Title
A Deep Learning Framework for Precipitation Estimation from Bispectral Satellite Information

Author
Tao, Yumeng

Publication Date
2017

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

A Deep Learning Framework for Precipitation Estimation from Bispectral Satellite Information

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Civil Engineering

by

Yumeng Tao

Dissertation Committee:
Professor Soroosh Sorooshian, Chair
Professor Alexander Ihler
Professor Kuolin Hsu

2017

DEDICATION

To

My parents, who taught me the secret of happiness,

and

Those I love

Thank you for being part of this journey.

Everything comes to you at the right moment.

Be patient.

Be grateful.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS AND ABBREVIATIONS

COR	Pearson's Correlation Coefficient
CSI	Critical Success Index
DNN	Deep Neural Network
FAR	False Alarm Ratio
FBI	Frequency Bias
GEO	geosynchronous-Earth-orbiting
GOES	Geostationary Operational Environmental Satellite
HSS	Heidke Skill Score
IR	Infrared
KL Divergence	Kullback-Leibler divergence
LEO	low-Earth-orbiting
MSE	Mean Squared Error
PERSIANN	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks
PERSIANN-CCS	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks - Cloud Classification System
POD	Probability of Detection
R/NR	Rain/No-Rain
SDAEs	Stacked Denoising Auto-Encoders
Stage IV	Stage IV Radar and Gauge Precipitation Data
WV	Water Vapor
VIS	Visible

ACKNOWLEDGMENTS

I would like to thank sincerely my advisor Professor Soroosh Sorooshian for giving me the opportunity to join the Center for Hydrometeorology and Remote Sensing (CHRS) at University of California, Irvine (UCI). His guidance, in both my research and many aspects of life, has been invaluable to me during my Ph.D. study. Without his support and open mind, I would not have been able to acquire and appreciate the rich experiences I underwent beyond the Ph.D. dissertation. I would also like to thank Professor Xiaogang Gao and Professor Kuolin Hsu for their continuous encouragement, patience, and passion. Their advice was immensely helpful throughout the various stages of my study.

My special thanks go to Professor Alexander Ihler, who led me to the world of machine learning. He has been a great mentor and friend on my research journey. His intuitions and insights were extremely valuable during my study, especially when we reached bottlenecks. I was also truly inspired by his intelligence, diligence, generosity, and passion for research.

I would also like to thank Professor Qingyun Duan. It was his kind recommendation that initiated my journey at UCI. It was also his guidance that introduced me to the world of research.

My sincerest gratitude goes to my colleagues at CHRS, who were always supportive and caring. I was very lucky to work in an environment permeated by positive energy. I would like to give special acknowledgment to Dr. Tiantian Yang, who has been like a big brother to me since my first day at Irvine. Many thanks are also due to Diane Hohnbaum and Dan Braithwaite, who are the true VIPs of the center. Thank you for making everything so smooth and easy.

I would also like to thank all my friends here in Southern California and around the world. You guys made me feel secure in unfamiliar places. A special acknowledgment goes to my dearest friend, Yiqing Lin, who has always been like my sister and who has always been there for me.

My deepest gratitude goes to my family, who offered me endless love and support. To the world's best parents, Feng Tao and Xuefang Liu, thank you for making me the happy person I am and for teaching me the most important things in life. To my grandparents, thank you for continuously spoiling me. And to my special one, Zhuozhou Liu, thank you for always being a great friend and for starting this new journey with me.

Primary financial support for this work was made available by the National Science Foundation Cyber-Enabled Sustainability Science and Engineering (Grant CCF-1331915) and the NASA Earth and Space Science Fellowship (Grant NNX15AN86H).

CURRICULUM VITAE

Yumeng Tao

EDUCATION

Ph.D., Civil Engineering, University of California, Irvine, June 2017

M.S., Civil Engineering, University of California, Irvine, March 2016

M.S., Statistics, University of California, Irvine, June 2015

B.S., Statistics, Beijing Normal University, Beijing, China, June 2012

PROFESSIONAL EXPERIENCE

2012 – 2017 Graduate Student Researcher, Center for Hydrometeorology and Remote Sensing, The Henry Samueli School of Engineering, University of California, Irvine, Irvine, California, United States

2016 Data Science for Social Good Fellow (Research Intern), IBM Thomas J. Watson Research Center, Yorktown Heights, New York, United States

2015 Research Scientist Intern, Amazon, Seattle, Washington, United States

2010 – 2012 Undergraduate Research Assistant, College of Global Change and Earth System Science, Beijing Normal University, Beijing, China

HONORS & AWARDS

- NASA Earth and Space Science Fellowship (NESSF) Award, 2015 – 2017

- IEEE Computational Intelligence Society (IEEE CIS) Students Travel Grant, 2016
- Outstanding Student Paper Award of American Geophysical Union 2015 Fall Meeting, San Francisco, California, December, 2015
- University of California Irvine Graduate Student Research and Travel Grant, 2015
- University of California Irvine Associated Graduate Students Travel Grant, 2014 & 2015
- Graduate Student Fellowship, Department of Civil and Environmental Engineering, University of California, Irvine, 2012 – 2013
- Fellowship of Chinese National University Student Innovation Research Plan, 2010 – 2012
- Academic Scholarships of Beijing Normal University, 2008 – 2011
- Honorable Mention in Mathematical Contest in Modeling, February 2011
- 1st National Prize of Contemporary Undergraduate Mathematical Contest in Modeling, October 2010
- Outstanding Student Leader, September 2010

PUBLICATIONS

- **Yumeng Tao**, K. Hsu, A. Ihler, X. Gao, and S. Sorooshian, 2017: A Two-Stage Deep Neural Network Framework for Precipitation Estimation from Bispectral Satellite Information. *Journal of Hydrometeorology.* (Submitted)
- **Yumeng Tao**, X. Gao, A. Ihler, S. Sorooshian, and K. Hsu, 2017: Precipitation Identification with Bispectral Satellite Information Using Deep Learning Approaches. *Journal of Hydrometeorology*, 18(5), 1271–1283.

- **Yumeng Tao**, D. Bhattacharjya, A. Heching, A. Vempaty, M. Singh, F. Lam, K. Varshney, and A. Mojsilović, 2017: Statistical Analysis of Peer Detailing for Children's Diarrhea Treatments, *AAAI technical report.*
- **Yumeng Tao**, D. Bhattacharjya, A. Heching, A. Vempaty, M. Singh, F. Lam, J. Houdek, M. Abubakar, A. Abdulwahab, T. Braimoh, N. Ihebuzor, A. Mojsilović, and K. Varshney, 2017: Effectiveness of Peer Detailing in a Diarrhea Program in Nigeria*. IBM Journal of Research and Development.* (Accepted)
- T. Yang, **Yumeng Tao**, J. Li, Q. Zhu, L. Su, and X. Zhang, 2017: Multi-criterion Model Ensemble of CMIP5 Surface Air Temperature over China. *Theoretical and Applied Climatology.* (Accepted)
- **Yumeng Tao**, T. Yang, M. Faridzad, L. Jiang, and X. Zhang, 2017: Non-stationary Bias Correction of Monthly CMIP5 Temperature Projections over China using a Residual-based Bagging Tree Model*. International Journal of Climatology.* (Reviewed and revised)
- **Yumeng Tao**, X. Gao, A. Ihler, K. Hsu, and S. Sorooshian, 2016: Deep Neural Networks for Precipitation Estimation from Remotely Sensed Information. *2016 IEEE Congress on Congress on Evolutionary Computation (CEC),* 1349–1355*.*
- **Yumeng Tao**, X. Gao, K. Hsu, S. Sorooshian, and A. Ihler, 2016: A Deep Neural Network Modeling Framework to Reduce Bias in Satellite Precipitation Products. *Journal of Hydrometeorology*, 17(3): 931–945.
- **Yumeng Tao**, Q. Duan, A. Ye, W. Gong, Z. Di, M. Xiao, and K. Hsu, 2014: An Evaluation of Post-processed TIGGE Multimodel Ensemble Precipitation Forecast in Huai River Basin. *Journal of Hydrology,* 519: 2890–2905.

- P. Nguyen, S. Sellars, A. Thorstensen, **Yumeng Tao**, H. Ashouri, D. Braithwaite, K. Hsu, and S. Sorooshian, 2014: Satellites Track Precipitation of Super Typhoon Haiyan. *Eos, Transactions American Geophysical Union*, 95(16), 133–135.
- Y. Liu, Q. Duan, L. Zhao, A. Ye, **Yumeng Tao**, C. Miao, X. Mu, and J. Schaake, 2013: Evaluating the Predictive Skill of Post-processed NCEP GFS Ensemble Precipitation Forecasts in China's Huai River Basin, *Hydrological Processes*, 27(1): 57–74.

PRESENTATIONS

- **Yumeng Tao**, X. Gao, A. Ihler, S. Sorooshian, and K. Hsu: Rain/No-Rain Identification from Bispectral Satellite Information using Deep Neural Networks. American Geophysical Union 2016 Fall Meeting, San Francisco, California, United States, December 2016
- **Yumeng Tao**, K. Varshney, A. Mojsilović, D. Bhattacharjya, A. Heching, M. Singh, and A. Vempaty: Evaluation of CHAI's Dissemination of the Best Treatment for Children's Diarrhea in Nigeria. IBM Research Intern Symposium, Yorktown Height, New York, United States, August 2016
- **Yumeng Tao**, X. Gao, A. Ihler, K. Hsu, and S. Sorooshian: Deep Neural Networks for Precipitation Estimation from Remotely Sensed Information. 2016 IEEE World Congress on Computational Intelligence, Vancouver, Canada, July 2016
- **Yumeng Tao**, X. Gao, K. Hsu, S. Sorooshian, and A. Ihler: Precipitation Estimation from Multi-Source Information using Advanced Machine Learning Methods. Coalition for National Science Funding Exhibition, Washington, DC, United States, April 2016

- **Yumeng Tao**, X. Gao, K. Hsu, S. Sorooshian, and A. Ihler: Precipitation Estimation from Remotely Sensed Data using Deep Neural Networks. American Geophysical Union 2015 Fall Meeting, San Francisco, California, United States, December 2015
- **Yumeng Tao**, K. Li, and C. Jones: Visualization Enhancement for Fraud Monitoring of Pay with Amazon. Amazon Research Intern Symposium, Seattle, Washington, United States, August 2015
- **Yumeng Tao**, X. Gao, K. Hsu, S. Sorooshian, and A. Ihler: Precipitation Estimation from Remotely Sensed Data using Deep Neural Networks. Workshop on Understanding Climate Change from Data, Minneapolis, Minnesota, United States, August 2015
- **Yumeng Tao**, X. Gao, and S. Sorooshian: Precipitation Estimation from Remotely Sensed Data using Deep Neural Networks. Data Science Initiative, University of California, Irvine, Irvine, California, United States, May 2015
- Q. Duan, **Yumeng Tao**, and M. Xiao: Statistical Post-processing to Improve Hydrometeorological Forecasts. 93rd American Meteorological Society Annual Meeting, Austin, Taxes, United States, January 2013
- M. Xiao, **Yumeng Tao**, Q. Duan, G. Zhang, and L. Zhao: Multi-model Ensemble QPF for Southeastern China Using Bayesian Model Averaging. WMO 3rd International Conference on Quantitative Precipitation Estimation and Quantitative Precipitation Forecasting and Hydrology, Nanjing, China, October 2010

SERVICE & VOLUNTEERING

- Guest lecturer for UCI undergraduate/graduate Hydrology class, Irvine, California, United States, October 2016
- Reviewer of Journal of Hydrology: Regional Studies, 2015

- Volunteer for American Geophysical Union 2014/2015 Fall Meeting, San Francisco, California, United States, December 2014 /2015
- Reviewer of PeerJ, 2014
- Volunteer for UCI Chinese students & scholars association, 2012 – 2013
- Vice-president of student union of School of Mathematical Sciences, Beijing Normal University, 2009 – 2010

ABSTRACT OF THE DISSERTATION

A Deep Learning Framework for Precipitation Estimation

from Bispectral Satellite Information

By

Yumeng Tao

Doctor of Philosophy in Civil Engineering

University of California, Irvine, 2017

Professor Soroosh Sorooshian, Chair

Compared to ground-based precipitation measurements, satellite-based precipitation estimates have the advantage of global coverage and high spatiotemporal resolutions. However, the accuracy of satellite-based precipitation observations is still insufficient to serve many weather, climate, and hydrologic applications. In the development of a satellite-based precipitation product, the two most important aspects are the provision of sufficient precipitation-related information in the selected satellite data and the use of the proper methodologies to extract such information and link it to precipitation estimates.

In this dissertation, a state-of-the-art deep learning framework for precipitation estimation using bispectral satellite information, Infrared (IR) and water vapor (WV) channels, is

developed. I explore the effectiveness of deep learning techniques in extracting useful features from the satellite information and demonstrate the value of incorporating multiple satellite channels.

Specifically, I first provide a bias reduction model for satellite-based precipitation products using deep learning approaches to demonstrate their capability of extracting additional useful information from the satellite data. I then design a two-stage framework for precipitation estimation from bispectral information, consisting of an initial rain/no-rain (R/NR) binary classification, followed by a second stage estimating the non-zero precipitation amount. In the first stage, the model aims to eliminate the large fraction of NR pixels and to precisely delineate precipitation regions. In the second stage, the model aims to estimate the point-wise precipitation amount accurately while preserving its heavy-tailed distribution. Stacked denoising auto-encoders (SDAEs), a commonly used deep learning method, are applied in both stages.

The operational product, Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks Cloud Classification System (PERSIANN-CCS), serves as a baseline model throughout this dissertation. I evaluate performance along a number of common performance measures, including both R/NR and real-valued precipitation accuracy. Case studies focusing on the model's performance for specific events are also included. The experiments show that our proposed two-stage model outperforms original

PERSIANN-CCS in different verification periods over the central United States and in large-scale application. Therefore, the two-stage deep learning framework has the potential to serve as a more accurate and more reliable satellite-based precipitation estimation algorithm.

Chapter 1

Introduction

1.1 Satellite-Based Precipitation

Weather forecasts, climate variability, hydrology, and water resource management require sufficient information regarding precipitation, which is one of the most active variables in the natural water cycle. Precipitation observing, monitoring, and analyzing tools provide fundamental information needed for society to manage the recent increased number of extreme hydrometeorological events of recent decades. For instance, prediction and management of flooding events, the leading cause of severe-weather-related deaths, require instant and reliable precipitation information in high-spatiotemporal resolution, such as rainfall intensity and duration (AghaKouchak and Nakhjiri, 2012; Ajami et al., 2008; Anderson et al., 2008). Therefore, it is essential to provide near-real-time, accurate precipitation measurements with high resolution and near-global coverage.

