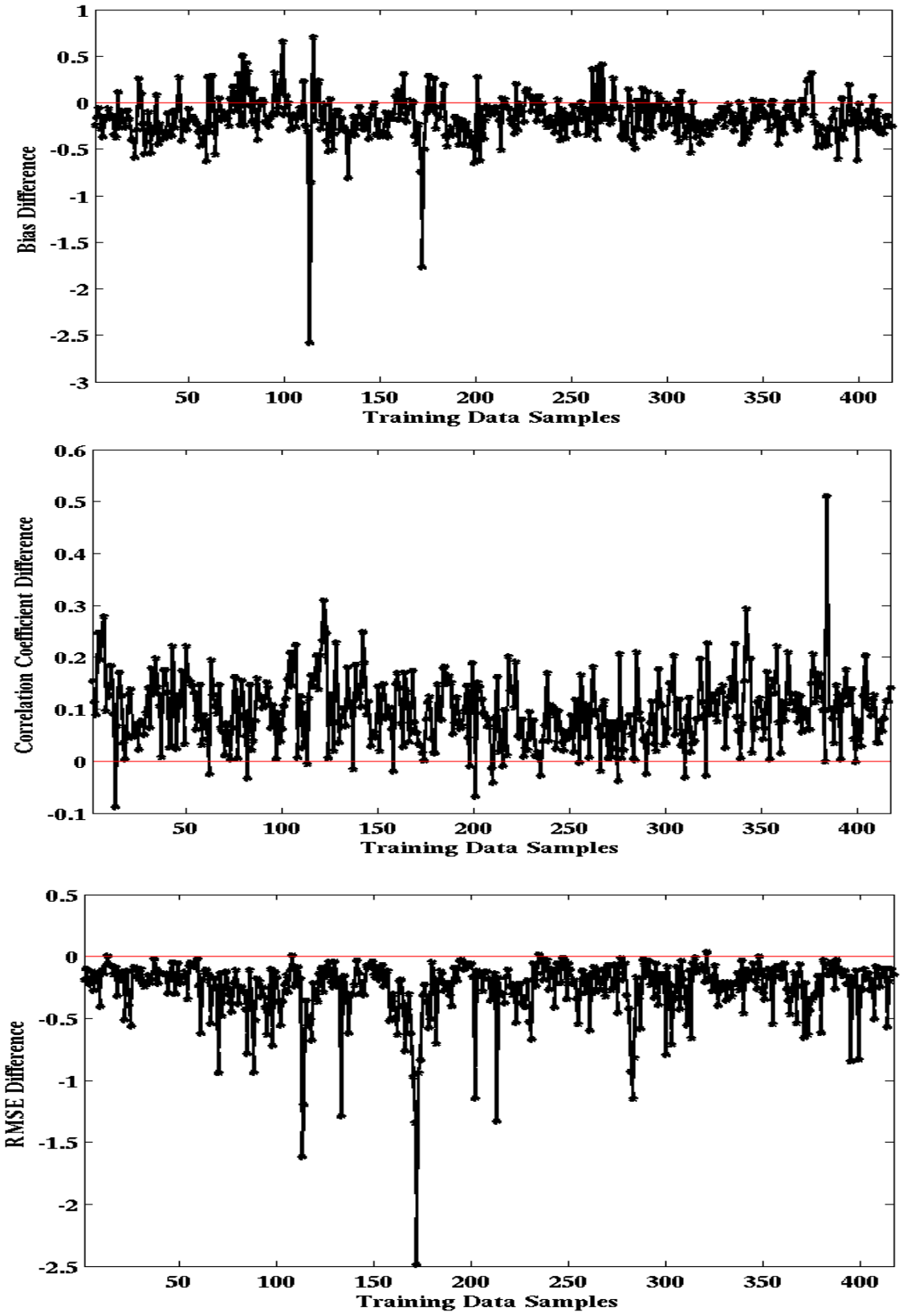


Figure 5.8: Difference between relative bias, correlation coefficient, and RMSE values of probabilistic model and PERSIANN-CCS during training period (June-September 2014).

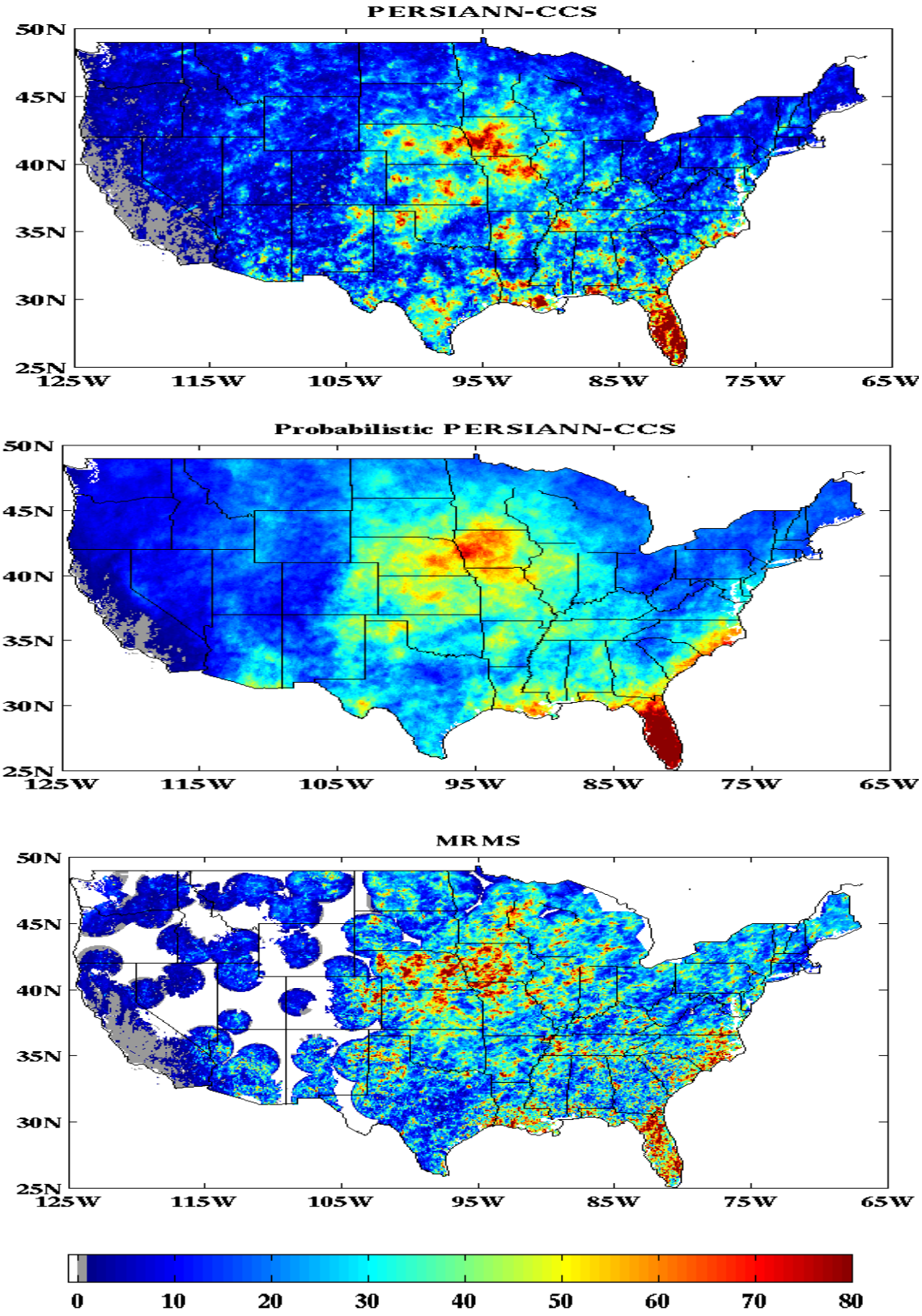


Figure 5.9: Total accumulated rainfall during training period (mm)

In addition to the training period, the model is evaluated during July, August, and September 2015 over the CONUS as well. Figures 5.10, 5.11, and 5.12 display the time series of binary analysis during July, August and September 2015 and Figures 5.16, 5.17, and 5.18 show the difference between binary parameters calculated for probabilistic model and PERSIANN-CCS algorithm. As it shown in the figures the POD values are better for all the time steps during validation period since the difference amount is positive. The FAR difference rate is also positive for most of the cases indicating that the FAR values are greater for probabilistic model however the FAR does not increase as much as POD increases. The performance of the model in detecting rainfall (in terms of true and false detection) can be measured by HSS. As is shown in Figures 5.16, 5.17, and 5.18, the HSS is improved for almost all time steps ($[HSS_{probabilistic} - HSS_{PERSIANN\text{-}CCS}] > 0$), which indicates that the ability of the probabilistic model at detecting rainfall in terms of true and false detection is improved. In addition to binary analysis, we evaluated the performance of the model by calculating the statistical parameters such as average relative bias, average correlation coefficient, and average RMSE. Figures 5.13, 5.14, and 5.15 show, the average relative bias, average correlation coefficient, and average RMSE calculated for time steps during July, August, and September 2015. According to Figure 5.19, 5.20, and 5.21 the statistical parameters are improved for probabilistic model since the difference between relative bias, correlation coefficient and RMSE indicate the better performance of the probabilistic model. Also, we validated the model in the monthly scale over the CONUS during July, August, and September 2015. Figure 5.22 through 5.24 demonstrate the total monthly rainfall during validation period. The statistical parameters calculated to evaluate the probabilistic model are shown in table 5.2.

During July 2015 the average of relative bias did not change significantly, but the average correlation coefficient and average RMSE are improved. Also, during August and September 2015 the average relative bias, average correlation coefficient, and average RMSE are improved, showing that the probabilistic model performs better than PERSIANN-CCS for the validation period.

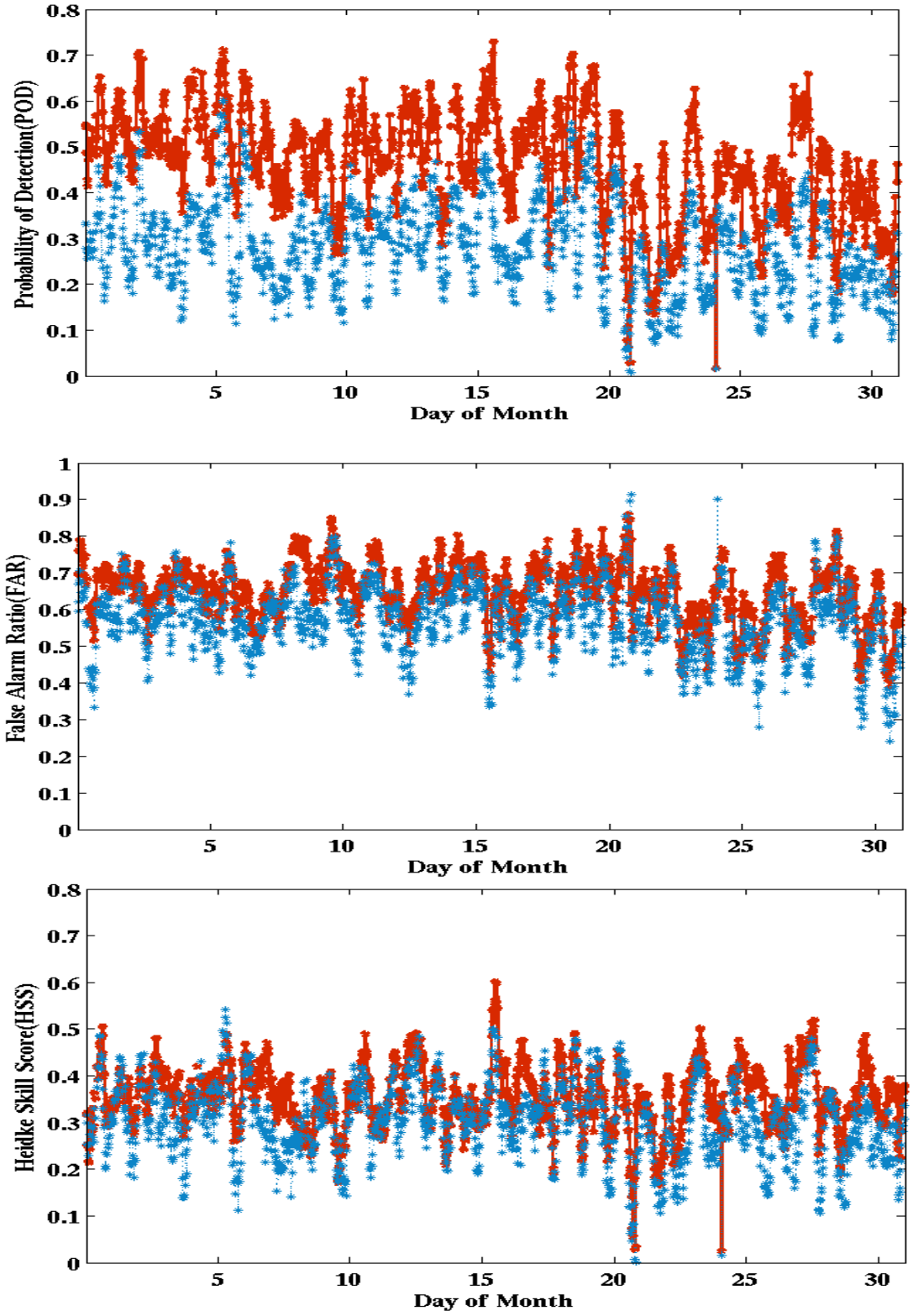


Figure 5.10: Time series of binary analysis of probabilistic model and PERSIANN-CCS model during validation period (July 2015). Blue dashed line represents values for PERSIANN-CCS and red solid line represents values for probabilistic model.