Most common precipitation measurements consist of ground-based measurements, including gauge networks and radars and satellite-based measurements. Traditional

ground-based precipitation measurements suffer from restricted coverage and fail to provide precipitation data with a high-spatiotemporal resolution, especially for remote and unequipped regions (Ajami et al., 2008; Bellerby and Sun, 2005; Habib et al., 2014; Huffman et al., 2007; Nasrollahi et al., 2013). More specifically, gauges and radars are distributed unevenly with a positive correlation with population density, and there is little to no coverage over the ocean and sparsely populated areas. In addition, radar suffers from blockages in mountainous regions, such as the Western United States, and provides limited coverage in such regions. In extreme events, gauges and radars are likely to fail due to damage caused by the event.

By contrast, near-real-time satellite-based precipitation estimation products show the advantage of providing global fine-resolution precipitation maps, including oceans and developing regions. Unlike ground-based measurements, satellite instruments can monitor the full evolutionary picture of an extreme event without being affected by the event itself (Nguyen et al., 2014; Scofield and Kuligowski, 2003). For example, Nguyen et al. (2014) emphasized the value of satellite-based precipitation during Super Typhoon Haiyan in November, 2013, when all the ground-based measuring facilities were destroyed. Near-real-time satellite-based precipitation estimation products can provide valuable information for global weather and climate monitoring and for disaster planning and management.

To retrieve precipitation values, the most commonly used satellite information includes infrared (IR) data from geosynchronous-Earth-orbiting (GEO) satellites and passive microwave (PMW) data from low-Earth-orbiting (LEO) satellites (Hsu et al., 1997; Weng et al., 2003). PMW data have the advantage of being directly retrieved from actual hydrometeor content, whereas IR data are limited to cloud-top information (Behrangi et al., 2009a; Kummerow and Giglio, 1995). However, one main drawback of PMW data is their low-temporal resolution (Marzano et al., 2004). IR data have high spatial and temporal resolution and thus can monitor the complete evolution of a local precipitation event (Behrangi et al., 2009b). Many satellite-based precipitation estimation products have been developed and made operational in the past few years with the general concept of combining IR and PMW data (Hong et al., 2004; Hsu et al., 1997; Huffman et al., 2007; Joyce et al., 2004; Kidd et al., 2003; Kuligowski, 2002). In addition, most algorithms estimate precipitation from IR information and then post-process it with PMW information. With their high-spatiotemporal resolutions, GEO satellite observations continuously serve as fundamental sources in precipitation estimation (Behrangi et al., 2009b; Huffman et al., 2007).

Some studies have incorporated multiple channels of GEO satellite data because IR cloud-top brightness temperature data alone do not contain sufficient information for accurate precipitation retrieval (Ba and Gruber, 2001; Behrangi et al., 2009b; Behrangi et

al., 2010). One common choice is visible (VIS) wavelength cloud albedo data because of the data's high quality in daytime (Behrangi et al., 2010; Capacci and Conway 2005; Cheng et al., 1993; Grassotti and Garand, 1994; Griffith et al., 1978; Hsu et al., 1999). The obvious drawback for VIS data is that they are not available at nighttime. On the other hand, WV channel information has also proven to be quite effective for precipitation retrieval by a few previous studies, especially when used in combination with IR data (Behrangi et al., 2009a; Martin et al., 2008; Tjemkes et al., 1997).

One other important component for a satellite-based precipitation estimation product is its algorithms to produce precipitation estimations from the input data. Several studies emphasize that the key to making the best use of these datasets is to promote advanced methods that better assist in the extraction of valuable information from the raw data than the existing algorithms (Nasrollahi et al., 2013; Sorooshian et al., 2011). In recent years, deep learning techniques, also known as deep neural networks (DNNs), have been developed and widely applied in the machine learning and computer vision areas, constituting a breakthrough in managing large and complex datasets, especially for feature extraction from a large amount of image data (Bengio, 2009; Hinton et al., 2006; LeCun et al., 2015; Lee et al., 2007; Zhao et al., 2014). The techniques have proven to be effective in managing many real-world data mining problems (Glorot et al., 2011a; Hinton et al., 2006; Lu et al., 2013; Vincent et al., 2008). One particular advantage of a DNN is that it helps

extract representative features automatically and further assists estimation. The power of deep learning for image processing and feature extraction provides an opportunity to improve the accuracy of satellite-based precipitation estimation.

1.2 Research Motivation and Objectives

As various satellites' information and techniques to link such information to precipitation are available, the goal of this research is to develop a satellite-based precipitation estimation product with high accuracy and reliability. In this dissertation, I explore the application of state-of-the-art deep learning techniques and develop a framework for precipitation estimation from bispectral satellite information (IR and WV channels).

Specifically, this dissertation addresses the following objectives:

1) Demonstrate the effectiveness of deep learning algorithms for precipitation-related information extraction from satellite infrared imageries by developing a bias reduction system on satellite-based precipitation estimation products.

2) Design a deep learning framework that extracts useful features from bispectral satellite information to produce a consistent and high-quality precipitation estimation product.

3) Evaluate the effectiveness of the methodology and the value of additional satellite information on both binary precipitation identification and precipitation amount estimation by comparing its performance with operational products.

4) Evaluate the performance of the developed DNN model at a large-scale application and thus demonstrate its potential of future use at a near-global scale.

5) Assess the advantages and limitations of the developed DNN structure and identify potential future developments and investigations for operational applications.

1.3 Organization of the Dissertation

The general outline and dependence of this dissertation is described in Figure 1-1. The remainder of this dissertation is organized into six chapters. Chapter 2 summarizes related background information on precipitation estimation and deep learning techniques. Then, Chapters 3–6 describe the main novel contributions of this dissertation. Finally, summary, conclusions, and future directions are discussed in Chapter 7.

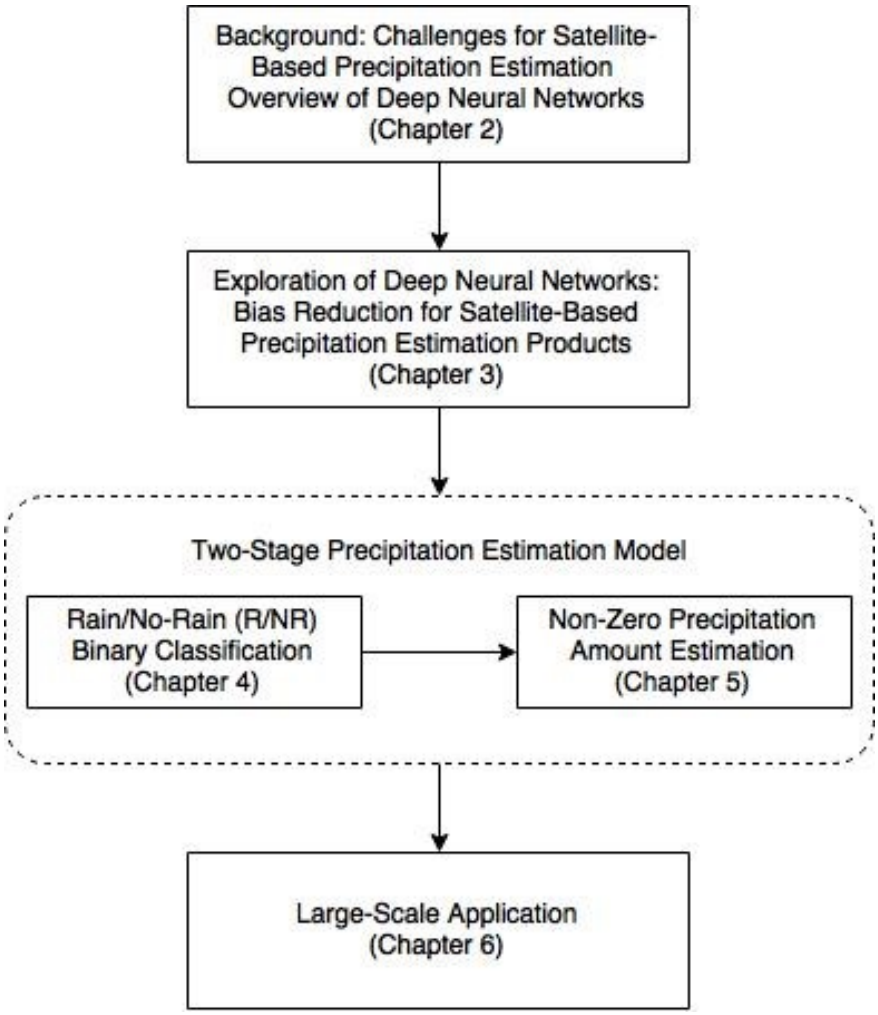

Figure 1-1 Outline and dependence of the dissertation

More specifically, each chapter contains the following materials:

Chapter 2 *provides background information on properties of precipitation data, satellite-based precipitation estimation products, and recent development of deep learning techniques. I first introduce the high imbalance of precipitation and the difficulties of its estimation. Then I overview the previous work on satellite-based precipitation estimation products and discuss in detail the operational product used as a*

baseline model in this study. Current development on DNNs is then discussed with a focus on stacked denoising auto-encoders (SDAEs), which is the primary training technique used in this study. Finally, the main study region and data used are described.

Chapter 3 *explores the effectiveness of DNNs for precipitation-related information extraction. As a first step, I take advantage of precipitation estimation made available by current satellite-based precipitation estimation products and produce a bias reduction process without adding extra satellite information. The current products are used to eliminate most no-rain samples and thus the DNN model can focus on capturing the residuals of the samples with a non-zero precipitation amount. I conclude with experiments that the DNN model can extract additional information from the same input data to reduce bias on the current products. Thus, by properly designing and training a DNN, a more effective satellite-based precipitation estimation model can be developed.*

Chapter 4 *develops a rain/no-rain (R/NR) binary classification neural network. This is the first step of the two-stage precipitation estimation model. This step focuses on correctly classify pixels to rain or no-rain pixel and thus eliminate the massive amount of no-rain cases. I first experimentally demonstrate the effectiveness of a DNN structure on the R/NR classification compared to the baseline model. Then I incorporate additional satellite information (water vapor channel) and adjust the DNN structure accordingly.*

The further improvement in the classification results proves the value of the additional information, which is the foundation of the bispectral precipitation estimation.

Chapter 5 *focuses on the part of real-valued precipitation amount estimation of the two-stage precipitation estimation model. As the second step of the two-stage model, I only deal with pixels classified as with a precipitation in the first step. The main challenge in this step is that the highly skewed the precipitation amount generally prevents data-driven model predicting large rainfall amount. To tackle with this issue, I incorporate the Kullback-Leibler divergence (KL divergence) as part of the objective function of the DNN and combine its effect with the conventional mean squared error (MSE). Combining with the binary classification structure, this is the complete two-stage model for precipitation estimation from bispectral satellite information. I then present experimental results to demonstrate the effectiveness of this model and compare it with the baseline model.*

Chapter 6 *presents a large-scale application of the developed two-stage model. One important property of satellite-based precipitation estimation products is their near-global coverage. Therefore, it is important to extend the model's application and evaluation beyond its main study region. I verify the model in the coverage of the whole United States with its calibration remained in the central United States to demonstrate its capability to preserve cloud and precipitation relationship regardless of experiment*

regions. It shows the potential of the model to serve as an operational product with near-global coverage.

Chapter 7 *gives the summary of the dissertation, the main conclusions, and the potential future research directions for the model further development and for preparation of an operational product.*

Chapter 2

Background

2.1 Challenges for Satellite-Based Precipitation Estimation

2.1.1 Imbalance of Precipitation Data

Satellite-based precipitation estimation products require high-spatiotemporal resolution, often at hourly scale and finer than 0.25° × 0.25° (approximately 25 km × 25 km) longitude and latitude. At this scale, the precipitation amount is expectably highly skewed. No rain is expected in most cases while heavy rain (> 10 mm/hr) is very unlikely to happen.

Table 2-1 Rainfall intensity percentages of radar observations of the central United States in summer 2013

	Intensity	Percentage
No Rain	0 – 0.1 mm/hr *	95.21%
Light Rain	0.1 – 2.5 mm/hr	3.59%
Moderate Rain	2.5 – 10 mm/hr	0.95%
Heavy Rain	> 10 mm/hr	0.25%

* 0.1 mm/hr is commonly used as a threshold to determine if there is precipitation due to radar sensors' uncertainty in very light rainfall.

As an example, Table 2-1 provides rainfall intensity spreads of radar observation of the central United States (30–45 N°, 90–105 W°) in summer (Jun.–Aug.) 2013. As shown in Table 2-1, over 95.2% of pixels show no rain (< 0.1 mm/hr), while only 1.2% of pixels show moderate to heavy rain (> 2.5 mm/hr, 0.95% of pixels at 2.5–10 mm/hr and 0.25% of pixels at > 10 mm/hr).

Such imbalance of precipitation data is one large challenge in applying the data-driven method towards precipitation estimation. Specifically, it makes the MSE, the most commonly used objective measure in supervised machine learning models, not suitable to work as the sole optimization objective for precipitation estimation. Optimized by the MSE solely, the models tend to be conservative and avoid predicting large rainfall amounts with great uncertainties. However, one main purpose of the satellite-based precipitation estimation is to track rainfall events, especially extreme events that may lead to floods and other environmental disasters. Therefore, in previous studies, the use of advanced machine learning approaches usually was limited to auxiliary steps, such as cloud unsupervised clustering. On the other hand, idealized assumptions are often imposed to ensure that large values can be predicted by the models. For example, PERSIANN assumes strictly that pixels with lower brightness temperature at the top of clouds have heavier precipitation on the ground (Arkin and Meisner, 1987; Hsu et al., 1997).

2.1.2 Multiple Objectives of Precipitation Estimation

Because of the imbalance of precipitation data and the use of precipitation observation, a reliable satellite-based precipitation estimation product needs to satisfy multiple objectives, including accurate binary R/NR identification, minimal error in pixel-wise point estimates, and similarly distributed as skewed precipitation observations. First, a successful precipitation estimation model needs to be capable of distinguishing R/NR pixels, which are usually measured by probability of detection (POD) and false alarm ratio (FAR). Then, such model needs to minimize the point-wise differences, which are usually measured by MSE, while predicting extreme events by capturing the distribution of precipitation values.

To develop a reliable structure of satellite-based precipitation estimation product, multiple objectives need to be considered and multiple performance measurements are provided. Table 2-2 provides the commonly used performance verification measurements for precipitation estimation, including performance for R/NR classification and precipitation amount estimation regression. These measurements are used to evaluate the performance of a precipitation estimation model with a reference to ground observation. In this research, the National Centers for Environmental Prediction (NCEP) Stage IV Radar and Gauge Precipitation Data (Stage IV) is served as ground observation (Lin and Mitchell, 2005; http://www.emc.ncep.noaa.gov/mmb/ylin/pcpanl/stage4/).

Table 2-2 Common verification measurements for satellite-based precipitation estimation products

	Verification Measures	**Formulas**	**Range and Desirable Value**
Classification Performance	Probability of Detection (POD)	$POD = \frac{TP}{TP + MS}$	Range: 0 to 1 Desirable Value: 1
	False Alarm Ratio (FAR)	$FAR = \frac{FP}{TP + FP}$	Range: 0 to 1 Desirable Value: 0
	Critical Success Index (CSI)	$CSI = \frac{TP}{TP + FP + MS}$	Range: 0 to 1 Desirable Value: 1
	Heidke Skill Score (HSS)	$HSS = \frac{2(TP \times TN - FP \times MS)}{[(TP + MS)(MS + TN) + (TP + FP)(FP + TN)]}$	Range: -∞ to 1 Desirable Value: 1
	Frequency Bias (FBI)	$FBI = \frac{TP + FP}{TP + MS}$	Range: 0 to +∞ Desirable Value: 1
Regression Performance	Bias	$Bias = \bar{x} - \bar{y}$	Range: -∞ to +∞ Desirable Value: 0
	Mean Squared Error (MSE)	$MSE = \frac{1}{N}\sum(x_i - y_i)^2$	Range: 0 to +∞ Desirable Value: 0
	Pearson's Correlation Coefficient (COR)	$COR = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2}\sqrt{\sum(y_i - \bar{y})^2}}$	Range: -1 to 1 Desirable Value: 1

TP denotes the number of true positive events, MS denotes the number of missing events, FP denotes the number of false positive events, TN denotes the number of true negative events, $\bar{x}$ denotes estimation average, $\bar{y}$ denotes observation average, x_i denotes pixel estimation, y_i denotes pixel observation, and N denotes amount of observations.

2.2 Overview of Satellite-Based Precipitation Estimation

In this section, I give a brief overview of important previous development on satellite-based precipitation estimation products. In addition, I review in detail one of the most widely used satellite-based precipitation estimation products, Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks - Cloud Classification System (PERSIANN-CCS), since it is used as a baseline model in this research. In this dissertation, the precipitation estimates from the operational PERSIANN-CCS is used for comparison. Appendix A provides details on the PERSIANN-CCS reproduced on the main study region and study periods and its comparison with operational product.

Several operational satellite precipitation estimation products are available for public use through their open websites. The Climate Prediction Center morphing method (CMORPH), developed by the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center, uses precipitation estimates derived from low-orbiter satellite PMW and IR data to transport the PMW precipitation features during periods when microwave data are not available at a given location (Joyce et al., 2004). The Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) blended IR information and PMW estimates and available rain gauge analyses to produce the final product with a calibration traceable to the single "best" satellite estimate (Huffman et al., 2007). Other satellite-based precipitation estimation products include the PMW-calibrated IR algorithm

(PMIR; Kidd et al., 2003), the Precipitation Estimation from Remotely Sensed Self-Calibrating Multivariate Precipitation Retrieval algorithm (SCaMPR; Kuligowski, 2002), and the Naval Research Laboratory Global Blended-Statistical Precipitation Analysis (NRLgeo; Turk and Miller, 2005).

The PERSIANN product takes advantage of machine learning techniques to estimate precipitation rates with features extracted from IR grids and a window of grids surrounding them (Hsu et al., 1997). Similarly, the PERSIANN-CCS, a revised PERSIANN product with finer resolution, also applies artificial neural networks to classify clouds based on IR information and estimate precipitation (Hong et al., 2004). A detailed description of PERSIANN-CCS can be found in Hong et al., 2004. Roughly speaking, there are two steps involved in the process: the unsupervised clustering of clouds using Self-Organizing Feature Map (SOFM) and the precipitation estimation of pixels using the probability-matching method and the exponential regression.

In the unsupervised clustering step, nine features are extracted from the cloud patches, as shown in Table 2-3. The features are designed to describe the temperature, geometric, and textural properties for the clouds and served as inputs to a 20×20 SOFM to cluster the cloud patches (Hong et al., 2004). One source of the bias in satellite-based precipitation estimation is that a few statistics of an IR image, such as mean and standard deviation of

nearby pixels, do not provide as much information as the raw image itself, where the full cloud-shape information is contained.

Table 2-3 Cloud features used for PERSIANN-CCS

Coldness Features	Minimum temperature* of a cloud patch Mean temperature of a cloud patch
Geometric Features	Cloud-patch area Cloud-patch shape index
Texture Features	Standard deviation of cloud-patch temperature Mean value of local standard deviation of cloud temperature Standard deviation of local standard deviation of cloud Gradient of cloud-top temperature Gray-image texture (maximum angular second moment)

* Temperature refers to the cloud-top brightness temperature provided by the satellite imagery.

After the clusters are developed, the probability-matching method is applied to each cluster respectively. In each cluster, the cloud-top brightness temperatures and the observed rain rates of the pixels are first redistributed separately and then re-matched with the assumption that pixels with lower cloud-top brightness temperature associate with higher rain rates. Then, an exponential regression is used to fit this relationship. One advantage of this process is that it preserves the ability of estimating extreme but rare events, which is substantial for the use of the precipitation values. However, the critical assumptions used

in this step are not necessarily true, since heavy rain rates may occur at the edges of a cloud patch, where the brightness is often not lower than that inside the patch.

Various validation studies have been conducted to address the errors in satellite-based precipitation estimation products and to investigate potential approaches to improve the algorithms (AghaKouchak et al., 2011; Bellerby and Sun, 2005; Moazami et al., ,2014; Tian et al., 2009). Overestimation with many false alarms is identified as a common drawback for most satellite-based precipitation estimation products, especially in warm seasons (Sapiano and Arkin, 2009). In addition, precipitation from warm clouds is often missed in satellite-based products (Behrangi et al., 2009a; Nasrollahi et al., 2013).

2.3 Overview of Deep Neural Networks

Deep learning techniques are a new version of artificial neural networks, breaking the traditional limitation of the number of hidden layers and hidden nodes. The deep learning algorithms aim at learning feature hierarchies with features from higher levels of the hierarchy formed by the composition of lower-level features (Bengio, 2009). In other words, instead of depending on human-crafted features, DNNs are capable of automatically extracting useful features from raw information, especially for image data. Additionally, such models are designed to learn complex functions that directly map the original input to

the output. This deep architecture corresponds to the mammalian brain, where an image is processed in sequence of stages: detection of edges, primitive shapes, and moving up to gradually more complex visual shapes (Bengio, 2009).

Successful use of a DNN depends largely on how well to optimize the parameters (weights) of neural nodes and calibrate a neural network with given data samples. In general, the parameter values are established through supervised or unsupervised training, which the former requires data samples of input-output pairs and an effective search program to calculate the parameter values, while the latter only needs input data samples and a self-organization program to set the parameters. Because of the large number of parameters in a DNN, the traditional supervised training process could use a large amount of time and may not successfully optimize the parameters from a random initialization (Glorot and Bengio, 2010). To obtain better estimates, unsupervised pre-training techniques were developed in recent years to initialize the weights to capture information in the input data (Bengio et al., 2007; Bengio, 2009; Vincent et al., 2008). Afterwards, the traditional supervised gradient-based backpropagation is applied to fine-tune the whole neural network (Bengio, 2009; Rumelhart et al., 1986; Vincent et al., 2008). A detailed review of deep learning methodologies and their applications can be found in Bengio (2009).

Specifically, SDAEs, widely used deep learning technique introduced by Vincent et al. (2008, 2010), are used here. SDAEs have been proven to effectively construct high-level representations from image patches (Glorot et al., 2011b; Lu et al., 2014; Netzer et al., 2011; Vincent et al., 2010; Wang and Yeung, 2013; Xie et al., 2012; Zhou and Lee, 2012). The method mainly involves two steps: (1) unsupervised feature extraction pre-training; and (2) supervised neural network fine-tuning.

The first step of the SDAEs is the unsupervised feature extraction process. The structure of neural networks for the process is an Auto-Encoder (AE), as shown in Figure 1. It uses a set of weights to convert the input vector $\boldsymbol{x}$ into an internal representation (the hidden nodes) $\boldsymbol{h}$ and then another set of weights to produce $\hat{\boldsymbol{x}}$, the reconstructed estimate of $\boldsymbol{x}$ (Hinton and Zemel, 1993). The process can be expressed by

$$\boldsymbol{h} = f(\boldsymbol{x}) = s(\boldsymbol{W_{12}}\boldsymbol{x} + \boldsymbol{b_1})$$

and

$$\hat{\boldsymbol{x}} = g(\boldsymbol{h}) = s(\boldsymbol{W_{23}}\boldsymbol{h} + \boldsymbol{b_2})$$

where $f(\cdot)$ and $g(\cdot)$ are deterministic mappings and s is a non-linear operator, known as the activation function; $\boldsymbol{W_{12}}$ denotes the weights between the input and hidden layers; $\boldsymbol{W_{23}}$ denotes the weights between the hidden and output (reconstruction) layers; and $\boldsymbol{b_1}$ and $\boldsymbol{b_2}$ are offset vectors. Common choices of activation functions include sigmoid,

hyperbolic tangent, and rectified linear functions. In this study, I chose the rectified linear function, since it has been proven to be effective for deep architectures, especially in regression problems (Zeiler et al., 2013). When used as a pre-training process for a DNN, the representations are commonly designed to be over-complete (**h** has higher dimension than **x**) but sparse (Le et al., 2011; Ranzato et al., 2007).

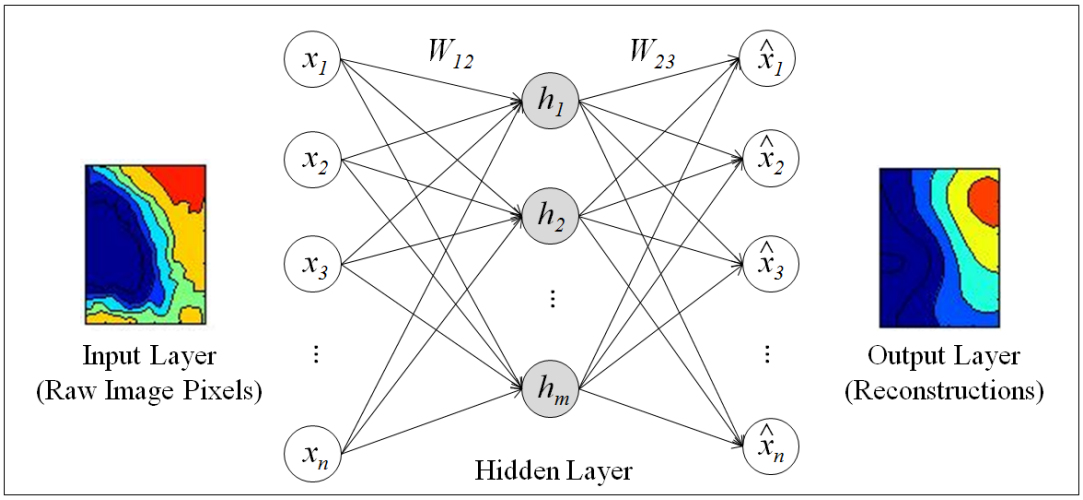

Figure 2-1 Structure of an Auto-Encoder (AE), which reconstructs the input information by learning the internal representations (hidden layer)

To extract sparse yet informative features, rather than "simply copy the input," a DAE is designed based on the idea that robust representations can reconstruct the input from a corrupted version of it (Vincent et al., 2010). In other words, the input layer of a DAE is the raw data with extra noise. It learns the hidden representations by "de-noising" it to a clean version. There are multiple ways to do the corruption, as suggested in (Vincent et al., 2010). In this study, I used "masking noise," which randomly forces a fraction of the input

elements to be 0. When DAEs are combined in a stacked fashion, it is called SDAEs (Hossain et al., 2015; Vincent et al., 2010). Starting from the input layer, a DAE is applied to initialize weights between it and the first hidden layer. Afterwards, the trained first hidden layer is taken as the input layer to the next DAE. The process runs sequentially until it reaches the last hidden layer.

The unsupervised layer-wise pre-training is computationally efficient but does not provide the optimal weights for the overall prediction task (Hinton et al., 2006). Thus, the fine-tuning step is applied to optimize the whole neural network by updating all the weights jointed using the labeled data. In this step, uncorrupted data is used as the input (Vincent et al., 2010).

In this dissertation, for the classification and regression problems, the activation functions used are sigmoid and rectified linear functions, respectively. The activation functions remain the same for all internal nodes. In pre-training, the masking fraction is 40%, where masked pixels are randomly selected. The objective function used in pre-training is the mean squared error between the raw and reconstructed images. In both steps, we used mini-batch gradient descent optimization algorithm with mini-batch size of 10,000. The learning rates for the classification and regression are 0.001 and 0.0001, respectively. These parameters are used in the rest of this research, unless otherwise stated. This work was implemented in Matlab.

2.4 Data Used and Study Region

The data used in this research incorporates satellite imageries, ground observations, and precipitation estimates from operational products. The satellite imageries are the inputs of the precipitation estimation model. The IR (10.8 μm) channel from the Geostationary Operational Environmental Satellite (GOES), capturing cloud-top brightness temperature, is the primary input in this research and most operational products. In addition, WV (6.7 μm) channel is used as supplementary inputs to IR data, since they are proven to be helpful for precipitation identification and estimation in conjunction with IR data (Ba and Gruber, 2001; Behrangi et al., 2009a; Behrangi et al., 2009b; Kurino, 1997). More specifically, the conversion of water vapor is a necessary condition in precipitation formation. By adding satellite precipitable water measurement, which measures the vertical integral of water vapor, the performances of precipitation estimation products are expected to improve. Specifically, both channels' data are processed to the hourly scale in this research. The Stage IV at the same spatial and temporal resolution serves as ground observations. As mentioned in Section 2.2, PERSIANN-CCS serves as a baseline model in the performance comparisons.

Table 2-4 Basic statistics of PERSIANN-CCS and Stage IV for the warm seasons

		June–August 2012	**June–August 2013**
Ave. Precip. (mm/hr)	PERSIANN-CCS	0.100	0.185
	Stage IV	0.077	0.116
Max. Precip. (mm/hr)	PERSIANN-CCS	55.0	75.2
	Stage IV	85.1	112.0

Table 2-5 Basic statistics of PERSIANN-CCS and Stage IV for the cold seasons

		December 2012–February 2013	**December 2013–February 2014**
Ave. Precip. (mm/hr)	PERSIANN-CCS	0.104	0.100
	Stage IV	0.080	0.040
Max. Precip. (mm/hr)	PERSIANN-CCS	51.2	62.56
	Stage IV	48.4	34.1

To properly validate the methodologies, I divided the data into training and verification datasets. This structure remains the same for the following sections. During the training process, the training datasets are used to calibrate the parameters and prevent overfitting. More specifically, the model is trained with 75% of the training data, and measures are calculated with the 25% hold-out data to determine the performances of the different

combination of hyper-parameters. The data of the summer and winter seasons of 2012–2013 are used to calibrate the models, and the data of the next year serve as verification data. I did not distinguish between the seasons for the models, because the potential uses of such models are global. In other words, such climate differences are left for the models to detect based on the satellite information.