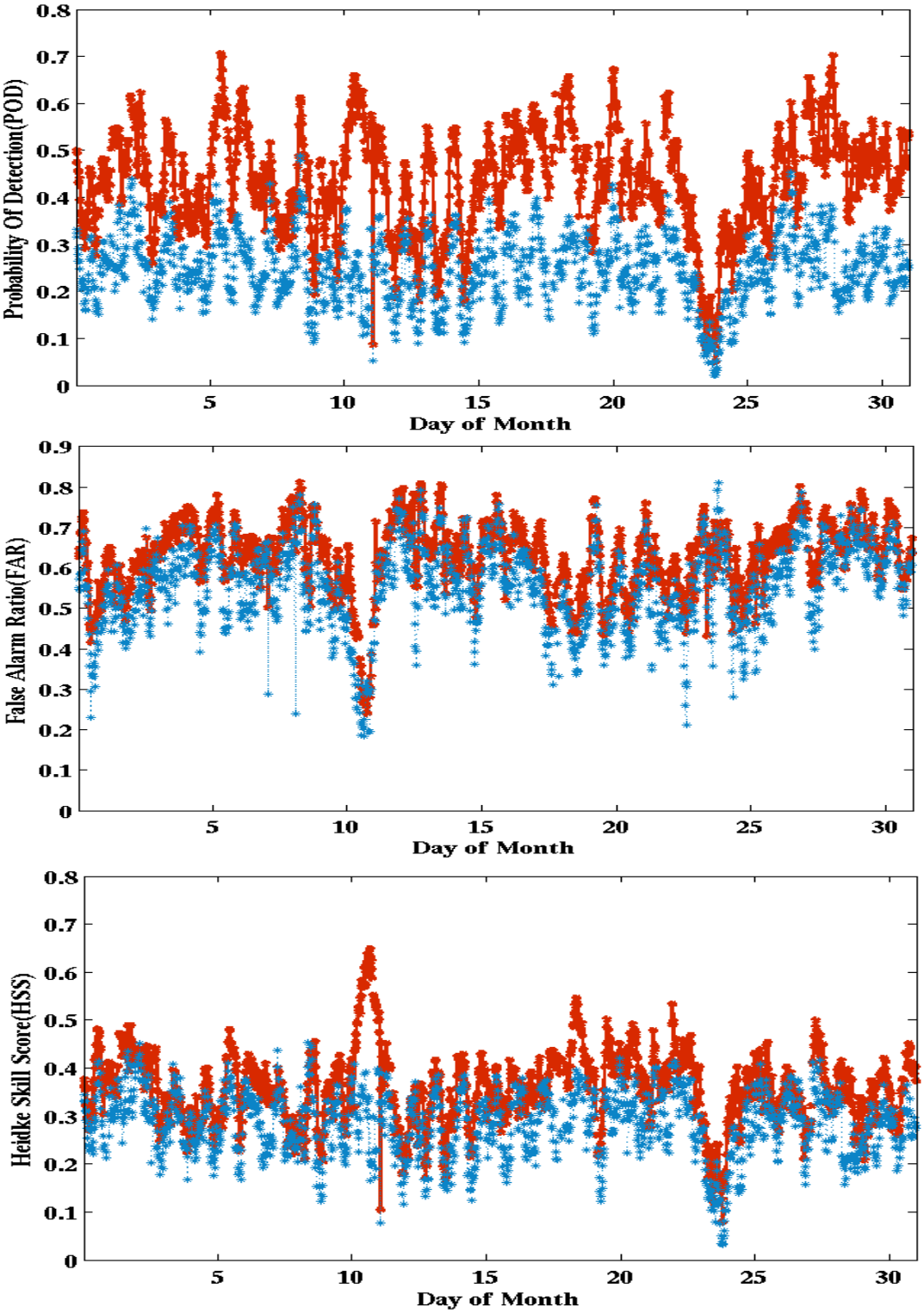


Figure 5.11: Time series of binary analysis of probabilistic model and PERSIANN-CCS during training period (August 2015). Blue dashed line represents values for PERSIANN-CCS and red solid line represents values for probabilistic model.

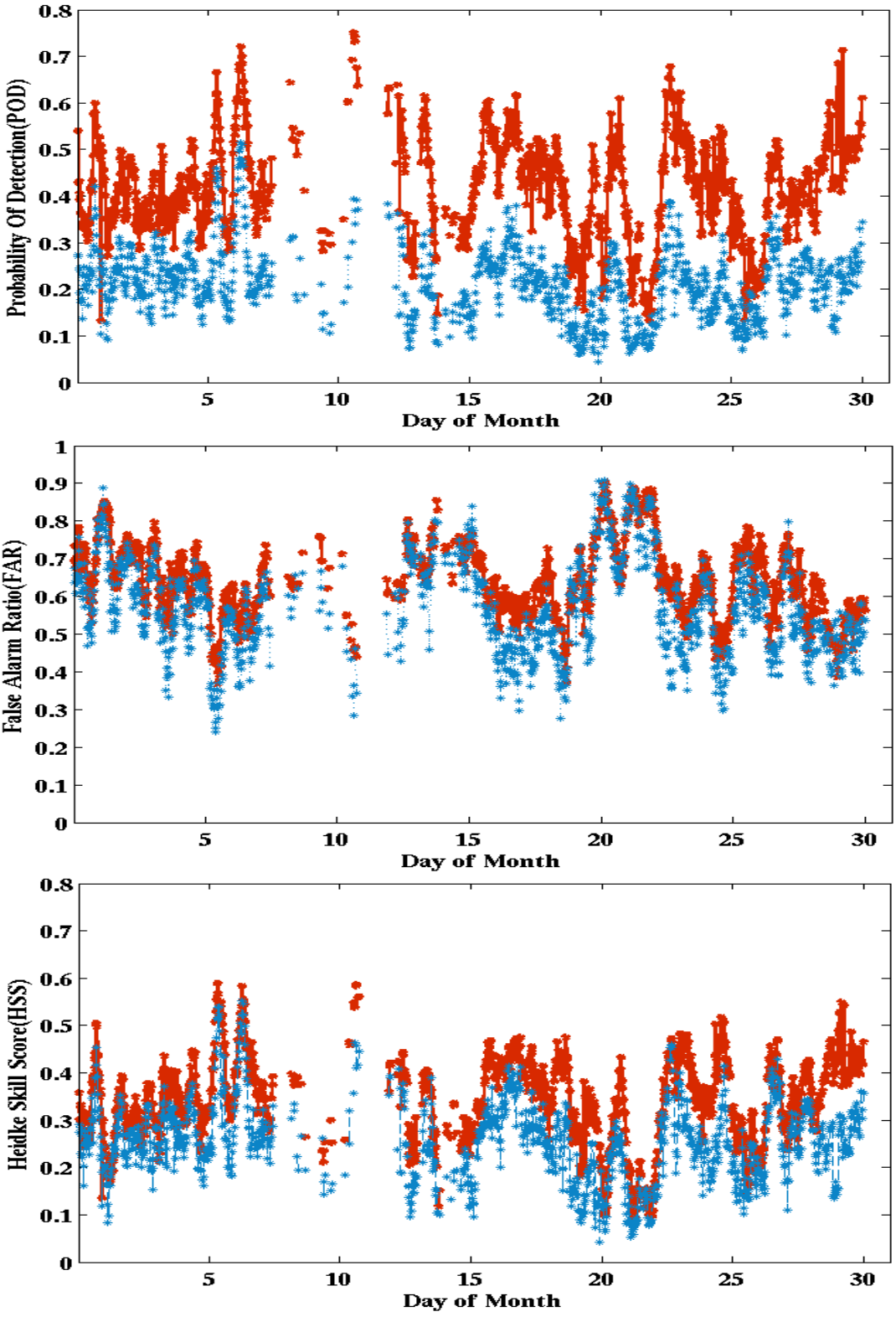


Figure 5.12: Time series of binary analysis of probabilistic model and PERSIANN-CCS model during validation period (September 2015). Blue dashed line represents values for PERSIANN-CCS and red solid line represents values for probabilistic model.

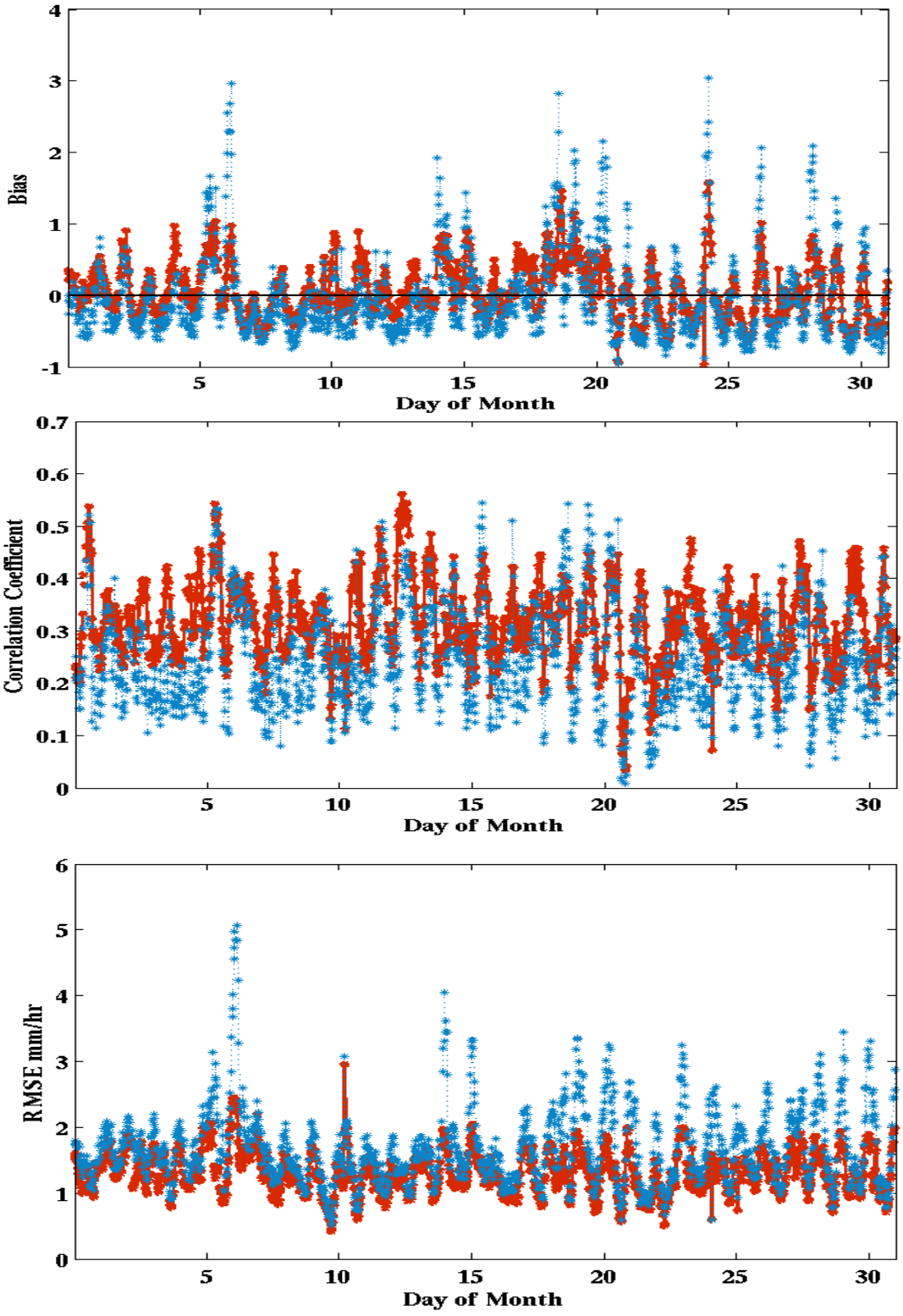


Figure 5.13: Time series of relative bias, correlation coefficient, and RMSE of Probabilistic model and PERSIANN-CCS during validation period (July 2015). Blue dashed line represents values for PERSIANN-CCS and red solid line represents values for Probabilistic model.