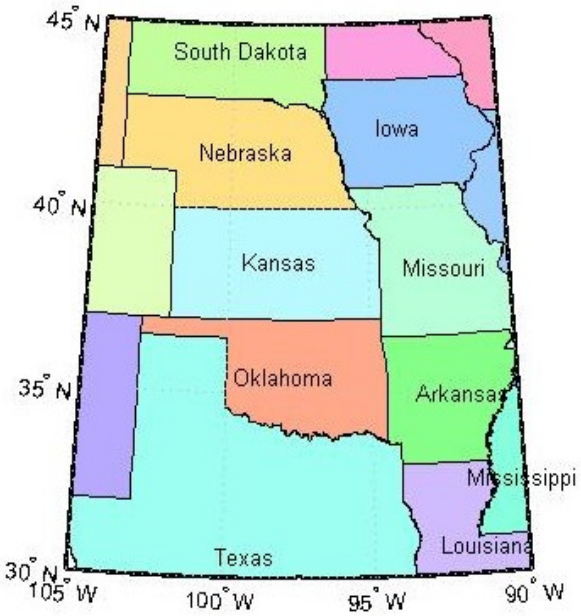

Figure 2-2 Map of the main study region, the central United States

Chapter 3

Exploration of Deep Neural Networks and Bias Reduction of Satellite-Based Precipitation Estimation

3.1 Introduction

To develop a deep learning framework for satellite-based precipitation estimation, the first step is to demonstrate of the effectiveness of DNNs in the application of satellite imageries. More specifically, I want to explore the DNNs' capability of extracting the useful precipitation information buried in the satellite imageries. In this Chapter, I develop a bias reduction model for satellite-based precipitation estimation products using DNNs without incorporating additional input-data sources.

As discussed in Section 2.2, the accuracy of current satellite-based products remains insufficient, despite the efforts of linking multiple satellite information to surface precipitation in various ways. To manage this problem, a variety of bias correction methods have been developed, mainly by incorporating additional available datasets, such as rain gauge or radar information (Boushaki et al., 2009; McCollum et al., 2002). However,

ground-based measurements are only available in specific regions with a sufficient number of instruments. Therefore, several proposed bias correction methodologies are limited to a regional scale and are very difficult to extend to global applications. On the other hand, research also requires more satellite datasets to help reduce biases in the products. For instance, Behrangi et al. (2009) used multi-spectral data from GOES and proved their effectiveness in precipitation detection. Li et al. (2007) and Nasrollahi et al. (2013) also showed the value of the moderate resolution imaging spectroradiometer (MODIS) in identifying high clouds and thus reducing false alarms.

In this section, I reduce the bias of satellite-based precipitation estimation products with a focus on taking advantage of advanced machine learning algorithms instead of incorporating additional observations. The power of deep learning techniques for image processing and feature extraction provides an opportunity to improve the accuracy of satellite-based precipitation estimation. As preliminary DNN exploration, I aim to perform the following tasks: (1) develop a bias correction system focusing on overestimation and false alarms, with a case study on the PERSIANN-CCS product; (2) demonstrate the effectiveness of deep learning for precipitation-related information extraction from satellite infrared imagery without adding any extra data from other sources; and (3) evaluate and analyze the case study results in the summer and winter seasons, respectively.

The remainder of this section is organized as follows. Section 3.2 describes the detailed methodology and model setup for this study. Section 3.3 presents a comparison between the output of this study and the original satellite product. Finally, the main conclusions are summarized in Section 3.4.

3.2 Methodology and Model Setup

The design of the process is presented in Figure 3-1. In this study, the input data for the DNN is IR imagery, which is the same raw information used by PERSIANN-CCS. Instead of using cloud image features designed by researchers, as used in PERSIANN-CCS, I allow the neural network to extract a useful representation for precipitation estimation by itself. As shown in Figure 3-1, the input to DNN is a matrix $T_{15\times15}$ containing the IR image in a 15 × 15 pixel window centered in pixel $t_{8,8}$, at which PERSIANN-CCS indicates a positive precipitation rate r_{p}. To produce a training data pool, the window is moved across the image sequentially, shifting the location one grid-box at a time in each hourly IR image of the study region.

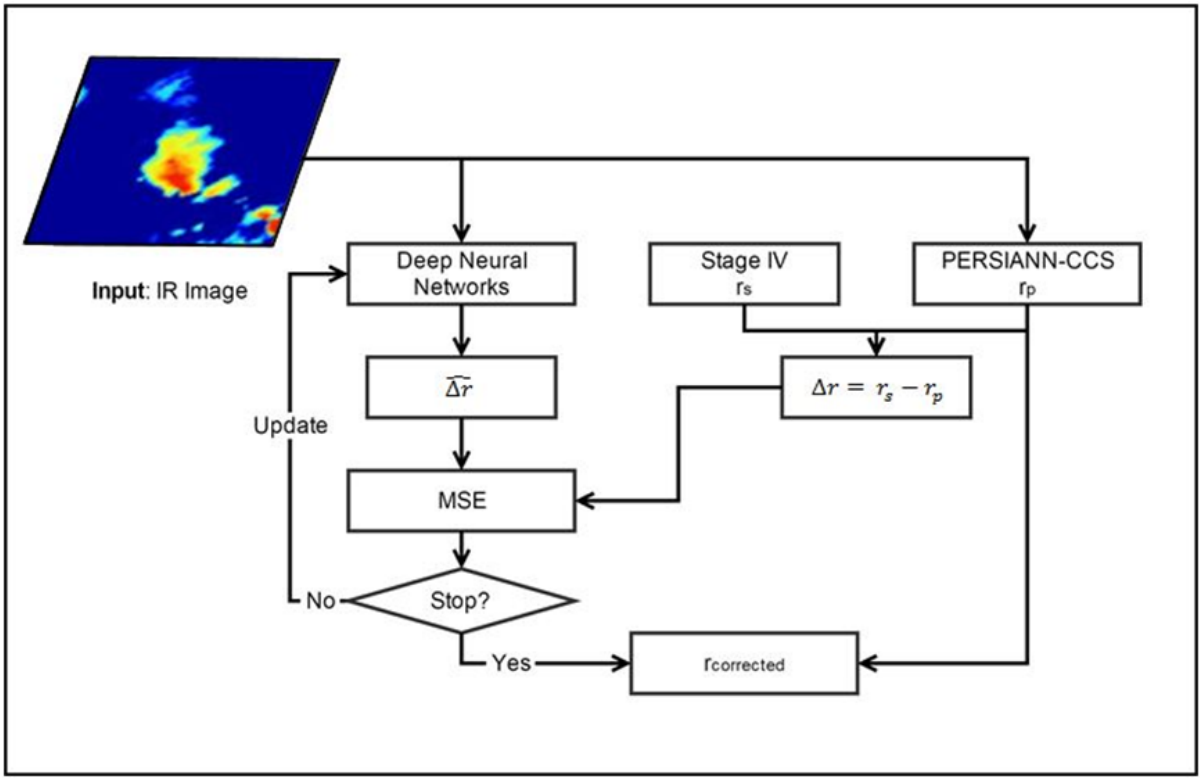

Figure 3-1 Experimental design process: the input to the neural network is the IR image, and the output is the difference between the PERSIANN-CCS estimates and the Stage IV measurements

The outputs/targets (at the same spatiotemporal resolutions as the input data) are the differences between the PERSIANN-CCS estimates and the Stage IV observations. The output is the value of the centered pixel of the 15 × 15 pixel window. In other words, it is the adjusted quantity needed for the PERSIANN-CCS estimate to match the Stage IV observed rate (r_s) at pixel $t_{8,8}$ ($\Delta r = r_p - r_s$). The reason I choose the differences, instead of directly estimating the Stage IV precipitation rates, is that PERSIANN-CCS, as well as other satellite-based products, has successfully screened out a large number of no-rain pixels

(Hong et al., 2007). Therefore, the input data are much more balanced and, thus, benefit the training process. However, the disadvantage of this design is that it does not help reduce the missing cases in PERSIANN-CCS. In addition, both inputs and outputs of the training data are normalized before training to shrink the range of the quantity and make it easier to operate.

After being properly trained, the DNN produces $\widehat{\Delta r}$ (estimated difference) given the IR imagery without information from Stage IV and PERSIANN-CCS. With this property, the model can be potentially applied to areas without Stage IV information and thus can offer global bias correction. On the other hand, the use of Δr (real difference) during the training process, which optimizes MSE between $\widehat{\Delta r}$ and Δr, is shown in Figure 3-1. Lastly, when producing adjusted precipitation, all negative values are forced to zero.

In addition to the main study region introduced in Section 2.4, to demonstrate the effectiveness of the methodology with variability in space, three extra regions with Stage IV measurements are selected for additional verification. These regions are (1) the Colorado area (35–40°N, 105–110°W), (2) the Arizona area (30–35°N, 110–115°W), and (3) the Georgia area (30–35°N, 80–85°W), which the maps of the regions are shown in Figure 3-2.

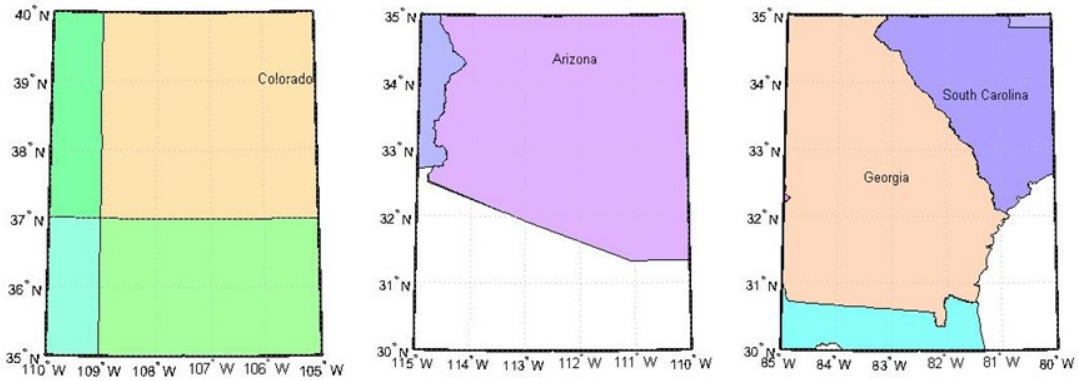

Figure 3-2 Maps of selected additional validation regions

The objective function of the SDAEs is to obtain the optimal weight values of MSE on the output. In addition, in this study, I use the rectified linear activation function, which is the most popular choice for real-value estimation (Glorot et al., 2011b). After various combinations were tested and compared, a 4-layer neural network with 1,000 hidden nodes for each hidden layer with 40% input corruption in the training was selected for this study. Table 3-1 gives some common hyper-parameters needed for training a DNN with SDAEs. The choices considered in this study are typical choices for the corresponding parameters (Vincent et al., 2010). In addition, I only considered neural networks with an equal number of hidden nodes at all hidden layers for processing convenience. The result should not be fundamentally different from other possible combinations. Other potential hyper-parameters, such as learning rate and training iterations, were decided manually to optimize the result within the training process.

Table 3-1 Hyper-parameters considered for the SDAEs in this section

Hyper Parameter	Description	Considered Value
nHLayers	Number of hidden layers	{1, 2, 3}
nHNodes	Number of hidden nodes	{500, 1,000, 2,000}
pCorruption	Percentage of corruption for pre-training	{0.1, 0.2, 0.4}

3.3 Results and Discussion

The results presented here show the performances of the SDAEs model in the verification periods (summer 2013 and winter 2013–2014) in comparison with the original PERSIANN-CCS data. The evaluation includes detection of R/NR pixels and intensity of the precipitation amount for warm and cold seasons. In addition, as an example, results of the rainfall event on August 4, 2014 are analyzed and compared with both PERSIANN-CCS estimation and Stage IV observation. In this section, "DNN-corrected" refers to the bias-corrected precipitation using SDAEs.

3.3.1 Rain/No-Rain Classification

Table 3-2 provides the binary R/NR detection performance of PERSIANN-CCS and DNN-corrected precipitation, including the average hourly number of precipitation pixels, false positive pixels, and misclassified pixels. The performance is evaluated for hourly

estimation on the study area and averaged over the validation periods for warm and cold seasons separately. The bias correction process is very effective at identifying false alarm pixels and balancing the number of precipitation pixels. Specifically, the average hourly number of false alarm pixels drops from 395 to 264 and from 598 to 339 for summer and winter (i.e., 33% and 43% correction), respectively. The model also properly reduced the overestimation of the number of pixels with precipitation in PERSIANN-CCS (30% more to 3% less and 90% more to 14% more relative to Stage IV observations).

The overall number of misclassified pixels is reduced for both warm and cold seasons (i.e., 13% and 28% correction), respectively. The Heidke skill score of DNN-corrected is similar to PERSIANN-CCS in summer and slightly better in winter. The model's incapability of dealing with missing cases may prevent it from improving the score and thus shows the necessity of moving on to a DNN that directly estimates precipitation from IR imagery. Moreover, the frequency bias shows that the forecast biases are reduced for both seasons compared to PERSIANN-CCS. This suggests that the model can identify false alarm pixels in the original PERSIANN-CCS.

*Table 3-2 R/NR classification performance of PERSIANN-CCS and DNN-corrected precipitation**

Season		**Summer**	**Winter**
Number of precipitation pixels (/hour)	Stage IV	533	358
	PERSIANN-CCS	694 (30% more)	681 (90% more)
	DNN-corrected	516 (3% less)	409 (14% more)
Number of false positive pixels (/hour)	PERSIANN-CCS	395	598
	DNN-corrected	264 (33% corrected)	339 (43% corrected)
Number of misclassified pixels (/hour)**	PERSIANN-CCS	629	873
	DNN-corrected	545 (13% corrected)	627 (28% corrected)
Heidke Skill Score	PERSIANN-CCS	0.478	0.148
	DNN-corrected	0.473	0.174
Frequency Bias	PERSIANN-CCS	1.27	1.86
	DNN-corrected	0.93	1.09

* Values in the parenthesis are the relative performance of DNN-corrected and PERSIANN-CCS. The same definition applies to all the following tables in this section.