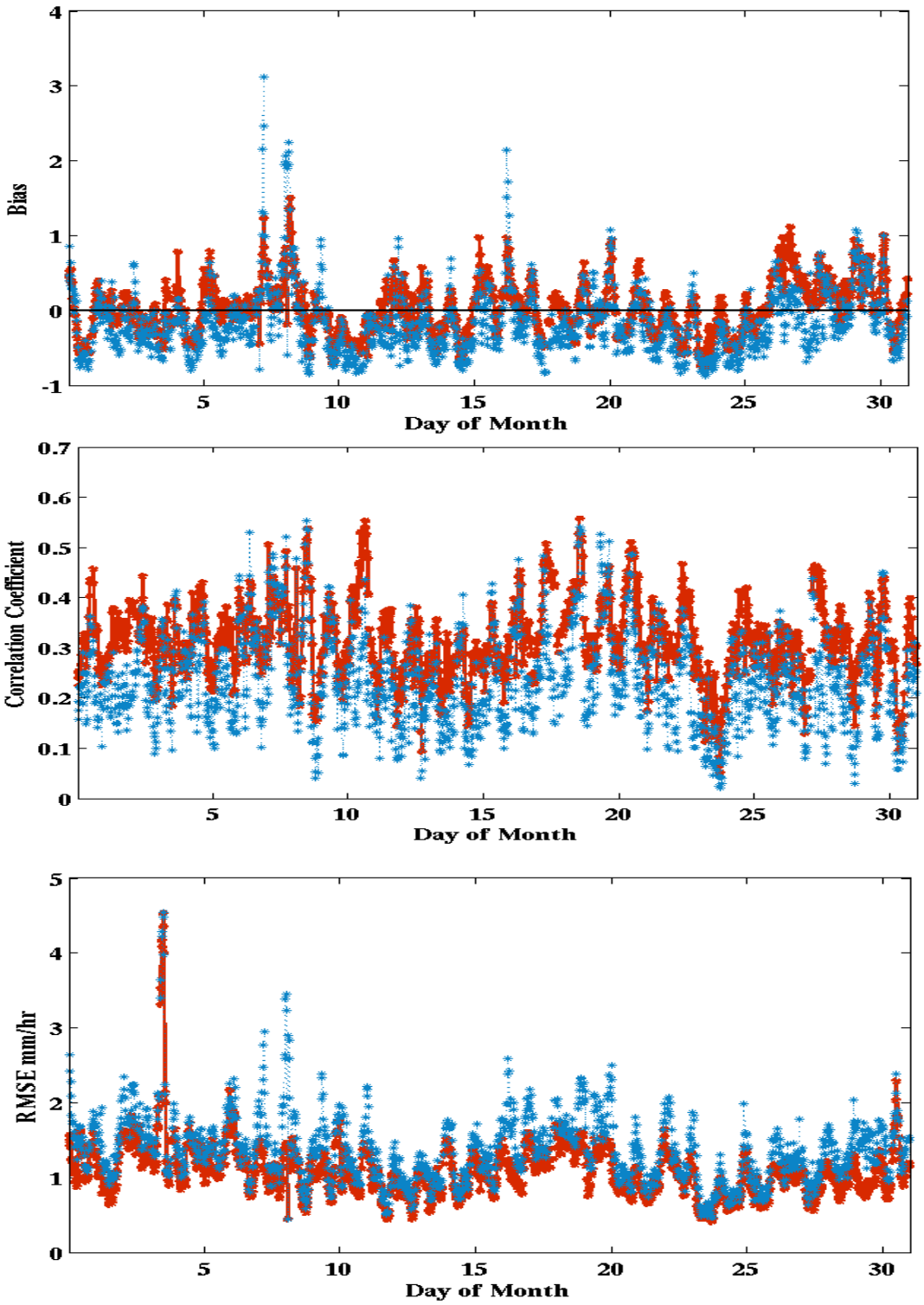


Figure 5.14: Time series of relative bias, correlation coefficient, and RMSE of Probabilistic model and PERSIANN-CCS during validation period (August 2015). Blue dashed line represents values for PERSIANN-CCS and red solid line represents values for Probabilistic model.

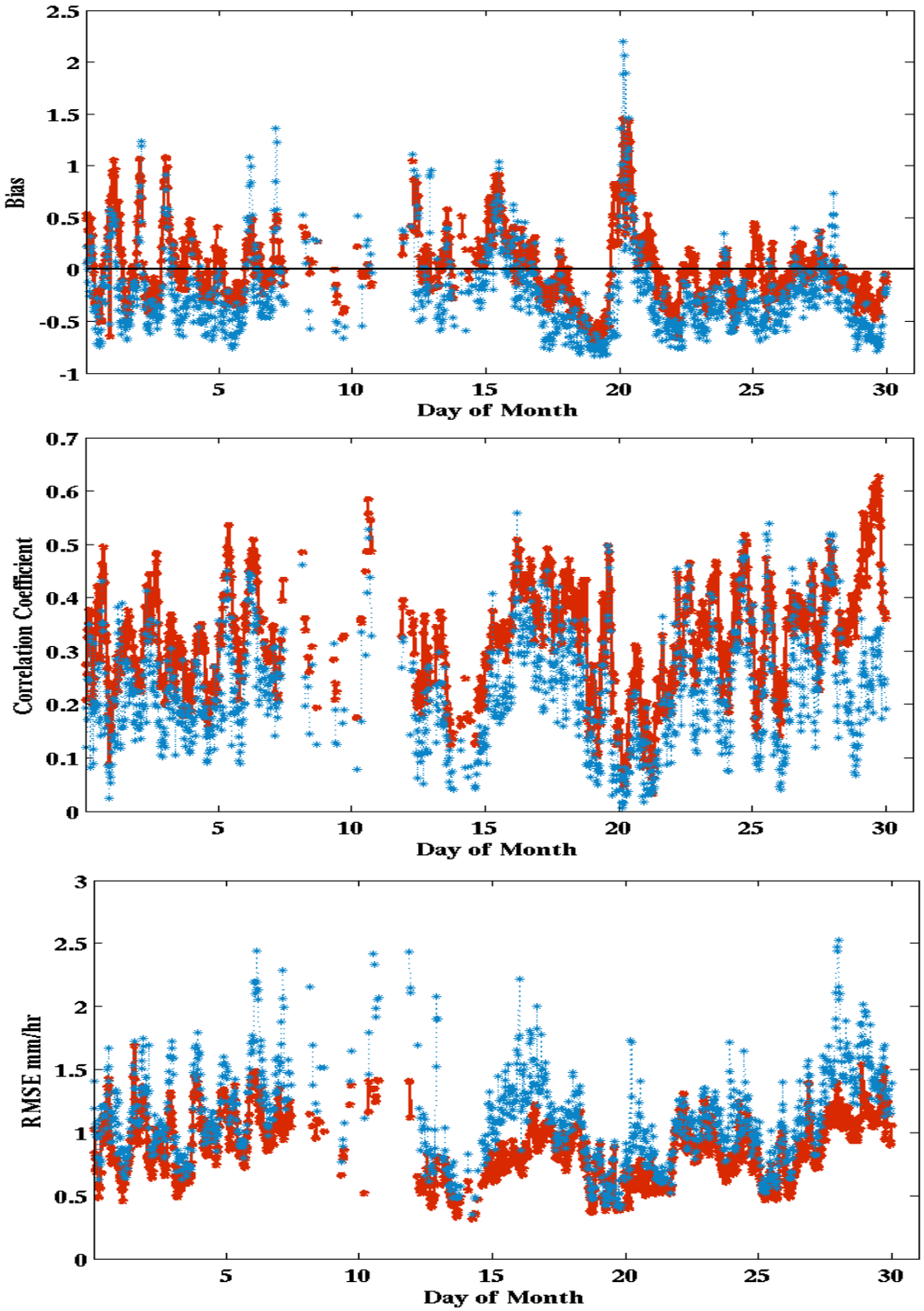


Figure 5.15: Time series of relative bias, correlation coefficient, and RMSE of Probabilistic model and PERSIANN-CCS during validation period (September 2015). Blue dashed line represents values for PERSIANN-CCS and red solid line represents values for Probabilistic model.

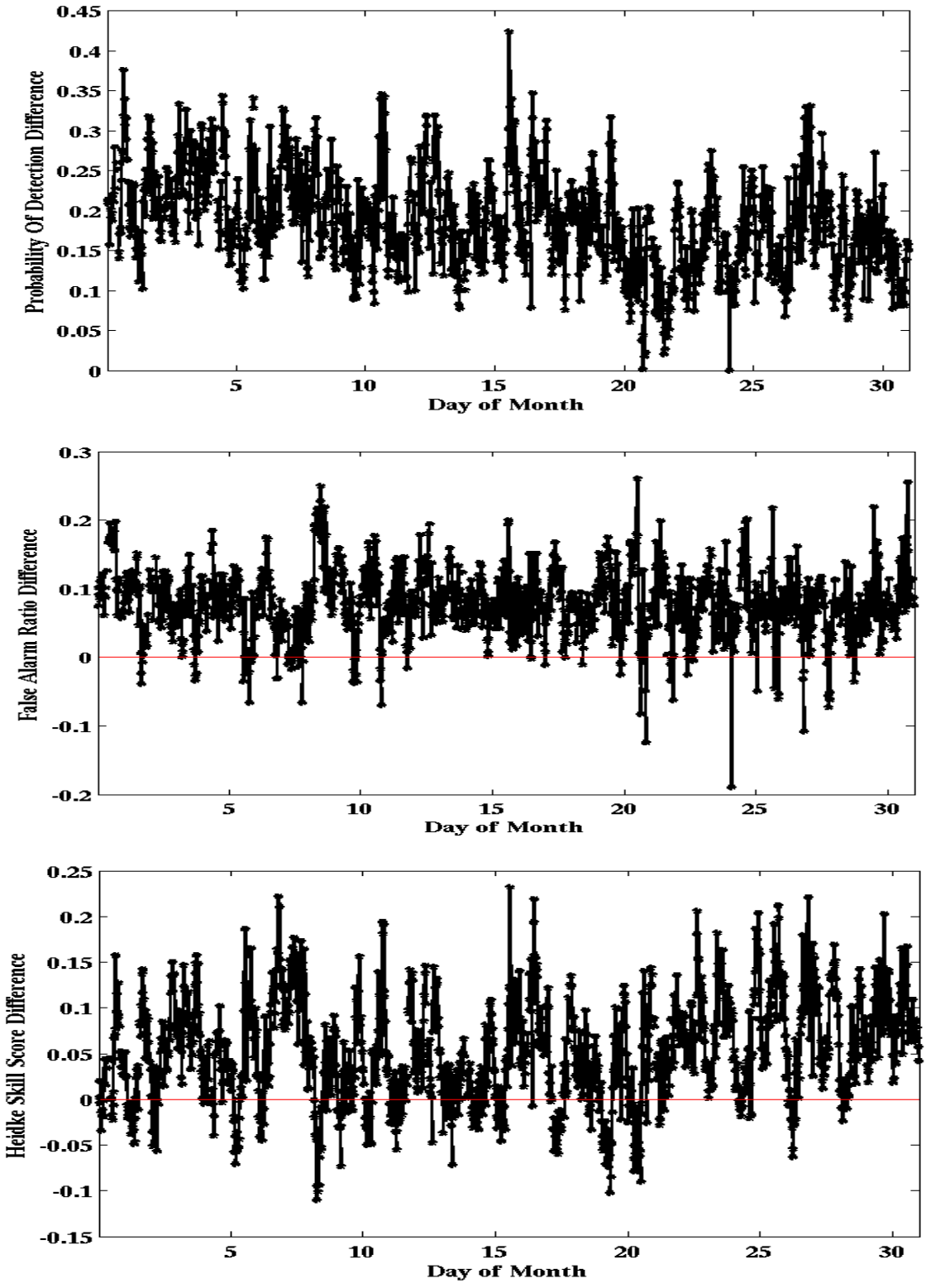


Figure 5.16: Difference between POD, FAR, and HSS values of probabilistic model and PERSIANN-CCS during validation period (July 2015).