** Number of misclassified pixels includes both false positive (false alarm) and false negative (missing) events.

3.3.2 Precipitation Amount

Table 3-3 Average bias, variance, and MSE of PERSIANN-CCS and DNN-corrected precipitation

Season		**Summer**	**Winter**
Average bias	PERSIANN-CCS	0.091	0.054
	DNN-corrected	0.002 (98% corrected)	0.012 (78% corrected)
Average variance (mm/hr)^2	PERSIANN-CCS	2.330	0.442
	DNN-corrected	1.596 (31% corrected)	0.306 (31% corrected)
Average MSE (mm/hr)^2	PERSIANN-CCS	2.338	0.445
	DNN-corrected	1.596 (32% corrected)	0.306 (31% corrected)

Figure 3-3 presents maps of the bias of PERSIANN-CCS and DNN-corrected precipitation over the study region averaged on the warm and cold verification periods, respectively. The white color indicates very small bias and shows that the DNN model has made relatively significant corrections to the PERSIANN-CCS precipitation pixels, especially in the summer season, and the overestimation produced by the PERSIANN-CCS product is mostly removed, which specific calculations are displayed in Table 3-3. The average biases are only 0.002 mm/day and 0.012 mm/day after bias correction, compared to 0.091

mm/day and 0.054 mm/day before bias correction for summer and winter (98% and 78% correction), respectively.

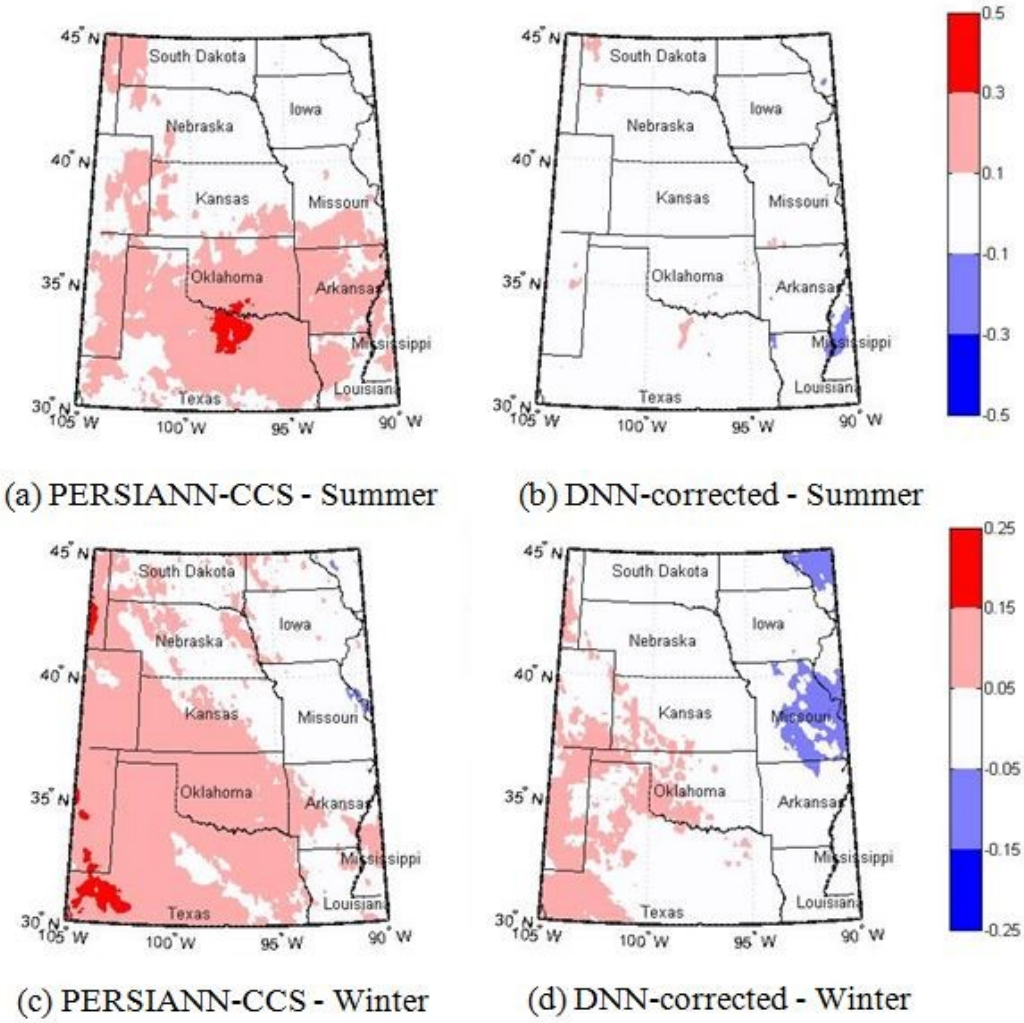

Figure 3-3 Average Bias (mm/hr) of PERSIANN-CCS [(a), (c)] and DNN-corrected [(b), (d)] output over the central United States (30–45N, 90–105°W): (a), (b) summer (June–August 2013); (c), (d) winter (December 2013–February 2014)

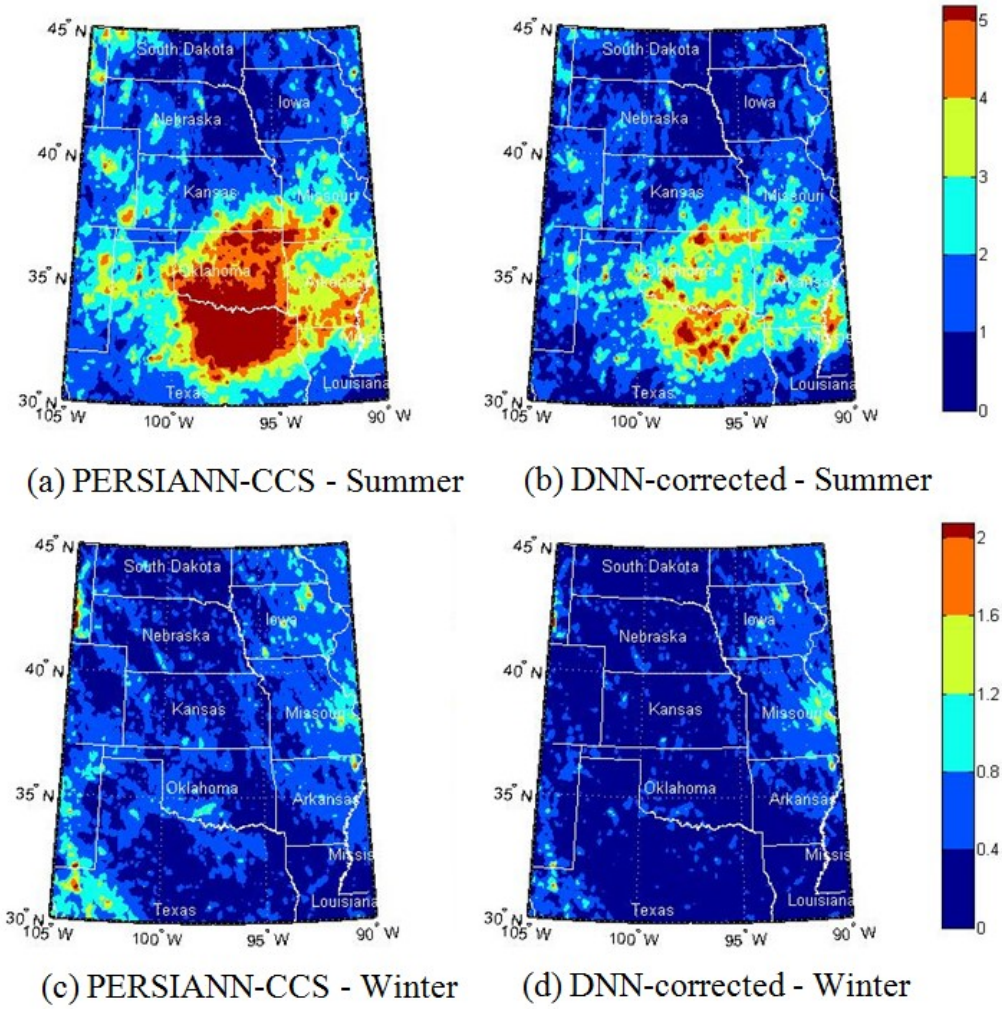

Figure 3-4 Average MSE ([mm/hr]^2) of PERSIANN-CCS [(a), (c)] and DNN-corrected [(b), (d)] output over the central United States (30–45°N, 90–105°W): (a), (b) summer (June–August 2013); (c), (d) winter (December 2013–February 2014)

Similar results can be seen in Figure 3-4, which shows the MSE of PERSIANN-CCS and DNN-corrected precipitation over the study region averaged over the verification periods. The warm colors indicate strong differences compared to Stage IV observations, while the cold colors indicate small differences. The heavy errors shown in PERSIANN-CCS over the

summer verification period (Figure 3-4(a)) are strongly reduced by the model (Figure 3-4(b)). Similar results can also be observed for the winter period (Figure 3-4(c, d)). However, as Table 3-3 shows, over 30% correction in average MSE is observed for both seasons, and the absolute improvement in summer is more significant. The results indicate the model's ability to correct the bias of the overall precipitation intensity for both warm and cold seasons by automatically extracting useful features from satellite data.

3.3.3 Case Studies

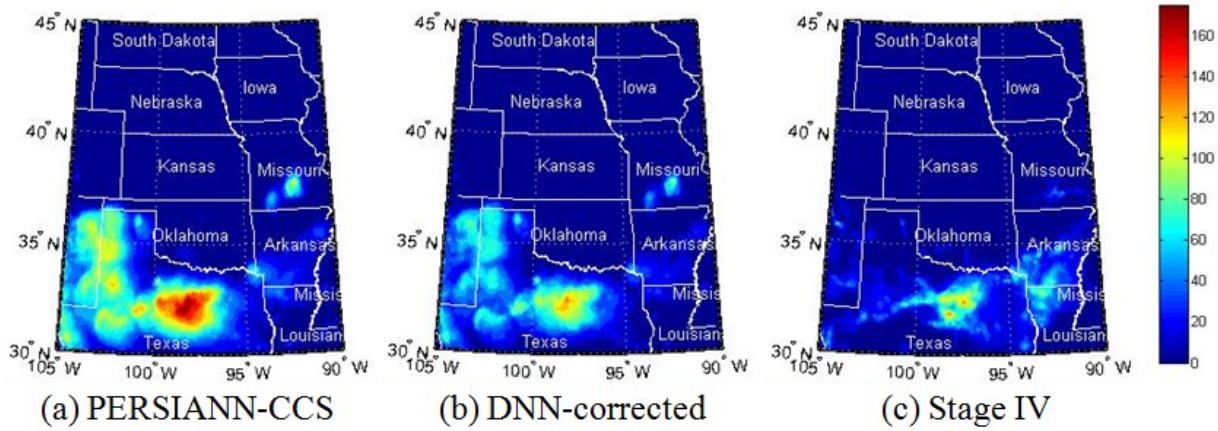

Figure 3-5 Cumulative precipitation amounts (mm/day) of PERSIANN-CCS estimation, DNN-corrected estimation, and Stage IV observation on August 4, 2014, over the central United States (30–45°N, 90–105°W): (a) PERSIANN-CCS; (b) DNN-corrected; and (c) Stage IV

Table 3-4 Performance of PERSIANN-CCS and DNN-corrected on August 4, 2014

Number of precipitation pixels (/hour)	Stage IV	1,433
	PERSIANN-CCS	1,732 (22% more)
	DNN-corrected	1,450 (2% more)
Number of false positive pixels (/hour)	PERSIANN-CCS	805
	DNN-corrected	622 (23% corrected)
Number of misclassified pixels (/hour)	PERSIANN-CCS	1,311
	DNN-corrected	1,227 (6% corrected)
Average bias (mm/hr)	PERSIANN-CCS	0.398
	DNN-corrected	0.164 (58% corrected)
Average MSE $(mm/hr)^2$	PERSIANN-CCS	8.267
	DNN-corrected	4.875 (41% corrected)

To demonstrate how SDAEs have significantly improved the estimates of individual precipitation events, Figure 3-5 and Table 3-4 present the analysis of a rainfall event on August 4, 2014. The event is randomly selected from noticeable rainfall events within the validation periods. The cumulative amounts of the rainfall event for PERSIANN-CCS, DNN-corrected, and Stage IV precipitation are all displayed in Figure 3-5. The figure shows that overestimation in the original PERSIANN-CCS is reduced remarkably, while the rainfall distribution pattern is also adjusted towards the observation to some extent. This effect is quantified in Table 3-4 for both R/NR detection and intensity. As for detection performance, the number of precipitation pixels in PERSIANN-CCS is reduced from 22%

overestimated to just 2%, while around 23% of false positive pixels are corrected. As for intensity, average bias and MSE decrease from 0.398 and 8.267 to 0.164 and 4.875 (i.e., 58% and 41% correction), respectively. This example demonstrates the effectiveness of the model in improving the precipitation estimation for typical storm events. Meanwhile, the scheme is unable to deal with the missing precipitation of the original PERSIANN-CCS. An area of future research for us will be to apply the method to direct precipitation estimation to start addressing this issue.

To validate the effectiveness of the methodology when the coefficients are applied in other locations, Table 3-5 summarizes average bias, variance, and MSE of DNN-corrected and PERSIANN-CCS precipitation over three areas outside of the study region on the warm and cold validation periods, respectively.

Generally, the model works effectively to reduce bias and variance of the original PERSIANN-CCS. For the Colorado area (35–40°N, 105–110°W) and the Arizona area (30–35°N, 110–115°W), MSEs are improved at least 29% for both warm and cold seasons, while improvement is only around for 4% for summer for the Georgia area (30–35°N, 80–85°W). One possible reason for this is that the original PERSIANN-CCS has a relatively large amount of missing and underestimation in this area. Therefore, our model is not helpful in those situations, as discussed above.

Table 3-5 Average bias, variance and MSE of PERSIANN-CCS and DNN-corrected precipitation over areas outside of the study region

Season			**Summer**	**Winter**
Colorado Area	Average bias (mm/hr)	PERSIANN-CCS	0.091	0.062
		DNN-corrected	0.002 (99% corrected)	-0.005 (92% corrected)
	Average variance $(mm/hr)^2$	PERSIANN-CCS	0.738	0.722
		DNN-corrected	0.498 (32% corrected)	0.421 (42% corrected)
	Average MSE $(mm/hr)^2$	PERSIANN-CCS	0.746	0.725
		DNN-corrected	0.498 (33% corrected)	0.421 (42% corrected)
Arizona Area	Average bias (mm/hr)	PERSIANN-CCS	0.028	0.085
		DNN-corrected	0.004 (86% corrected)	0.002 (98% corrected)
	Average variance $(mm/hr)^2$	PERSIANN-CCS	0.331	0.778
		DNN-corrected	0.235 (29% corrected)	0.437 (44% corrected)
	Average MSE $(mm/hr)^2$	PERSIANN-CCS	0.332	0.785
		DNN-corrected	0.235 (29% corrected)	0.437 (44% corrected)
Georgia Area	Average bias (mm/hr)	PERSIANN-CCS	-0.039	0.017
		DNN-corrected	-0.052 (33% worse)	-0.016 (6% corrected)
	Average variance $(mm/hr)^2$	PERSIANN-CCS	0.911	0.862
		DNN-corrected	0.875 (4% corrected)	0.655 (24% corrected)
	Average MSE $(mm/hr)^2$	PERSIANN-CCS	0.911	0.862
		DNN-corrected	0.878 (4% corrected)	0.657 (24% corrected)

3.4 Conclusion

The aim of this chapter is to apply a DNN framework to satellite-based precipitation estimation products to correct the estimation bias in a data-driven manner by extracting more useful features from satellite imagery. More specifically, SDAEs, a popular technique in image recognition, are employed to improve the PERSIANN-CCS product. The model is trained in 2012–2013 and evaluated during the 2013–2014 summer and winter seasons.