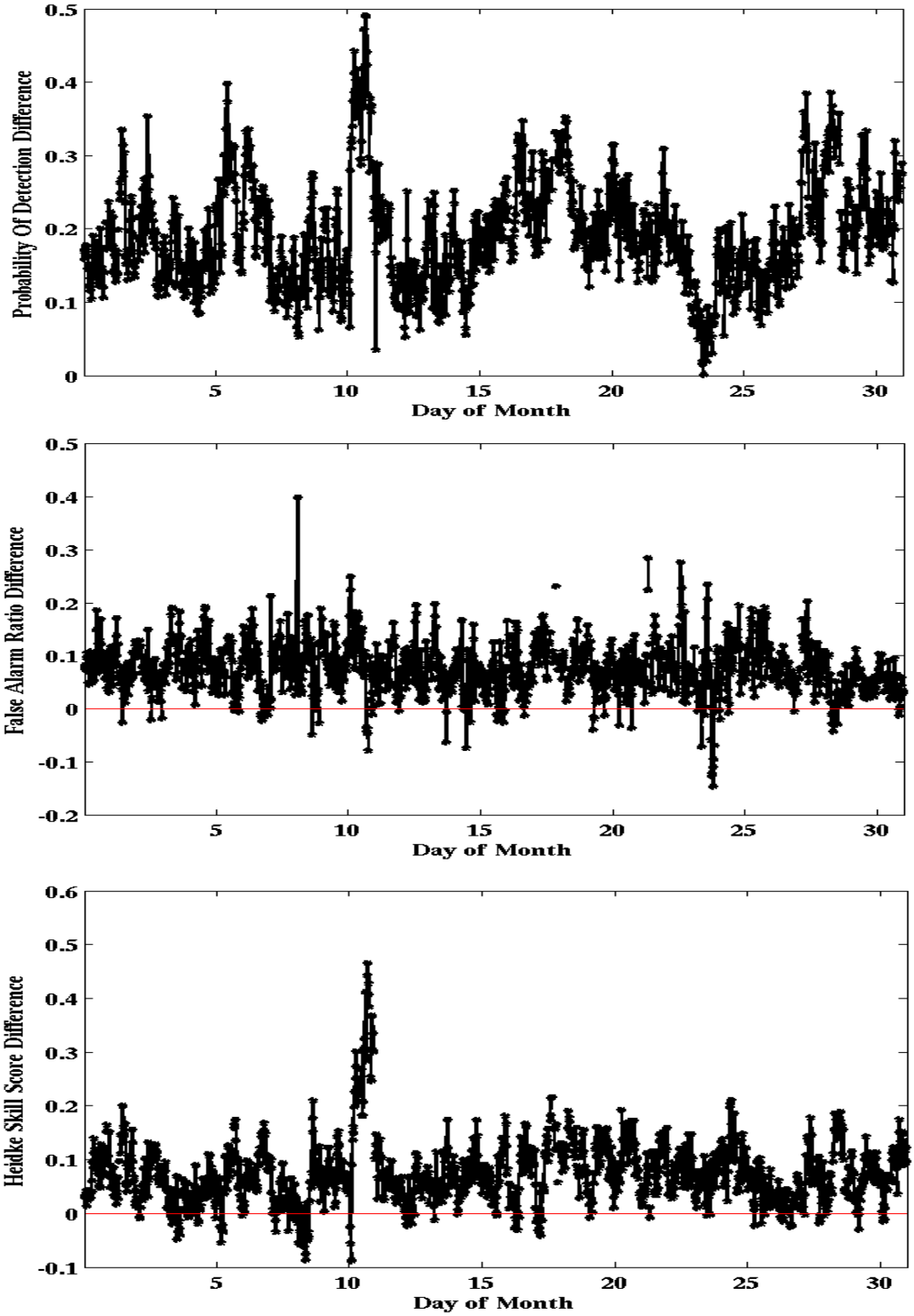


Figure 5.17: Difference between POD, FAR, and HSS values of probabilistic model and PERSIANN-CCS during validation period (August 2015).

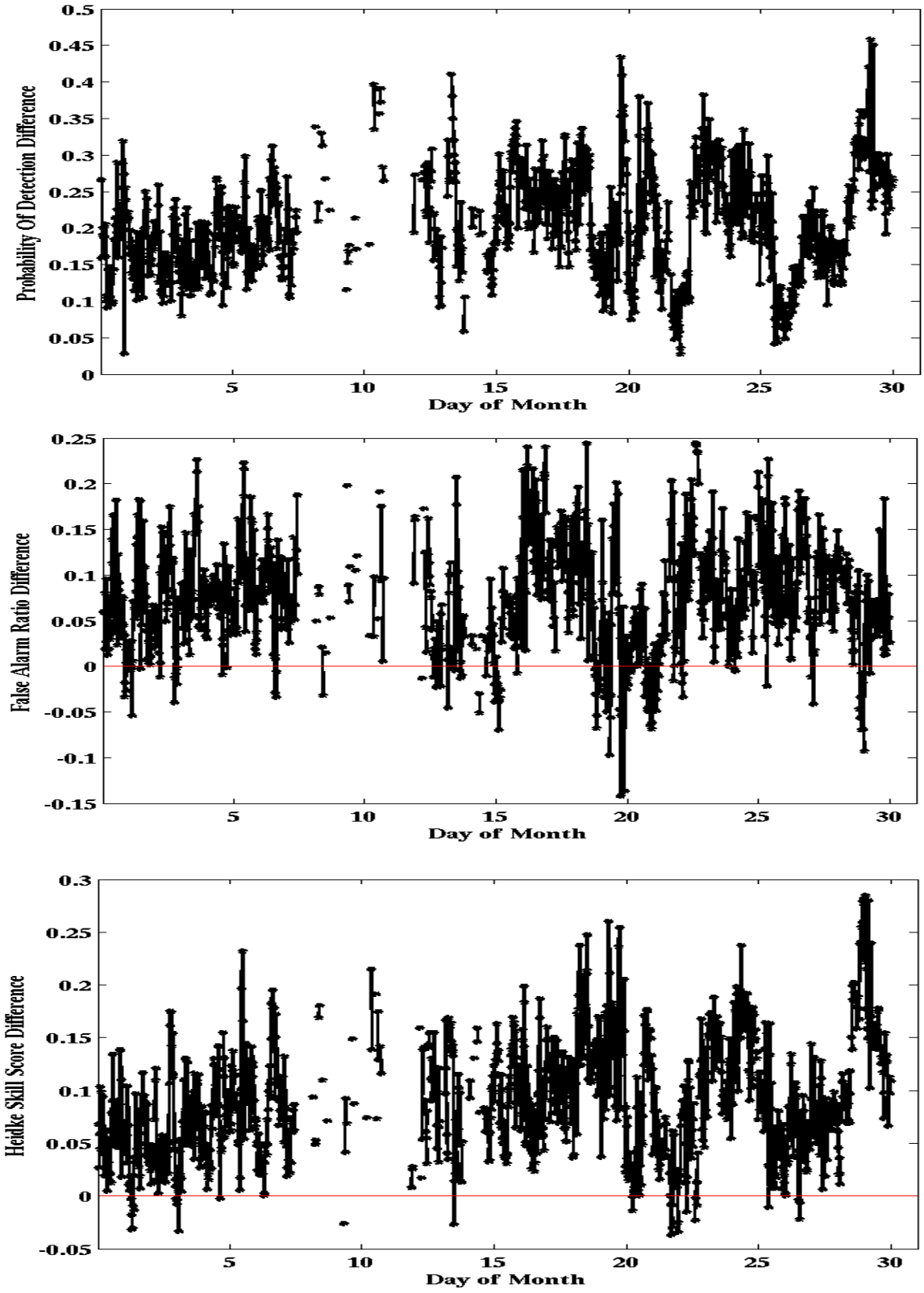


Figure 5.18: Difference between POD, FAR, and HSS values of probabilistic model and PERSIANN-CCS during validation period (September 2015).

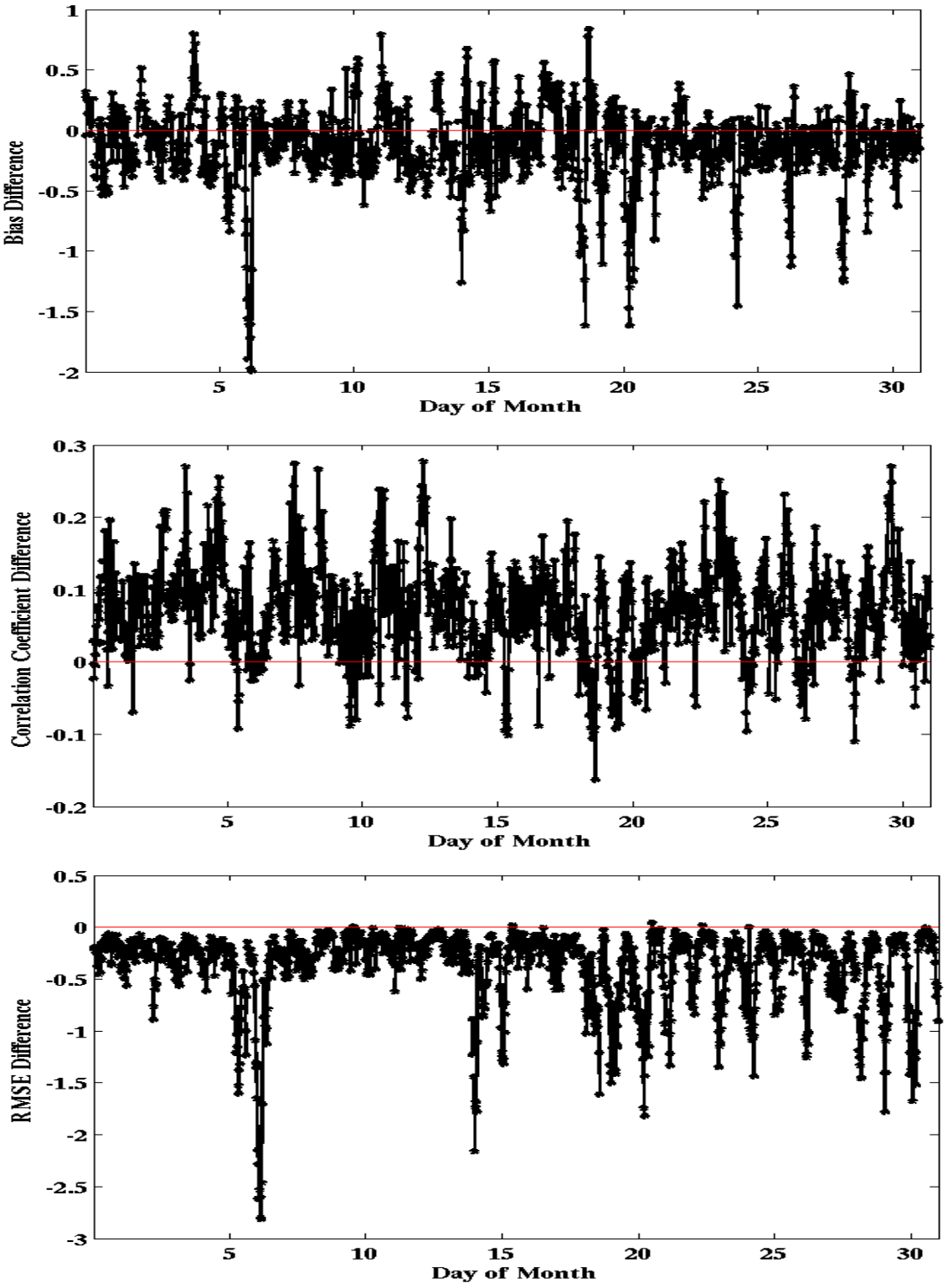


Figure 5.19: Difference between relative bias, correlation coefficient, and RMSE values of probabilistic model and PERSIANN-CCS during validation period (July 2015).

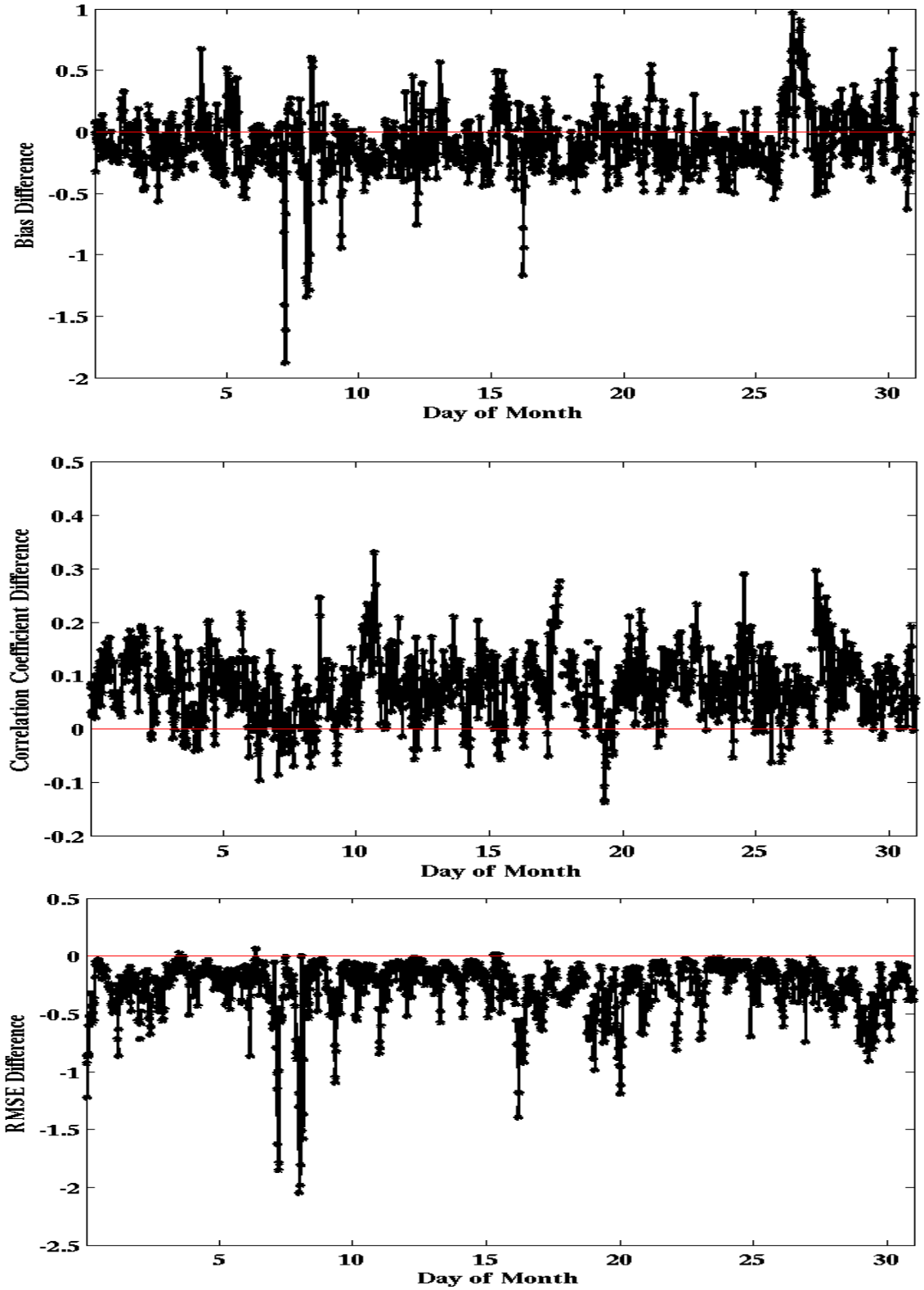


Figure 5.20: Difference between relative bias, correlation coefficient, and RMSE values of probabilistic model and PERSIANN-CCS during validation period (August 2015).

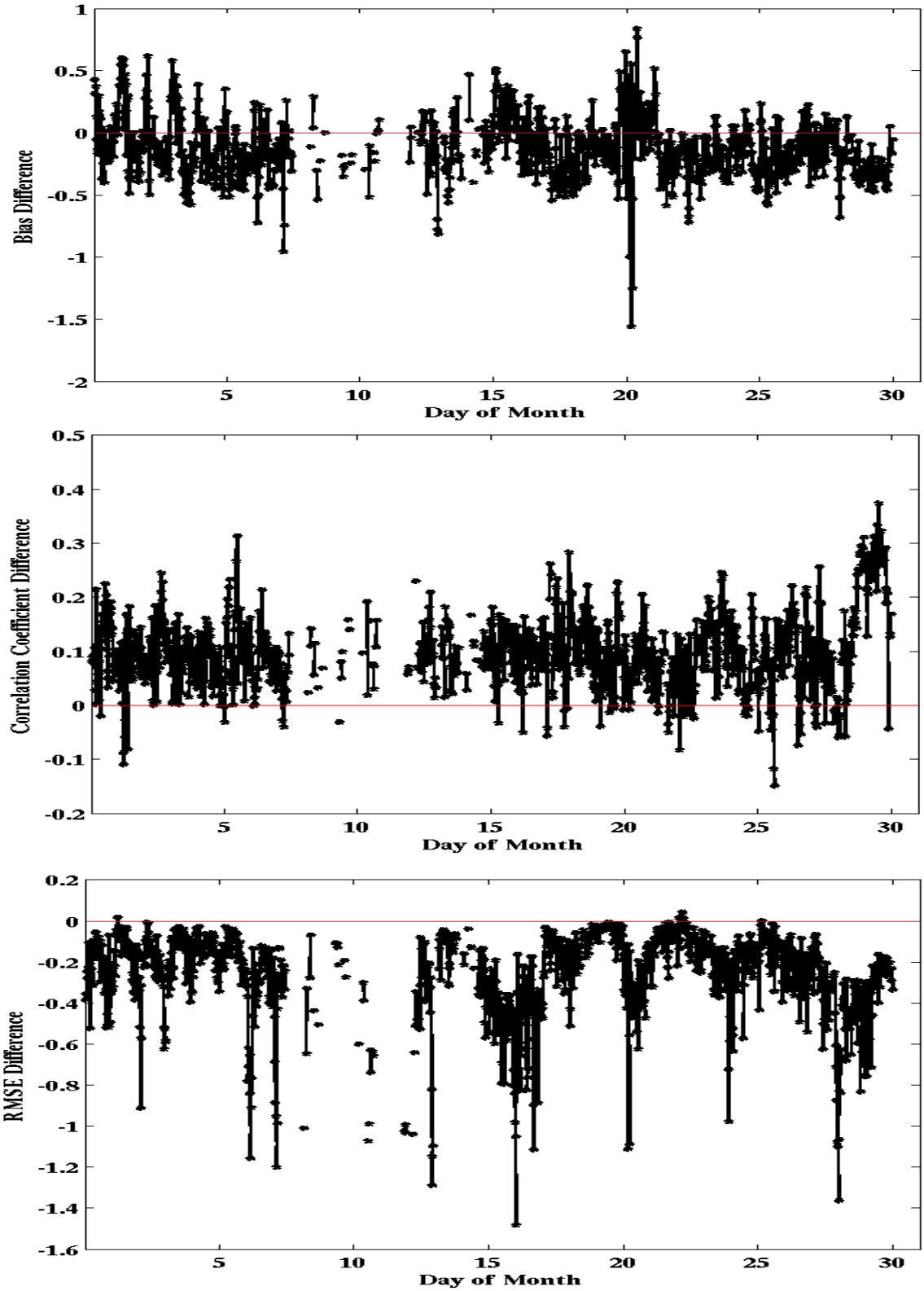


Figure 5.21: Difference between relative bias, correlation coefficient, and RMSE values of probabilistic model and PERSIANN-CCS during validation period (September 2015).

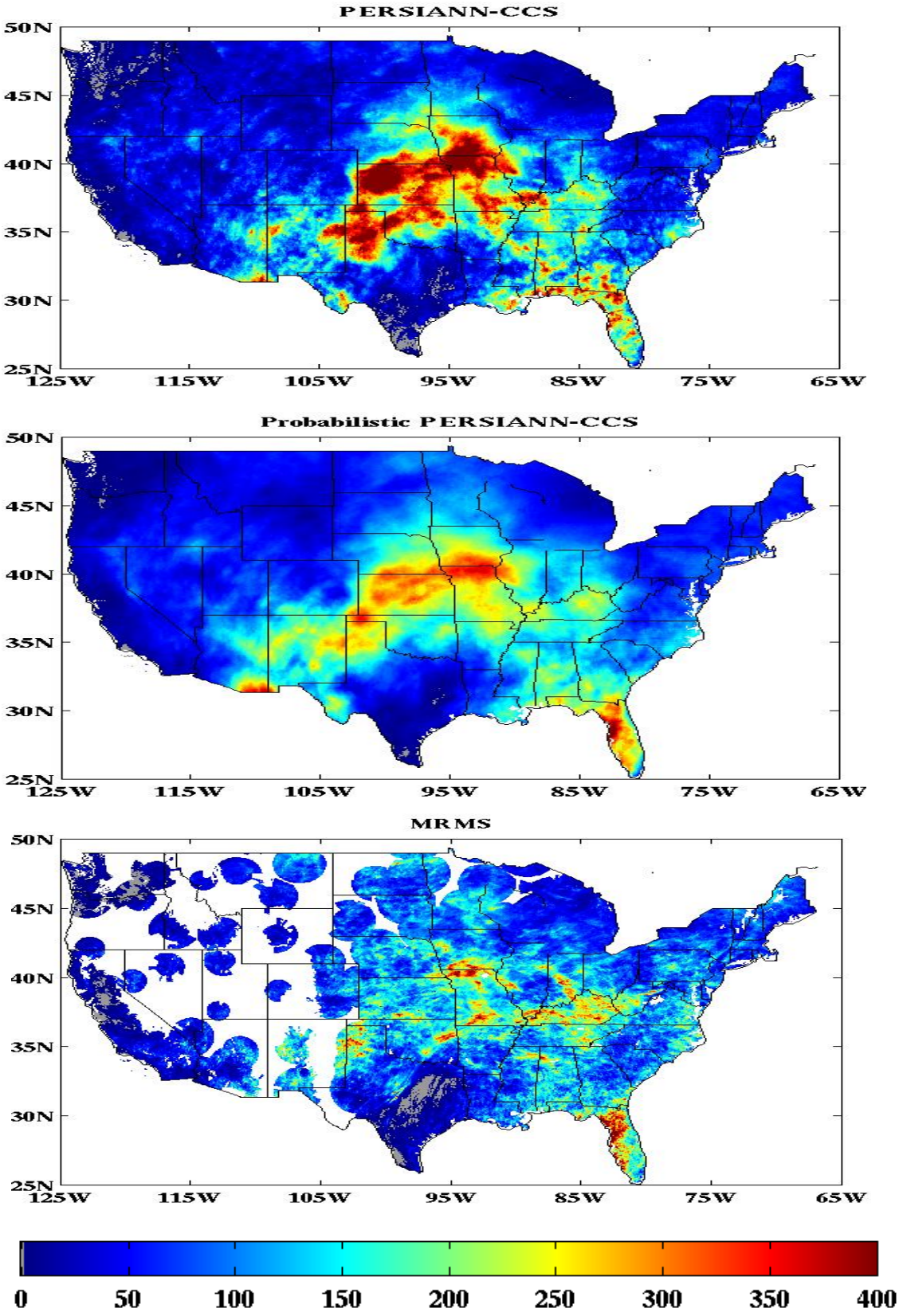


Figure 5.22: Total accumulated rainfall during validation period (July 2015-mm)