Verification studies show improved results in both R/NR detection and precipitation intensity over the validation period for both seasons. Binary R/NR detection resulted in the correction of a significant number of false alarm pixels, especially in the cold season. For precipitation intensity, the average daily biases are corrected by as much as 98% and 78% in the validation warm and cold seasons, respectively. These results are also illustrated for a specific rainfall event on August 4, 2014, for which visualization of the cumulative rainfall amount demonstrates the model's ability to correct false alarms and overestimation.

The results verify that useful information is available in IR imagery and can help improve the quality of satellite precipitation estimation products with respect to detecting R/NR pixels and quantifying the precipitation rates. More importantly, such useful information for precipitation estimation can be extracted automatically by DNNs. Moreover, the methodology can be easily integrated into near-real-time operational precipitation estimation products and help extract additional features from satellite datasets to reduce

bias. Meanwhile, the application of the technique is not limited to IR imagery, but should be extendable to multiple satellite datasets due to its ability to automatically extract information. The case study of PERSIANN-CCS proves its advantage compared to a few manually designed features.

In addition, our results suggest that GOES cloud IR imagery still contains valuable information that has not been utilized by most satellite-based precipitation retrieval algorithms. Our experiment demonstrates that the cloud IR image from a 15 × 15 pixel window is more informative than the nine IR statistic features used in PERSIANN-CCS as the input data for precipitation estimation. Such information can be extracted automatically by a well-designed DNN.

A natural next step for this work is to explore the possibility of using deep learning techniques to produce a precipitation estimation product directly instead of using it in a bias correction manner. In the next two chapters, I introduce a carefully designed two-stage DNN structure to fulfill the needs of direct precipitation estimation.

Chapter 4

Two-Stage Model Part I: Rain/No-Rain Classification

4.1 Introduction

Chapter 3 demonstrated the DNN models' capability of extracting additional useful features from satellite information for precipitation estimation. The next goal is to design a structure to properly use such models to develop a precipitation estimation product. As discussed in Section 2.1, the imbalance nature of precipitation and multi-objective needs of precipitation estimation products make it difficult to directly apply of data-driven methods. To fulfill these needs, a two-stage model is considered. The first stage is a binary R/NR classification model, which focuses on correctly identifying pixels with precipitation and eliminating the massive amount of no precipitation pixels. The second stage is a precipitation amount regression model, which focus on estimating precipitation amount accurately by properly adjusting the objective function.

On the other hand, as discussed in Section 1.1, satellite information, in addition to IR imageries, is helpful for precipitation estimation. Various previous studies have demonstrated that additional satellite channels may help better estimate precipitation by providing supplementary cloud characteristics or delineating the precipitation regions (Arking and Childs, 1985; Inoue and Aonashi, 2000; Lensky and Rosenfeld, 2003; Levizzani and Setva´k, 1996; Pilewskie and Twomey, 1987; Rosenfeld and Gutman, 1994; Turk and Miller, 2005; Tsonis et al., 1985). Most common choices include VIS and WV data from GEO satellites and PMW from LEO satellites (Ba and Gruber, 2001; Behrangi et al., 2010; Capacci and Conway, 2005; Hsu et al., 1999; Hong et al., 2004; Hsu et al., 1997; Huffman et al., 2007; Joyce et al., 2004; Kidd et al., 2003; Kuligowski, 2002). PMW is recognized as direct and reliable precipitation measurements from space (Adler et al., 2001; Ebert et al., 1996). VIS channel data measures cloud albedo and is often considered effective in assisting precipitation estimation in daytime (Behrangi et al., 2010; Capacci and Conway, 2005; Cheng et al., 1993; Grassotti and Garand, 1994; Griffith et al., 1978; Hsu et al., 1999; Lovejoy and Austin, 1979; O'Sullivan et al., 1990; Tsonis and Isaac, 1985). In this research, I used WV channel data as the supplementary data source for the model because of its availability in high-spatiotemporal resolution. More specifically, compared to WV channel data, PMW data suffer from low-temporal resolution, while VIS channel data is unavailable during night time and may cause consistency issues for the model (Marzano et al., 2004).

In this and the next chapters, I explore the application of deep learning techniques to precipitation estimation with bispectral information (IR and WV channels). In this chapter, I focus on R/NR identification because accurate precipitation areal delineation is essential for a precipitation estimation product and helps eliminate an extensive amount of NR pixels. Specifically, I aim to perform the following tasks: (1) design a DNN that can manage satellite imageries from multiple channels; (2) demonstrate the effectiveness of the methodology on precipitation identification by comparing its performance with an operational product, PERSIANN-CCS; (3) evaluate the value of adding data sources in addition to IR imageries; and (4) evaluate and analyze the case study results in the summer and winter seasons, respectively.

The remainder of this section is organized as follows. Section 4.2 describes the detailed methodology and model setup for the binary R/NR classification part of the two-stage model. In Section 4.3, the model is verified by comparisons between the outputs of this model using only IR information, this model using both IR and WV information, and the estimates from PERSIANN-CCS product. Finally, the main conclusions are summarized in Section 4.4.

4.2 Methodology and Model Setup

The overview of the experimental design is presented in Figure 4-1. In this study, two models are built to explore the effectiveness of the deep learning approaches and the additional information provided by the WV channel. Specifically, model 1 applies SDAEs to IR patches alone as input and predicts R/NR for the centered pixel. Model 1 (structure shown in Figure 4-2) is referred to as DL-IR only for the remainder of this dissertation. Model 2 is built on the well-calibrated model 1 and incorporates WV imageries as a combined neural network, as shown in Figure 4-3. Hereafter, model 2 is referred to as DL-IR+WV.

As shown in Figure 4-1, for both IR and WV images, the unsupervised feature extraction pre-training step of SDAEs is applied first. For the DL-IR only model (model 1), the supervised fine-tuning step is then applied to connect the IR images to the R/NR information for the centered pixel extracted from the Stage IV data. For the DL-IR+WV model (model 2), I combine the WV imageries and their extracted features into the calibrated DL-IR only model with the design shown in Figure 4-2 and apply the supervised fine-tuning step to it. The input layer includes two patches and connects to the first hidden layer separately to produce their extracted features, respectively. Then, all features act together to eventually produce R/NR predictions. I start with the DL-IR only model to take

advantage of the calibrated parameters, which contain more effective information to transform IR patches to R/NR probability compared to random parameter initialization.

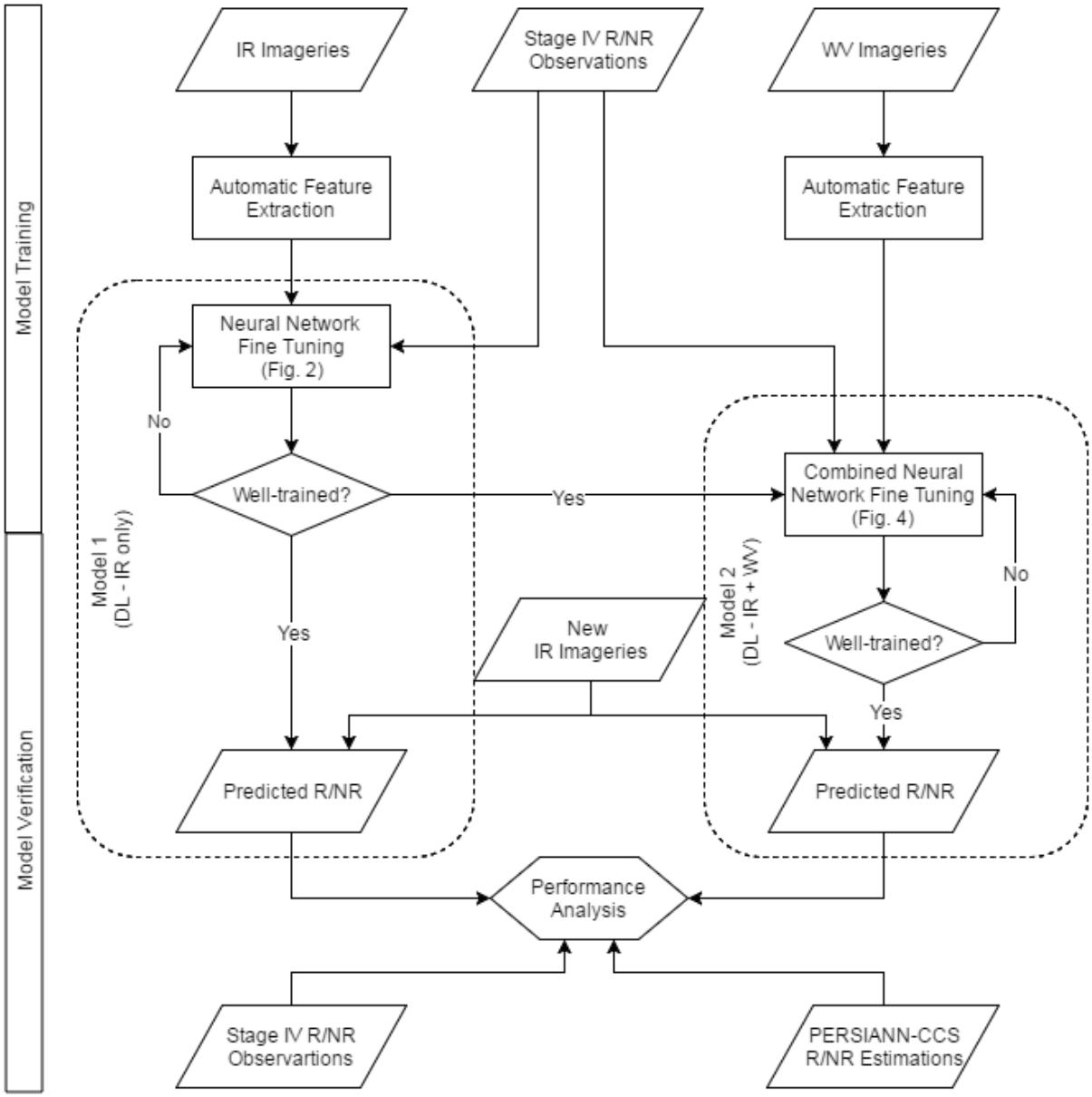

Figure 4-1 Overview of the model training and verification process: Model 1 is a DNN with only IR imageries as input, as shown in Figure 4-2. Model 2 is a combined DNN with both IR and WV imageries as inputs, as shown in Figure 4-3

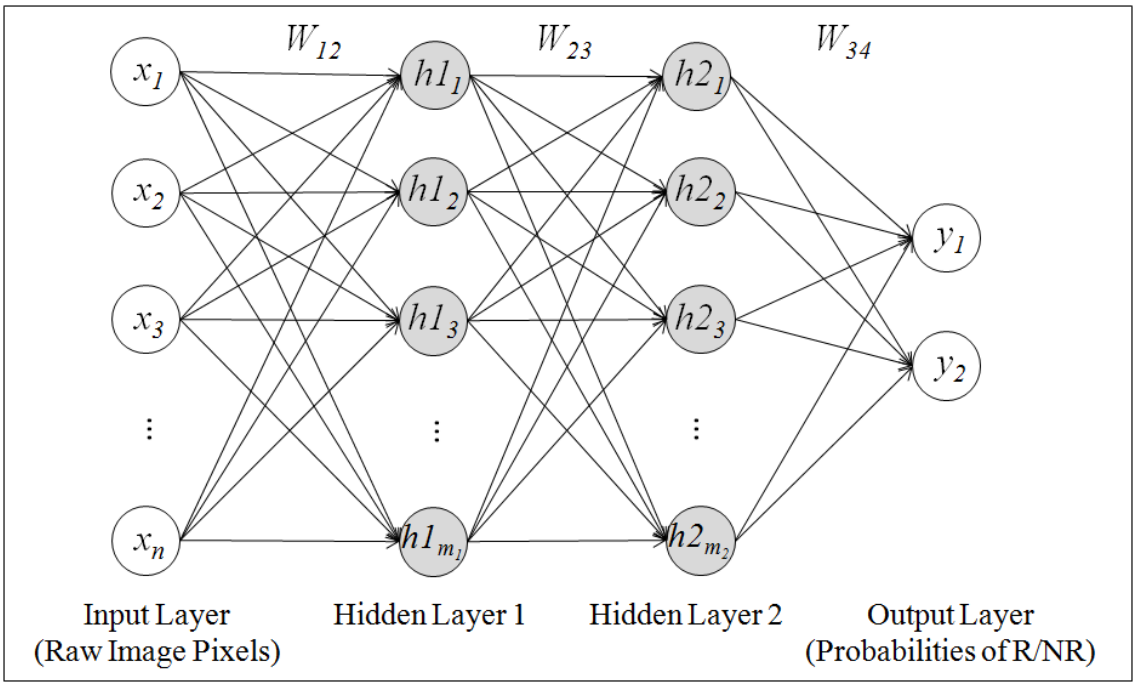

Figure 4-2 Structure of a four-layer, fully connected neural network: the first layer is the input layer, which is the IR patches used in this study; the next two layers are the hidden layers, which can be interpreted as extracted features; the last layer is the output layer, which contains the probabilities of R/NR for the centered pixel

Here, each input patch is a 15 × 15 pixel IR or WV patch, which covers approximately 120 km × 120 km. The patches are selected with overlapping pixels. The number of hidden nodes for the DL-IR only model is 1,000 for both hidden layers. Thus, the number of hidden nodes for the first hidden layer for the DL-IR+WV is 2,000, with 1,000 nodes connected to each input patch. The second layer remains 1,000 hidden nodes. The output layer for both models has two nodes, predicting the probability of R/NR for the centered pixel of the 15 × 15 pixel input patch. These hyper-parameters are selected based on previous study

parameter selection results and experiments with different number combinations (Tao et al., 2016a; Tao et al., 2016b; Vincent et al., 2010).