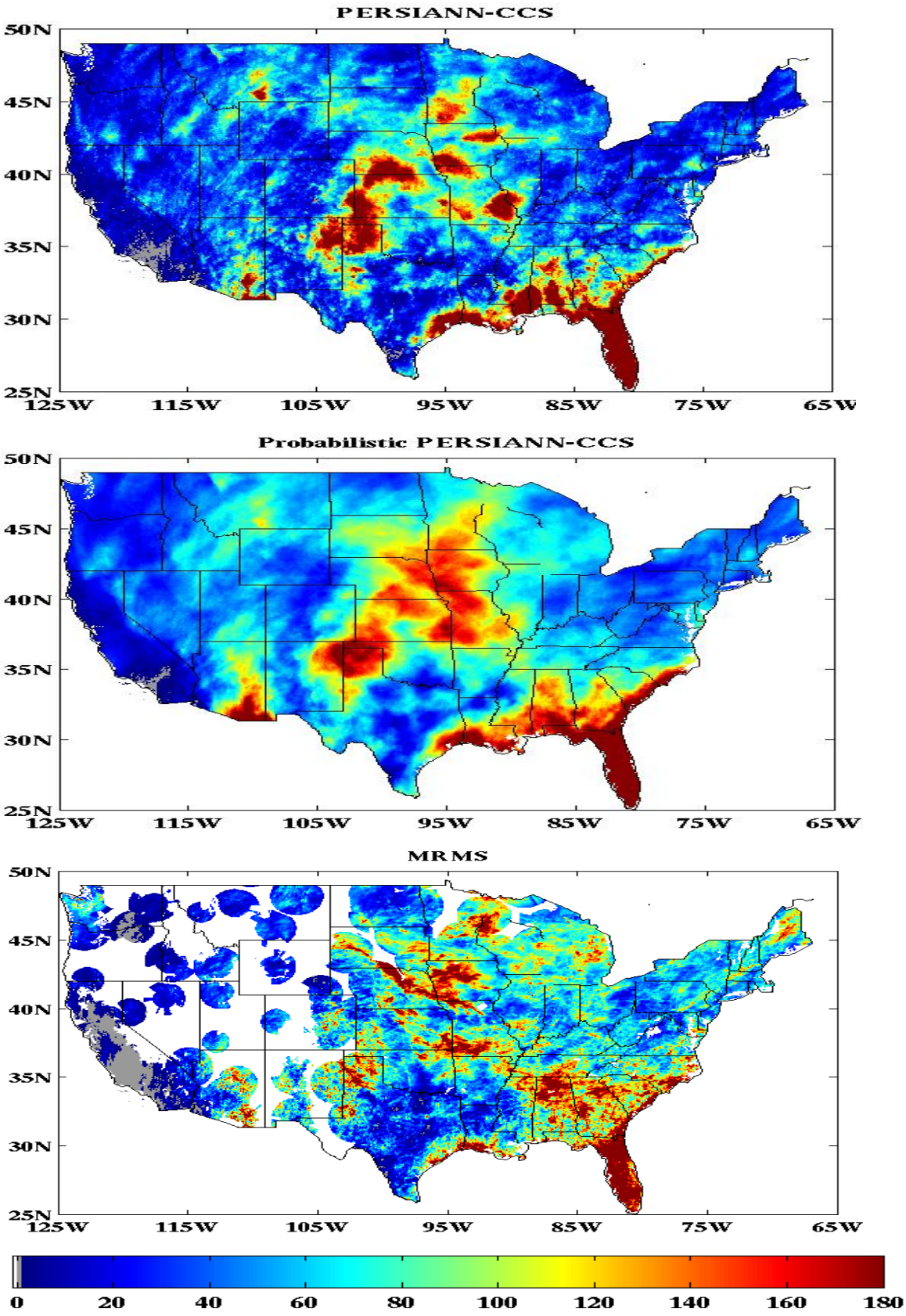


Figure 5.23: Total accumulated rainfall during validation period (August 2015-mm)

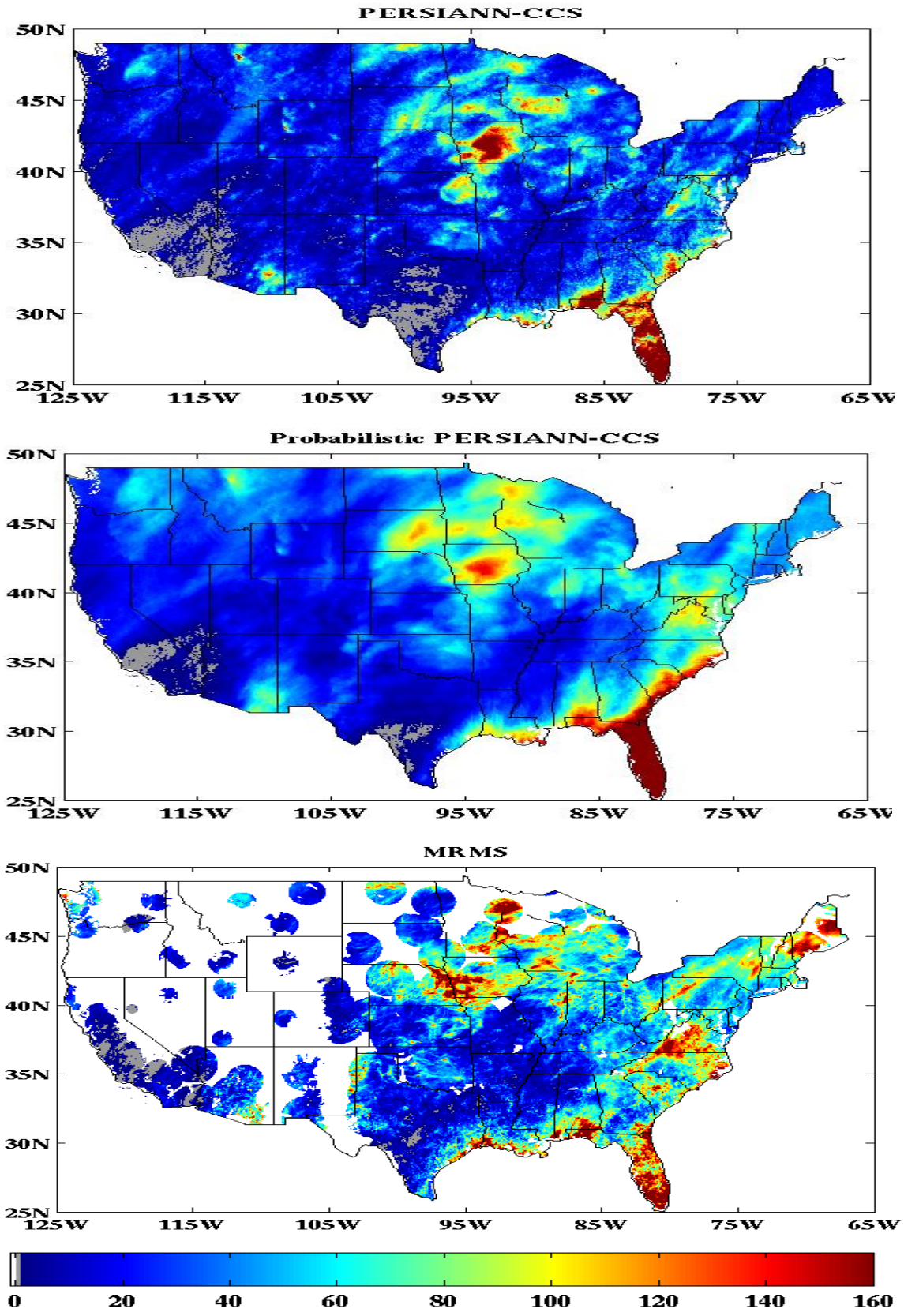
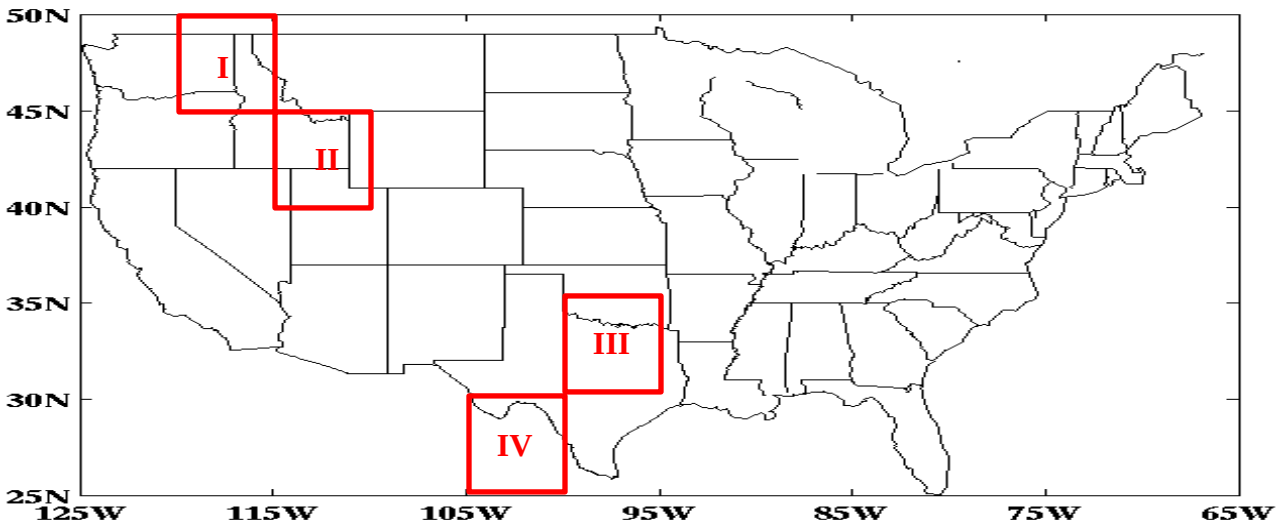


Figure 5.24: Total accumulated rainfall during validation period (September 2015-mm)

Table 5-2: Statistical evaluation of probabilistic model and PERSIANN-CCS using MRMS data during validation period. Numbers are calculated based on averaging over entire month.

	CCS			Probabilistic model		
	Relative Bias	Corr.	RMSE (mm/hr)	Relative Bias	Corr.	RMSE (mm/hr)
July	-0.04	0.24	1.68	0.03	0.31	1.31
August	-0.17	0.23	1.36	0.01	0.31	1.09
September	-0.23	0.22	1.12	-0.006	0.32	0.86

In addition to estimated rainfall comparison during training and testing period, the rainfall distribution from PERSIANN-CCS, PMW and probabilistic PERSIANN-CCS is calculated during July through September 2015. Figure 5.25 shows the CDFs for 4 selected Bins over the US. Based on CDF comparison we can conclude the probabilistic model cannot detect extreme rainfall values due to considering the expected value for rainfall retrieval.



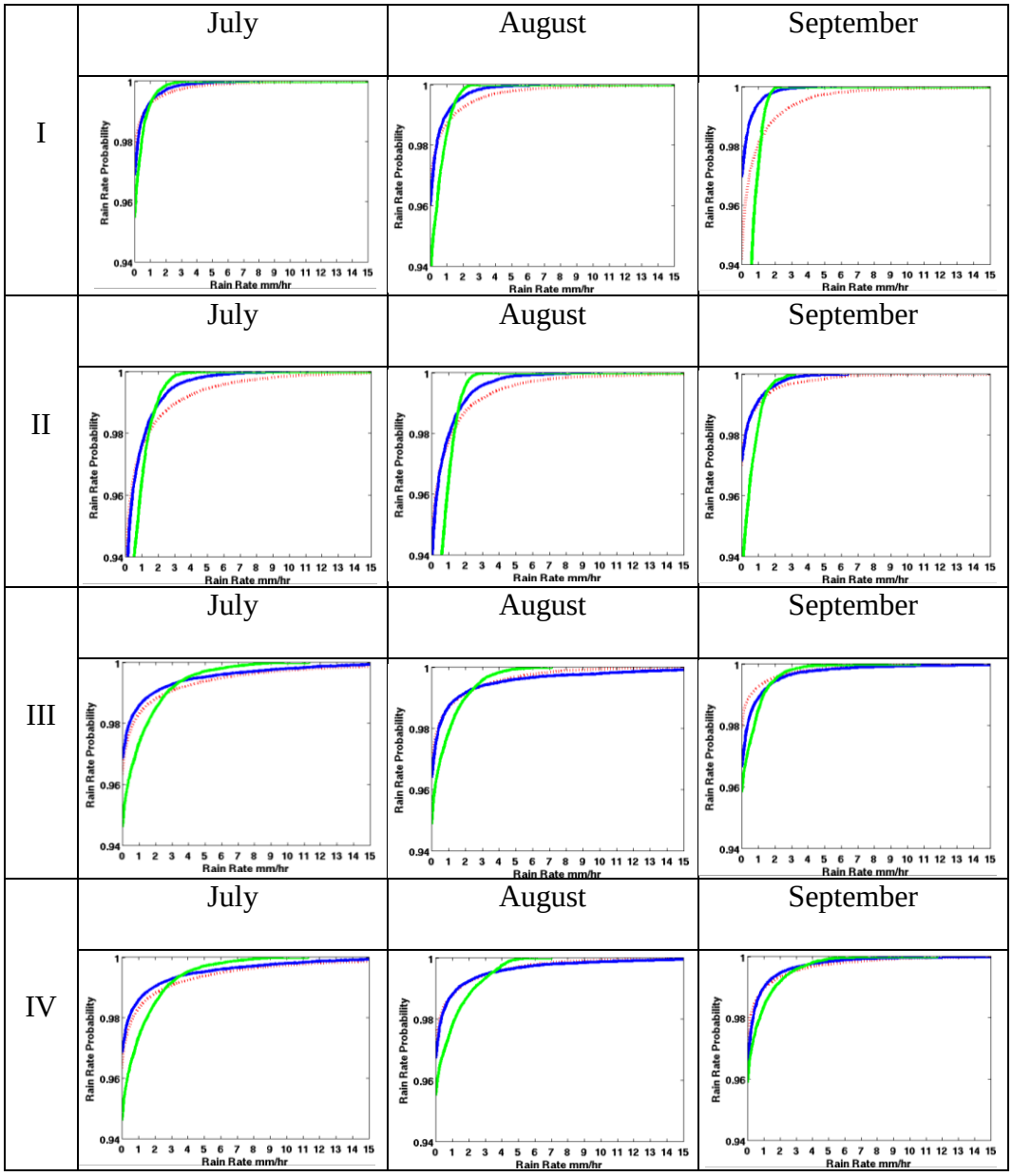


Figure 5.25: CDF comparison between PERSIANN-CCS & PMW and probabilistic model during validation year 2015. Blue solid line represents CDF for PMW data, red dashed line represents CDF for PERSIANN-CCS data, and green solid line represents CDF for probabilistic model.