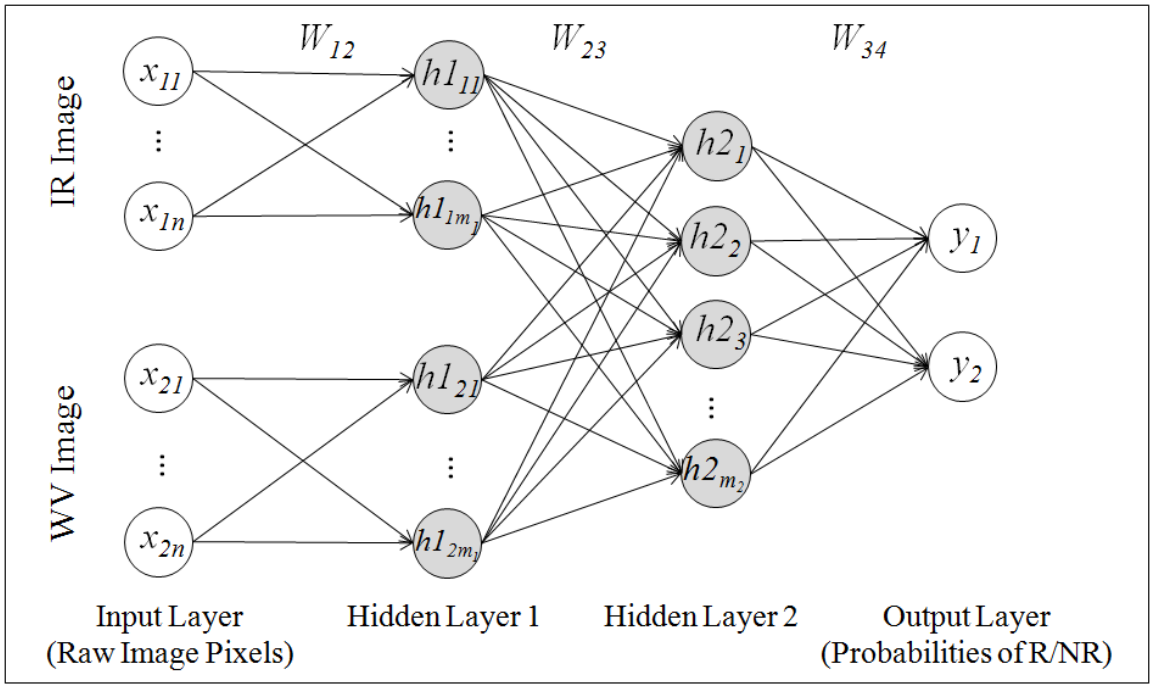

Figure 4-3 Structure of a four-layer neural network with bispectral imageries: the input layer has two images, which connect to half of the hidden nodes in the first hidden layer separately

4.3 Results and Discussion

In this section, I evaluate the performance of the DL-IR model and the DL-IR+WV model to validate the effectiveness of the deep learning techniques and additional information contains in the WV channel. The results of both models are compared with the data of PERSIANN-CCS, a current operational precipitation estimation product introduced in Section 2. The validation measurements I used to evaluate precipitation identification

performance are POD, FAR, CSI, and the performance gain of these measurements with respect to PERSIANN-CCS. The definition of performance gain is presented in Table 4-1. In addition, a few case studies are provided as examples to analyze the precipitation identification performance difference between the models. All evaluations are conducted for the verification periods (winter and summer 2013–2014).

Table 4-1 Definition of performance gain for POD, FAR, and CSI

Verification Measures	**Formulas**	**Range and Desirable Value**
Performance Gain for POD and CSI	$PG = \frac{Val_{model} - Val_{ref}}{Val_{ref}} \times 100\%$	Range: -1 to +∞ Desirable Value: +∞
Performance Gain for FAR	$PG = \frac{Val_{ref} - Val_{model}}{Val_{model}} \times 100\%$	Range: -1 to +∞ Desirable Value: +∞

POD, FAR, and CSI are defined in Table 2-2;

Val_{model} denotes the value of a measurement for a model; Val_{ref} denotes the value of a measurement for a reference.

4.3.1 Rain/No-Rain Classification

Table 4-2 provides the overall performances of PERSIANN-CCS, the DL-IR only model, and the DL-IR+WV model over the verification periods. Both models show improvement in all measurements compared to PERSIANN-CCS. In addition, the DL-IR+WV model has the best performance among them. Specifically, the DL-IR only model and the DL-IR+WV model

have 14.83% and 29.41% performance improvement in POD compared to PERSIANN-CCS, respectively (0.449 and 0.506 compared to 0.391). At the same time, these two models have 9.86% and 20.75% performance improvement in FAR, respectively (0.620 and 0.564 compared to 0.681). These improvements result in significant increases (21.60% and 43.66%) in CSI performance for the models (0.259 and 0.306 compared to 0.213). The improvement of the DL-IR only model, compared to PERSIANN-CCS, demonstrates the effectiveness of the deep learning techniques in automatically extracting useful features for precipitation identification from the raw IR images. The DL-IR+WV model outperforms both PERSIANN-CCS and the DL-IR only model, which shows that the WV channel contains additional information to better support delineating precipitation areas and that the methodology can take advantage of such information.

Table 4-2 Summary of R/NR classification performance over the verification periods (including both summer 2013 and winter 2013–2014)

		PERSIANN-CCS	**DL-IR only (Model 1)**	**DL-IR+WV (Model 2)**
POD	Value	0.391	0.449	0.506
	Performance Gain	-	14.83%	29.41%
FAR	Value	0.681	0.620	0.564
	Performance Gain	-	9.68%	20.74%
CSI	Value	0.213	0.259	0.306
	Performance Gain	-	21.60%	43.66%

Table 4-3 Summary of R/NR classification performance for summer 2013 and winter 2013–2014, respectively

		PERSIANN-CCS	DL-IR only (Model 1)	DL-IR+WV (Model 2)
Summer Season (June–August 2013)				
POD	Value	0.486	0.556	0.608
	Performance Gain	-	14.40%	25.10%
FAR	Value	0.552	0.575	0.554
	Performance Gain	-	-4.00%	-0.36%
CSI	Value	0.304	0.317	0.346
	Performance Gain	-	4.28%	13.82%
Winter Season (December 2013–February 2014)				
POD	Value	0.275	0.323	0.386
	Performance Gain	-	17.45%	40.36%
FAR	Value	0.804	0.687	0.581
	Performance Gain	-	17.03%	38.38%
CSI	Value	0.129	0.189	0.251
	Performance Gain	-	46.51%	94.57%

More detailed performances of the models over different seasons in the verification periods are provided in Table 4-3. Compared to PERSIANN-CCS, the overall improvements shown in CSI are much higher for the winter (46.51% and 94.57% performance gains, respectively) than the summer (4.28% and 13.82% performance gains, respectively) for

both the DL-IR only model and the DL-PR WV model. One reason for these improvements is that the performance of PERSIANN-CSS in the summer (CSI: 0.304) is much better than it is in the winter (CSI: 0.129) for the central United States because the strong convective storms in the summer are relatively easy to detect.

Figures 4-4 and 4-5 present the maps of POD, FAR, and CSI of the models in the warm and cold verification periods, respectively. The warm colors indicate high measurement values, while cold colors indicate low measurement values, which high values are desirable for POD and CSI, while low values are desirable for FAR.

Figure 4-4(a–c) shows that the DL-IR+WV model outperforms the other two in the summer, especially in Kansas, Missouri, and Oklahoma. For FAR (Figure 4-4(d–f)), no significant difference can be distinguished between the models, which is consistent with the FAR values presented in Table 4-3 (0.552, 0.575, and 0.554 for PERSIANN-CCS, the DL-IR only model, and the DL-IR+WV model, respectively). An ascending order can be observed in the maps of CSI of PERSIANN-CCS, the DL-IR only model, and the DL-IR+WV model (Figure 4-4(g–i)).

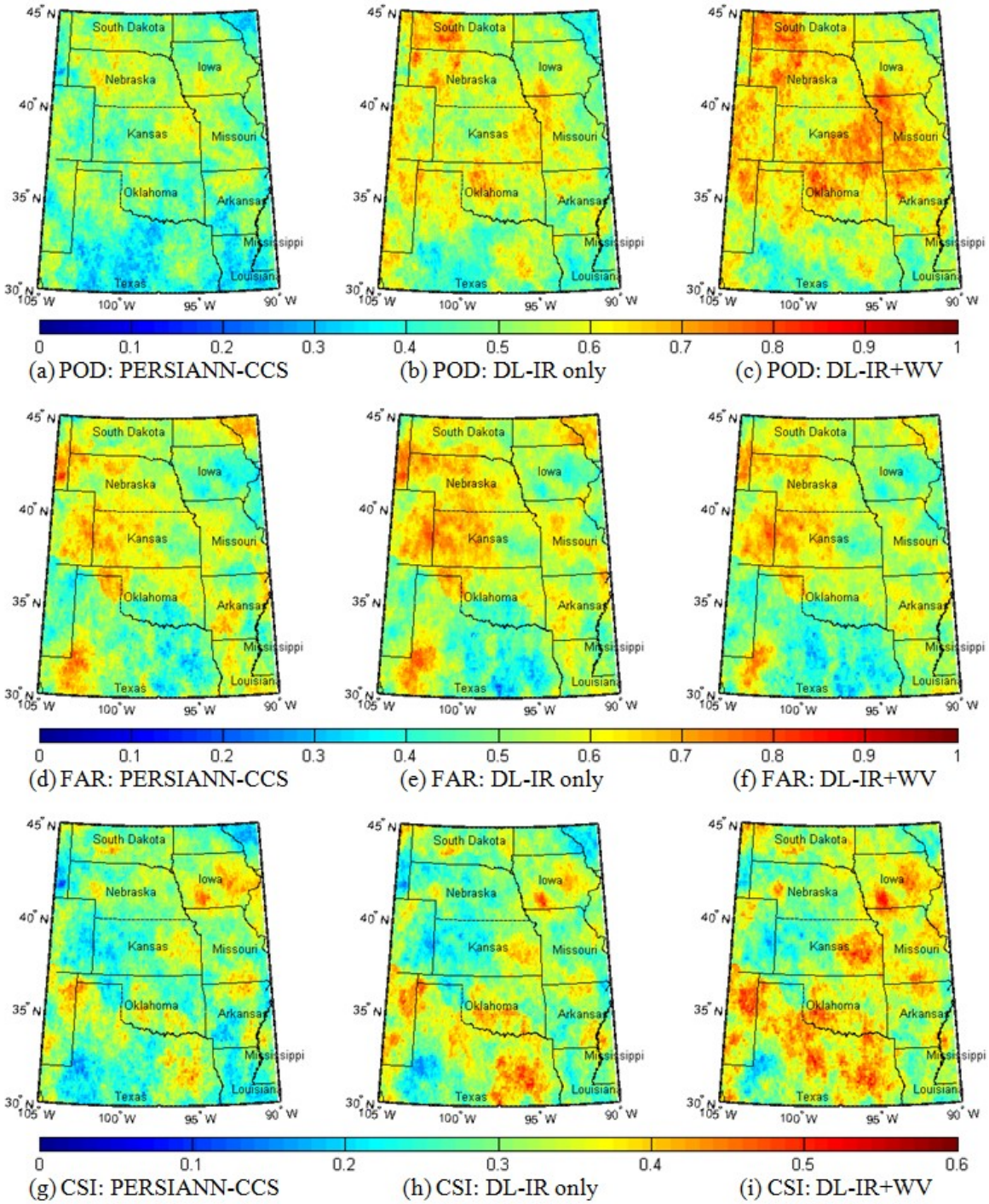

Figure 4-4 POD, FAR, and CSI of PERSIANN-CCS, the DL-IR only model, and the DL-IR+WV model over the central United States for summer 2013 (June–August). (a)–(c): POD; (d)–(f) FAR; (g)–(i) CSI

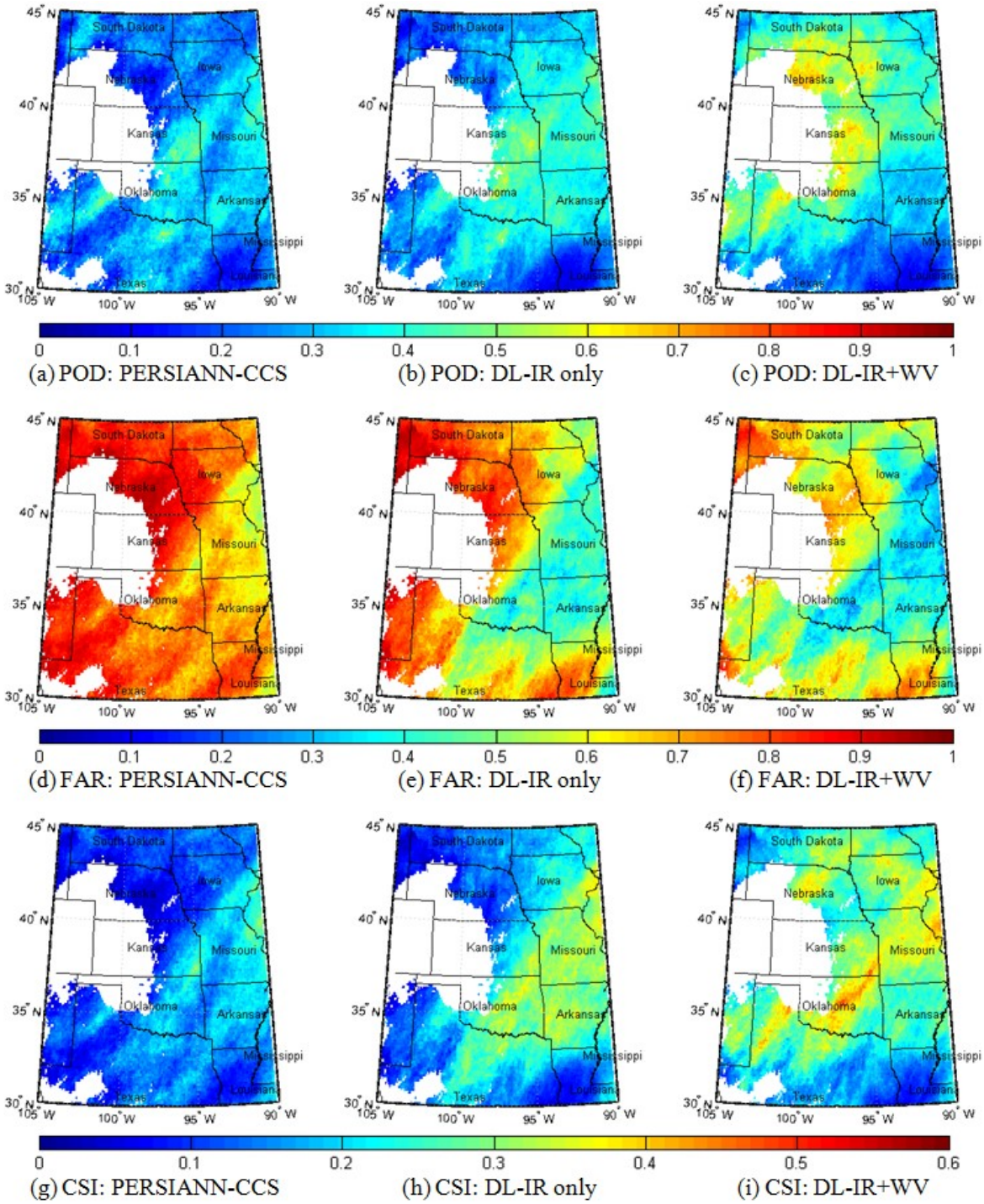

Figure 4-5 POD, FAR, and CSI of PERSIANN-CCS, the DL-IR only model, and the DL-IR+WV model over the central United States for winter 2013–2014 (December 2013–February 2014). (a)–(c): POD; (d)–(f) FAR; (g)–(i) CSI. The white color means that less than 50 precipitation pixels in the location were observed within the corresponding periods

Figure 4-5 shows very similar but more significant improvements in the winter. Moreover, noticeable decreases in FAR for the DL-IR only model and DL-IR+WV model can be observed, compared to PERSIANN-CCS (Figure 4-5(d–f)). Overall, performance improvements are consistent geographically for both models in both seasons. Furthermore, the DL-IR+WV model has the best performance for both seasons.