CHAPETR 6: Summary, Conclusion and Future work

6.1. Summary and Conclusions

In the recent years precipitation estimation from satellites has become very popular. Many researches are ongoing to enhance the quality of rainfall retrieval from satellites. This research is mainly done in support of GPM project. PERSIANN-CCS as an IR only based algorithm in combination with PMW satellite data provides rainfall estimation at fine spatial and temporal resolution through the IMERG algorithm. The PERSIANN-CCS is a patch based algorithm that for estimating precipitation relies on IR data from GEO satellites. This algorithm retrieves precipitation based on a maximum temperature threshold level equal to 253K which causes missing rainfall from warm clouds and falsely identifies rainfall for non-precipitating cold clouds. Another limitation of this algorithm is due to the method for pixel based rainfall estimation from patch based cloud classification. PERSIANN-CCS classifies cloud patches into 400 groups, for which each cloud group has a specific Tb-RR relationship based on fitting an exponential function to data. By fitting a nonlinear exponential function to the data of each group, we assume that the heavier rainfall rates are associated with clouds with lower brightness temperature. In order to make the data match this assumption we redistribute the data in a way that colder clouds contain more rainfall. The redistribution of data loses the actual relationship which results in increasing the error of rainfall measurement from PERSIANN-CCS. In order to improve the PERSIANN-CCS we have addressed three approaches in this research, which in summary achieved the following results in this research.

- ***Using multi- satellite data and combining GEO and LEO based satellite estimates to improve precipitation derived from PERSIANN-CCS algorithm over the globe. How much improvement is achieved by combing PERSIANN-CCS and PMW satellite data over different regions and seasons knowing that PMW sensors retrieve precipitation more accurately?***

GEO satellites have high spatial and temporal resolution but the precipitation estimation from this type of satellites is less reliable due to the information from IR, VIS, and water vapor channels are indirectly related to rainfall rate. On the other hand, PMW satellites have coarser temporal resolution but the information collected from PMW sensors are associated with physical cloud properties which results in more reliable precipitation estimates. To benefit from high spatial and temporal resolution of GEO-based data and also reliable estimation of LEO satellites, we combined the data from PERSIANN-CCS with PMW-COMB based on a statistical framework to reduce the bias in PERSIANN-CCS rainfall estimates. The methodology is based on Probability Matching Method (PMM). The lookup tables are built based on 4 years of climatological data from PERSIANN-CCS and PMW data for 5x5 degree sample size. The model is validated over the globe and also US for 3 winter month and 3 summer month during 2012. The results are evaluated by calculating statistical parameters such as bias, correlation coefficient, and RMSE. The results show that all three parameters are improved during summer and winter time over the globe. However, due to sample size problem from PMW data over high latitude regions during winter time we only rely on the improvement during summer time over high latitude regions. Also, PERSIANN-CCS and MA-PERSIANN-CCS are compared with radar data over the CONUS. The results show that the bias is improved for MA-PERSIANN-CCS during summer time but

there is no significant improvement for other parameters. During winter time, no significant improvement is observed since the PMW and radar data are not very well matched over the US.

- ***PERSIANN-CCS algorithm estimates precipitation up to certain brightness temperature threshold. This causes that PERSIANN-CCS to miss rainfall from clouds with temperature higher than 253 K. This issue is investigated to see how much precipitation can PERSIANN-CCS capture if we raise the temperature threshold to higher levels.***

The PERSIANN-CCS algorithm is a patch based algorithm that defines cloud patches based on segmentation technique. After the segmentation we extract coldness, textural, and geometrical features at 3 different threshold levels 220, 235, and 253. This means that technically, PERSIANN-CCS algorithm does not consider clouds with temperature higher than 253K resulting in underestimation of rainfall. To overcome the underestimation of the PERSIANN-CCS algorithm, we extended the temperature threshold up too 300K to be able to enhance the rainfall identification. The data was trained during summer time of 2014 and was tested during summer 2015. The results indicate that the bias is improved after extending the threshold but no significant improvement for correlation coefficient and RMSE. This is because PERSIANN-CCS assigns the rainfall to the coldest part of the clouds which is not always the correct locations. In addition to this, the fundamental assumption behind this model is colder clouds contain heavier rainfalls which cause underestimation of warm clouds rainfall even after raising the threshold to 300K.

- ***PERSIANN-CCS algorithm retrieves precipitation based on the exponential relationship between Tb-RR (cloud top temperature and rainfall intensity). The rational for using the exponential relationship is that the heavier rainfalls rates are associated with colder clouds. Two potential issues arise: 1) falsely estimate precipitation from non-precipitating cold clouds (cirrus clouds and) 2) missing rainfall from warm clouds (Behrangi et al. 2010a). In this research the degree of improvement which can be achieved for implementing a probabilistic rainfall estimation model based on the actual Tb-RR relationship is explored.***

As discussed earlier, PERSIANN-CCS estimates rainfall based on the Tb-RR relationship. The fundamental assumption for PERSIANN-CCS precipitation retrieval is that the cold clouds contain heavier rainfall so that non-linear exponential function can be defined for calculating rainfall rate from cloud brightness temperature. To fit the non-linear exponential function, we redistribute the data in such a way that assign higher intensity of rainfall to colder clouds. The redistribution of rainfall cause distorting the actual Tb-RR relationship and results in underestimation of rainfall for warmer clouds and overestimation of rainfall for colder clouds. To improve the capability of PERSIANN-CCS in rainfall estimation and also considering the actual Tb-RR relationship, we developed a probabilistic framework for rainfall estimation. The probabilistic model is generated based on the fitting a lognormal distribution to the data for each brightness temperature and using the conditional mean rainfall rate as the predicted value for each cloud temperature. The performance of the model was evaluated over a validation period. The binary and quantitative analysis shows that the probabilistic model works better than the current PERSIANN-CCS.

6.2. Future Work

The following suggestions can be used for future work towards improving precipitation estimation from remotely sensed data.

1. In order to improve the PERSIANN-CCS over high latitude regions during winter time we should increase the number of samples. Considering more climatology data as well as extending the sample size instead of 5x5 degree to 15x15 can improve the bias over high latitude regions during winter.
2. GOES-R is the next generation of GOES satellites that has significantly improved detection and observation of meteorological parameters. The satellite is equipped with advanced imaging with higher spatial and temporal resolution than current GOES satellite. The new data from the GOES-R can be used for developing GEO bases precipitation algorithm using multi-channel data.

APPENDIX A

Binary evaluation equations:

1) Probability Of Detection (POD)

$$POD = \frac{Hit}{Hit + Miss}$$

Range: 0-1; Perfect Score = 1;

2) False Alarm Ratio (FAR)

$$FAR = \frac{False}{Hit + Flase}$$

Range: 0-1; Perfect Score = 0;

3) Heidke Skill Score (HSS)

$$HSS = \frac{2 \times [(Hit \times Reject) - (False \times Miss)]}{(Hit + Miss) \times (Miss + Reject) + (Hit + False) \times (False + Reject)}$$

Range: 0-1; Perfect Score = 1

Hit= Number of pixels correctly identified as rain;
Miss = Number of pixels incorrectly identified as no rain;
False= Number of Pixels incorrectly identified as rain;
Reject= Number of pixels correctly identified as no rain;

Precipitation amount verification equations:

1) Relative Bias

$$\frac{\sum SIM - \sum OBS}{\sum OBS}$$

Range: $-\infty \; to \; +\infty$; Perfect Value = 0;

2) Correlation Coefficient

$$\frac{\sum(SIM_i - \overline{SIM})(OBS_i - \overline{OBS})}{\sqrt{\sum(SIM_i - \overline{SIM})^2}\sqrt{\sum(OBS_i - \overline{OBS})^2}}$$

Range: -1 to 1; Perfect Value = 1;

3) Root Mean Square Error (RMSE)

$$\sqrt{\frac{1}{N}\sum(SIM_i - OBS_i)^2}$$

Range: 0 to ∞; Perfect Value = 0;

SIM = Simulation; OBS = Observation

7. References

Adler R. F., Negri a. J., (1988). A Satellite infrared technique to estimate tropical convective and stratiform rainfall. *J. Appl. Meteor*., 27, 30-51, doi: 10.1175/1520-0450(1988)027<0030:ASITTE>2.0.CO;2

Adler, R.F., A.J. Negri, P.R. Keehn, and I.M. Hakkarinen, (1993): Estimation of Monthly Rainfall over Japan and Surrounding Waters from a Combination of Low-Orbit Microwave and Geosynchronous IR Data. *J. Appl. Meteor*., **32**, 335–356, doi: 10.1175/1520-0450(1993)032<0335:EOMROJ>2.0.CO;2.

Arkin, P.A. and P. Xie, (1994): The Global Precipitation Climatology Project: First Algorithm Intercomparison Project. *Bull. Amer. Meteor*. Soc., 75, 401–419, doi: 10.1175/1520-0477(1994)075<0401:TGPCPF>2.0.CO;2.

Barret E. C., (2001). Satellite remote sensing of precipitation: progress and problems. IAHS, publication number 267.

Behrangi, A., K. Hsu, B. Imam, S. Sorooshian, G.J. Huffman, and R.J. Kuligowski, (2009): PERSIANN-MSA: A Precipitation Estimation Method from Satellite-Based Multispectral Analysis. *J. Hydrometeor.,* **10**, 1414–1429, doi: 10.1175/2009JHM1139.1.

Behrangi, A., K. Hsu, B. Imam, and S. Sorooshian, (2010): Daytime Precipitation Estimation Using Bispectral Cloud Classification System. J. Appl. Meteor. Climatol., **49**,1015–1031, doi: 10.1175/2009JAMC2291.1.

Behrangi, A., B. Imam, K. Hsu, S. Sorooshian, T.J. Bellerby, and G.J. Huffman, (2010): REFAME: Rain Estimation Using Forward-Adjusted Advection of Microwave Estimates. J. Hydrometeor., 11, 1305–1321, doi: 10.1175/2010JHM1248.1.

Behrangi, A., K. Andreadis, J.B. Fisher, F. Turk, S. Granger, T. Painter, and N. Das, (2014): Satellite-Based Precipitation Estimation and Its Application for Streamflow Prediction over Mountainous Western U.S. Basins. J. Appl. Meteor. Climatol., **53**, 2823–2842, doi: 10.1175/JAMC-D-14-0056.1.

Bellerby, T., K. Hsu, and S. Sorooshian, (2009): LMODEL: A Satellite Precipitation Methodology Using Cloud Development Modeling. Part I: Algorithm Construction and Calibration. J. Hydrometeor., **10**, 1081–1095, doi: 10.1175/2009JHM1091.1.

Calheiros R. V., Zawadzki I., (1986). Reflectivity-Rain Rate Relationships for Radar Hydrology in Brazil. Journal of Climate and Applied Meteorology, 26, 118–132, doi: 10.1175/1520-0450(1987)026<0118:RRRRFR>2.0.CO;2.

Chen S., Gourley J. J., Hong Y., Kirstetter P. E., Zhang J., Howard K., Flamig Z. L., Hu J., and Qi Y., (2013). Evaluation and uncertainty estimation of NOAA/NSSL next-generation national mosaic quantitative precipitation estimation product (Q2) over the Continental United States. J. Hydrometeor, 14, 1308–1322, doi: 10.1175/JHM-D-12-0150.1

Ciach G. J., Krajewski W. F., and Smith J. A., (1997). Comments on "The window probability matching method for rainfall measurements with radar". J. Appl. Meteor., 36, 243–246, doi: 10.1175/1520-0450(1997)036<0243:COTWPM>2.0.CO;2

Ebert, E.E. and M.J Manton, (1998): Performance of Satellite Rainfall Estimation Algorithms during TOGA COARE. J. Atoms. Sci., 55,1537-1557, doi: 10.1175/1250-0469(1998)055<1537:POSREA>2.0CO;2.

Ferraro, R., N. Grody, D. Forsyth, R. Carey, A. Basist, J. Janowiak, F. Weng, G. F. Marks, and R. Yanamandra (1994). Microwave measurements produce global climatic, hydrologic data, Eos Trans. AGU, 75(30), 337–343, doi: 10.1029/94EO00988.

Ferraro, R.R. and G.F. Marks, (1995): The Development of SSM/I Rain-Rate Retrieval Algorithms Using Ground-Based Radar Measurements. *J. Atmos. Oceanic Technol.*,**12**, 755–770, doi: 10.1175/1520-0426(1995)012<0755:TDOSRR>2.0.CO;2.

Griffith C. G., Woodley W. L., Grube P. G., Martin D. W., Stout J., and Sikder D. N., (1978). Rain estimation from Geosynchronous satellite imagery—visible and infrared studies. *Mon. Wea. Rev.*, 106, 1153–1171, doi: 10.1175/1520-0493(1978)106<1153:REFGSI?2.0.CO;2.

Groiman, P. and D.R. Legates, (1994): The Accuracy of United States Precipitation Data. Bull. Amer. Meteor. Soc., 75,215-227, doi: 10.1175/1520-0477(1994)075<0215:TAOUSP>2.0.CO:2.

Hong Y., Hsu K., Sorooshian S., and Gao X., (2004). Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. J. Appl. Meteor., 43, 1834–1853, doi: 10.1175/JAM2173.1

Hong, Y., K.-L. Hsu, S. Sorooshian, and X. Gao (2005), Improved representation of diurnal variability of rainfall retrieved from the Tropical Rainfall Measurement Mission Microwave Imager adjusted Precipitation Estimation From Remotely Sensed Information Using Artificial Neural Networks (PERSIANN) system, J. Geophys. Res., 110, D06102, doi:10.1029/2004JD005301.

Hsu, K., X. Gao, S. Sorooshian, and H.V. Gupta, (1997): Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks. *J. Appl. Meteor.*, **36**,1176–1190, doi: 10.1175/1520-0450(1997)036<1176:PEFRSI>2.0.CO;2.

Hsu K., Hong Y., Sorooshian S., (2007): Rainfall estimation using cloud patch classification map. Measuring Precipitation From Space, Volume 28 of the series Advances In Global Change Resaerch pp 329-342.

Hou, A.Y., R.K. Kakar, S. Neeck, A.A. Azarbarzin, C.D. Kummerow, M. Kojima, R. Oki, K. Nakamura, and T. Iguchi, (2014): The Global Precipitation Measurement Mission.Bull. Amer. Meteor. Soc., **95**, 701–722, doi: 10.1175/BAMS-D-13-00164.1.

Huffman, G.J., R.F. Adler, M.M. Morrissey, D.T. Bolvin, S. Curtis, R. Joyce, B. McGavock, and J. Susskind, (2001): Global Precipitation at One-Degree Daily Resolution from Multisatellite Obsrvations. J. Hyrometeor. 2,36-50, doi: 10.1175/15257541(2201)002<0036: GPAODD>2.0.CO:2.

Huffman G. J., Adler R. F., Bolvin D. T., Gu G., Nelkin E. J., Bowman K. P., Hong Y., Stocker E. F., and Wolff D. B., (2007). The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear combined-sensor precipitation estimates at fine scales. Journal of Hydrometeorology, 8, 38–55, doi: 10.1175/JHM560.1.

Huffman G. J., Bolvin D. T., Braithwaite D., Hsu K., Joyce R., and Xie P. (2015). NASA Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG). Algorithm theoretical basis document, version 4.5 NASA (30 pp.)

Huffman G. J., Bolvin D. T., Nelkin E. J., (2015): Integrated Multi-satelliteE Retrievals for GPM (IMERG) Technical Documentation. [available online at https://pmm.nasa.gov/sites/default/files/document_files/IMERG_doc.pdf]

Joyce R. J., Janowaik J. E., Arkin P. A., and Xie P., (2004). CMORPH: A Method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. J. Hydrometeor, 5, 487–503, doi: 10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2

Joyce R. J., Xie P., (2011). Kalman filter–based CMORPH. J. Hydrometeor, 12, 1547–1563, doi: 10.1175/JHM-D-11-022.1.

Kidd, C., D.R. Kniveton, M.C. Todd, and T.J. Bellerby, (2003): Satellite Rainfall Estimation Using Combined Passive Microwave and Infrared Algorithms. J. Hydrometeor.,**4**, 1088–1104, doi: 10.1175/1525-7541(2003)004<1088:SREUCP>2.0.CO;2.

Kirstetter, P.-E., J. J. Gourley, Y. Hong, J. Zhang, S. Moazamigoodarzi, C. Langston, and A. Arthur (2015), Probabilistic precipitation rate estimates with ground-based radar networks, Water Resour. Res., 51, 1422–1442, doi:10.1002/2014WR015672.

Kirstetter, P.E., G. Delreu, B. Boudevilian, and C. Obled, (2010): Toward an error model for radar quantitative precipitation estimation in the Cevennes-Vivarias region, France, J. Hydrol., 394(1-2),28-41.

Krajewski, W.F. and J.A. Smith, 1991: On the Estimation of Climatological Z–R Relationships. *J. Appl. Meteor.*, **30**, 1436–1445, doi:10.1175/1520-0450(1991)030<1436:OTEOCR>2.0.CO;2.

Kummerow, C. and L. Giglio, (1995): A Method for Combining Passive Microwave and Infrared Rainfall Observations. J. Atmos. Oceanic Technol., **12**, 33–45, doi:10.1175/1520-0426(1995)012<0033:AMFCPM>2.0.CO;2.

Kummerow C., Hong Y., Olson W. S., Yang S., Adler R. F., McCollum J., Ferraro R., Petty G., Shin D-B, Wilheit T. T., (2001). The evolution of the Goddard Profiling Algorithm (GPROF) for rainfall estimation from passive microwave sensors. J. Appl. Meteor., 40, 1801–1820, doi: 10.1175/1520-0450(2001)040<1801:TEOTGP>2.0.CO;2

Kuligowski, R.J., (2002): A Self-Calibrating Real-Time GOES Rainfall Algorithm for Short-Term Rainfall Estimates. J. Hydrometeor., **3**, 112–130, doi:10.1175/1525-7541(2002)003<0112:ASCRTG>2.0.CO;2.

Kuligowski R. J., (2009). The self-calibrating multivariate precipitation retrieval (SCaMPR) for high resolution, low-latency satellite-based rainfall estimation. in satellite rainfall applications for Surface Hydrology, eds. M. Gebremichael and F.Hossain, PP.39-48, Springer, New York, doi:10.1007/978-90-481-2915-7_3.

Legates, D. R. and DeLiberty, T. L. (1993), PRECIPITATION MEASUREMENT BIASES IN THE UNITED STATES[1]. JAWRA Journal of the American Water Resources Association, 29: 855–861. doi:10.1111/j.1752-1688.1993.tb03245.x

Levizzani, V., Porcú, F., Marzano, F. S., Mugnai, A., Smith, E. A. and Prodi, F. (1996), Investigating a SSM/I microwave algorithm to calibrate Meteosat infrared instantaneous rainrate estimates. Met. Apps, 3: 5–17. doi:10.1002/met.5060030102.

Liu, Z., 2016: Comparison of Integrated Multisatellite Retrievals for GPM (IMERG) and TRMM Multisatellite Precipitation Analysis (TMPA) Monthly Precipitation Products: Initial Results. *J. Hydrometeor.,* **17**, 777–790, doi: 10.1175/JHM-D-15-0068.1.

Mahrooghy, M., V.G. Anantharaj, N.H. Younan, J. Aanstoos, and K. Hsu, (2012): On an Enhanced PERSIANN-CCS Algorithm for Precipitation Estimation. *J. Atmos. Oceanic Technol.,* **29**, 922–932, doi: 10.1175/JTECH-D-11-00146.1.

Mahroughi, M., V. G. Anantharaj, N. H. Younan, W. A. Petersen, K. Hsu, A. Behrangi, and J. Anantoos, (2013): Augmenting satellite precipitation estimation with lightning information. International Journal of Remote Sensing, 34:16, 5796-5811,doi: 10.1080/01431161.2013.796100.

Marzano, F.S.; Palmacci, M.; Cimini, D.; Giuliani, G.; Turk, F.J., (2004). "Multivariate statistical integration of satellite infrared and microwave radiometric measurements for rainfall retrieval at the geostationary scale," Geoscience and Remote Sensing, IEEE Transactions on , vol.42, no.5, pp.1018,1032, doi: 10.1109/TGRS.2003.820312.

Morin, E., W.F. Krajewski, D.C. Goodrich, X. Gao, and S. Sorooshian, 2003: Estimating Rainfall Intensities from Weather Radar Data: The Scale-Dependency Problem. *J. Hydrometeor.,* **4**, 782–797, doi: 10.1175/1525-7541(2003)004<0782:ERIFWR>2.0.CO;2.

Piman T., Babel M. S., Das Gupta A., and Weesakul S.,(2007). Development of a window correlation matching method for improved radar rainfall estimation, Hydrol. Earth Syst. Sci., 11, 1361-1372, doi:10.5194/hess-11-1361-2007.

Rosenfield D., Amitai E., (1998). Comparison of WPMM versus regression for evaluating Z-R relationships. Journal of Applied Meteorology, vol. 37, Issue 10, pp.1241-1249, doi: 10.1175/1520-0450(1998)037<1241:COWVRF>2.0.CO;2

Sorooshian, S., Hsu, K. L., Gao, X., Gupta, H. V., Imam, B., & Braithwaite, D. (2000). Evaluation of PERSIANN system satellite–based estimates of tropical rainfall. Bulletin of the American Meteorological Society, 81(9), 2035-2046.

Scofield R. A., (1987). The NESDIS operational convective precipitation estimation technique. Mon. Wea. Rev., 115, 1773–1793, doi: 10.1175/1520-0493(1987)115<1773:TNOCPE>2.0.CO;2.

Tapiador F. J., (2008). A physically-based satellite rainfall estimation method using fluid dynamics modelling. Int. J. Remote. Sens., 29 (20) (2008), pp. 5851–5862, doi: 10.1080/01431160802029677.

Todd, M.C., C. Kidd, D. Kniveton, and T.J. Bellerby, (2001): A Combined Satellite Infrared and Passive Microwave Technique for Estimation of Small-Scale Rainfall. J. Atmos. Oceanic Technol., **18**, 742–755, doi: 10.1175/1520-0469(2001)058<0742:ACSIAP>2.0.CO;2.

Turk, F. J., J. Hawkins, E. A. Smith, F. S. Marzano, A. Mugnai, and v. Levizzani, (2000): Combining SSM/I, TRMM and infrared. geostationary satellite data in a near realtime fashion for rapid precipitation updates: Advantages and limitations. Proc., EUMETSAT Meteorological Satellite Data Users' Conf., Vol. 2, Bologna, Italy,EUMETSAT, 705–707.

Turk F. J. , E. E. Ebert, H. J. Oh, B. J. Sohm, V. Levizzani, E. A. Smith, and R. R. Ferraro, (2003): Validation of an operational global precipitation analysis at short time scales, 12th Conf. Sat. Meteor. And Ocean., AMS, 9-13 February, Long Beach, CA, CD-ROM.

Vicente G. A., Scofield R. A., Menzel W. P., (1998). "The operational GOES infrared rainfall estimation technique." Bulletin of the American Meteorological Society 79.9, 1883-1898, doi: 10.1175/1520-0477(1998)079<1883:TOGIRE>2.0.CO;2.

Weng, F., R. R. Ferraro, and N. C. Grody (1994), Global precipitation estimations using Defense Meteorological Satellite Program F10 and F11 special sensor microwave imager data, J. Geophys. Res., 99(D7), 14493–14502, doi:10.1029/94JD00961.

Weng, F., L. Zhao, R. R. Ferraro, G. Poe, X. Li, and N. C. Grody (2003). Advanced microwave sounding unit cloud and precipitation algorithms, Radio Sci., 38, 8068, doi: 10.1029/2002RS002679.

Wu, R., J.A. Weinman, and R.T. Chin, (1985): Determination of Rainfall Rates from GOES Satellite Images by a Pattern Recognition Technique. *J. Atmos. Oceanic Technol.*, **2**, 314–330, doi: 10.1175/1520-0426(1985)002<0314:DORRFG>2.0.CO;2.

Wu R., Weinman J. A., and Chin R. T., (1985). Determination of rainfall rates from GOES satellite images by a pattern recognition technique. J. Atmos. Oceanic Technol., **2**, 314–330, doi: 1520-0426(1985)002<0314:DORRFG>2.0.CO;2.

Yilmaz, K.K., T.S. Hogue, K. Hsu, S. Sorooshian, H.V. Gupta, and T. Wagener, (2005): Intercomparison of Rain Gauge, Radar, and Satellite-Based Precipitation Estimates with Emphasis on Hydrologic Forecasting. J. Hydrometeor., 6, 497–517, doi: 10.1175/JHM431.1.

Zhang J., Qi Y., Langston C., Kaney B., (2012). Radar Quality Index (RQI): A combined measure of beam blockage and VPR effects in a national network. Proceedings International Symposium on Weatger Radar and Hydrology, IAHS Publ. 351-25.