4.3.2 Case Studies

Table 4-4 Summary of R/NR classification performance in the case studies

		PERSIANN-CCS	DL-IR only (Model 1)	DL-IR+WV (Model 2)
7/26/13 19:00 UTC				
POD	Value	0.265	0.258	0.554
	Performance Gain	-	-2.64%	109.06%
FAR	Value	0.456	0.396	0.237
	Performance Gain	-	15.15%	92.41%
CSI	Value	0.217	0.221	0.473
	Performance Gain	-	1.84%	117.97%
12/21/13 21:00 UTC				
POD	Value	0.459	0.702	0.844
	Performance Gain	-	52.94%	83.88%
FAR	Value	0.099	0.125	0.153
	Performance Gain	-	-20.80%	-35.29%
CSI	Value	0.437	0.638	0.733
	Performance Gain	-	46.00%	67.73%

To further investigate how the models identify precipitation pixels from the IR and WV channels for precipitation events, Figures 4-6 and 4-7 provide visualizations of two case studies in the summer and winter, respectively. The two case studies selected help to reason the improvement of the performance of the DL-IR only and DL-IR+WV models, compared to PERSIANN-CCS. Their performance statistics are presented in Table 4-4.

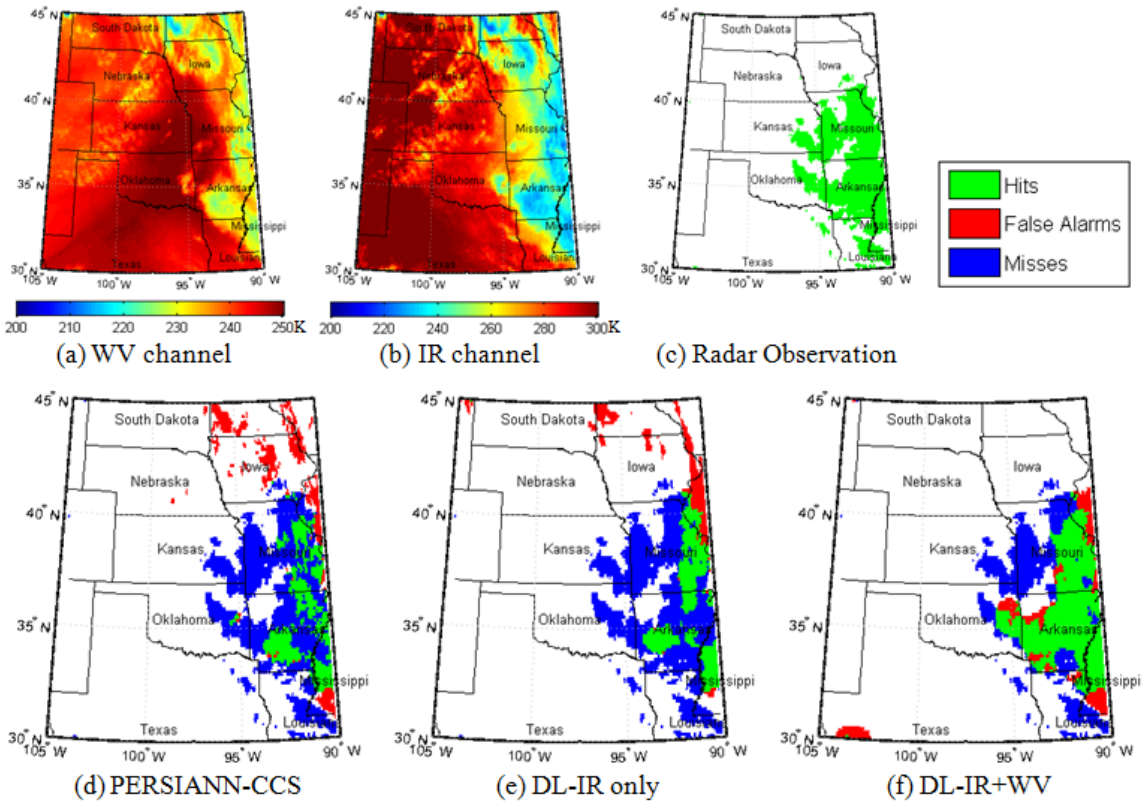

Figure 4-6 Visualization of WV and IR imageries and precipitation identification performance of PERSIANN-CCS, the DL-IR only model, and the DL-IR+WV model over the central United States for July 26, 2013, 19:00 UTC. (a), (b): snapshots of WV and IR channels; (c): radar R/NR observation; (d)–(f) hits, false alarms, and misses maps. Green, red, and blue indicate hits, false alarms, and misses, respectively

Figure 4-6 presents WV and IR imageries, Stage IV R/NR observations, and R/NR identification results for all models for a rainfall event on July 26, 2013, 19:00 UTC. The cloud patch on the east of the map is relatively warm with only a few cold pixels, as shown in Figure 4-6(b). Hence, only small sections of rainfall are correctly identified by PERSIANN-CCS (the green pixels in Figure 4-6(d)), while there is a large rainy area over Missouri and Arkansas captured by Stage IV observations (Figure 4-6(c)). On the other hand, the cold cloud over the northeast part (Figure 4-6(b)) leads to false alarms in PERSIANN-CCS (the red pixels in Figure 4-6(d)). Figure 4-6(e) shows that the DL-IR only model reduced the false alarm pixels while successfully identifying similar amounts of precipitation pixels. However, the overall improvement is marginal compared to PERSIANN-CCS (CSI performance gain: 1.84%). Nevertheless, the DL-IR+WV model shows a significant improvement in delineating the precipitation area (the green area in Figure 4-6(f)). Figure 4-6(d–f) shows that the DL-IR+WV model successfully connects the large area of precipitation captured by Stage IV observations instead of only identifying small sections similar to PERSIANN-CCS and the DL-IR model, especially the warm cloud over the center of Arkansas and the eastern region of Oklahoma. Moreover, the DL-IR+WV model avoided the false alarms on the northeast part of the map. The overall improvement is significant compared to PERSIANN-CCS (CSI performance gain: 117.97%). This case study shows supplementary information confirming that the WV channel supports precipitation identification.

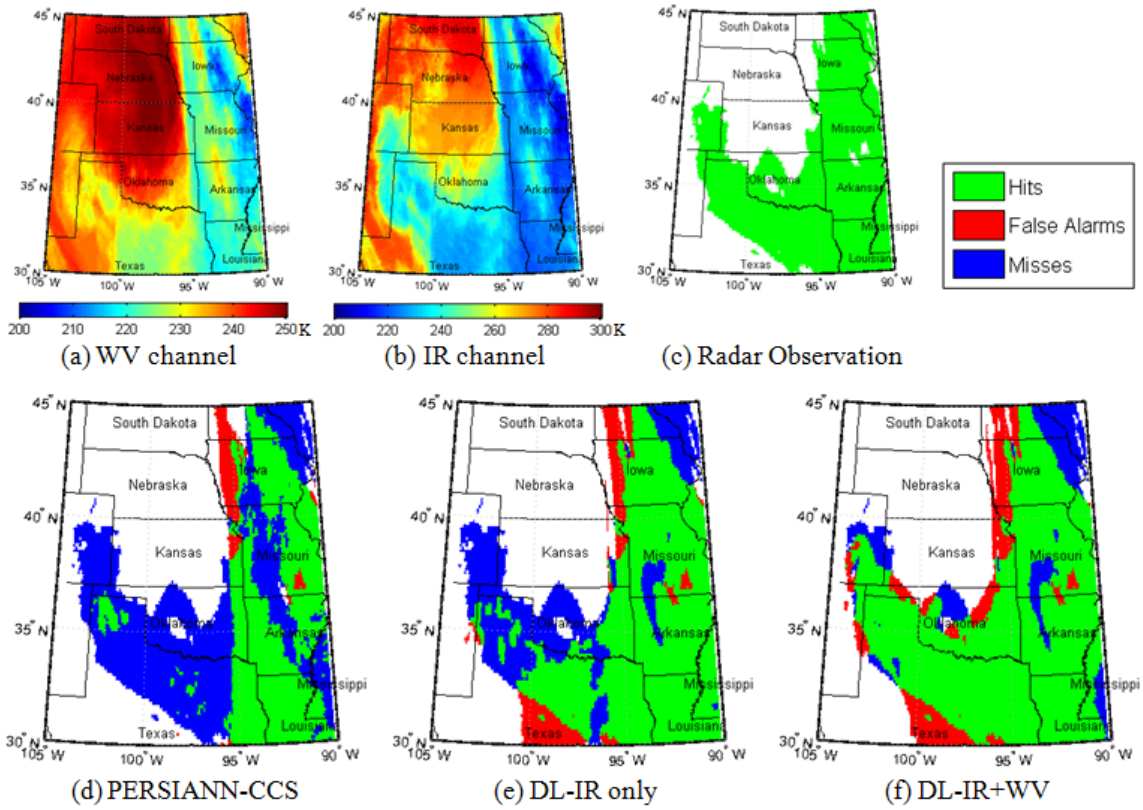

Figure 4-7 Visualization of WV and IR imageries and precipitation identification performance of PERSIANN-CCS, the DL-IR only model, and the DL-IR+WV model over the central United States for December 21, 2013, 21:00 UTC. (a)-(b): snapshots of WV and IR channels; (c): radar R/NR observation; (d)-(f) hits, false alarms, and misses maps. Green, red, and blue indicate hits, false alarms, and misses, respectively

Similarly, Figure 4-7 shows WV and IR imageries, Stage IV R/NR observation, and R/NR identification results for all models of a rainfall event on December 21, 2013, 21:00 UTC. As shown in Figure 4-7(d), PERSIANN-CCS misses a large area of precipitation in Oklahoma

and Texas, where the cloud is not as cold as that in the eastern part of the map (Figure 4-7(b)).

As shown in Figure 4-7(e), the DL-IR only model can extend the hits pixels in most parts of Texas, but still misses the pixels over the midwestern section of the map, where the cloud brightness temperatures are relatively high (Figure 4-7(b)). The overall improvement is already quite significant compared to PERSIANN-CCS (CSI performance gain: 46.00%). This proves the capability of the deep learning techniques to extract additional information in IR data, compared to PERSIANN-CCS. On the other hand, Figure 4-7(f) shows that the DL-IR+WV model almost correctly captures the entire precipitation region (green) with only a few false alarms and misses at the edges (blue and red). Its CSI performance gain is 67.73% regarding PERSIANN-CCS. It again shows value of the WV channel information.

4.4 Conclusion

As the first stage of the two-stage precipitation estimation model, this chapter explores the application of the deep learning techniques on precipitation identification with bispectral information (IR and WV channels). Two models are built to evaluate the effectiveness of the methodology and the value of the additional information provided by the WV channel. The first model, referred to as the DL-IR only model in this paper, applies deep learning

techniques to IR imageries to automatically extract useful features for precipitation identification. The second model, referred to as DL-IR+WV, incorporates WV imageries and its extracted features into the calibrated DL-IR only model to further improve the performance. Results of both models are compared with PERSIANN-CCS, an operational satellite-based precipitation estimation product.

The experiments show significant improvements for both models in R/NR detection compared to PERSIANN-CCS. The performance gains in CSI are 21.60% and 43.66% over the verification periods for the DL-IR only model and the DL-IR+WV model, respectively. For the winter, the performance gains in CSI are as high as 46.51% and 94.57% for the models. In addition, case studies in both seasons show that the deep learning techniques and the WV channel information can help delineate precipitation regions with relatively warm clouds, while reducing false alarms with cold clouds.

The improved performance of the DL-IR only model demonstrates that sparse automatically extracted features from the IR imageries can assist in better detecting precipitation compared to a limited number of hand-designed features. On the other hand, the significant improvement in the DL-IR+WV model, compared to both PERSIANN-CCS and the DL-IR only model, verifies that the information contained in WV imageries supplements the IR data to better delineate precipitation regions. The case studies show that the

additional WV information can help capture warm-cloud precipitation likely to be missed by other models and identify non-raining cold clouds.

With an effective R/NR binary classification model, the next step is to implement a DNN to estimate precipitation amount of the pixels classified with precipitation and eventually provide a complete satellite-based precipitation estimation product. In the next chapter, I illustrate the second stage of the two-stage precipitation estimation model that focuses on precipitation amount regression for those pixels with precipitation.

Chapter 5

Two-Stage Model Part II: Precipitation Amount Regression

5.1 Introduction

Chapter 4 introduced the motivation and background of developing a two-stage model for precipitation estimation from bispectral satellite information and then described and demonstrated the first step of the model, R/NR binary classification. In this chapter, I focus on the description of the second step of the model, precipitation amount regression.

As discussed in Section 2.1, in estimating the precipitation amount, there are multiple objectives: accurate R/NR classification, minimize point-wise differences, and preserve the distribution of the precipitation values that can capture large to extreme events. In Chapter 4, I focused on the first objective and eliminated a large amount of NR pixels. In this section, I explore the capability of DNNs to work as a regression model to estimate the real-valued precipitation amount.

In the development of DNNs in the field of machine learning and computer vision, DNNs are demonstrated to be very effective in classification problems (Ciregan et al., 2012; He et al., 2015; Karpathy et al., 2014; Krizhevsky et al., 2012; Sun et al., 2014; Szegedy et al., 2015). In contrast, there are relatively few regression applications for DNNs, especially for the highly skewed datasets. More specifically, for satellite-based precipitation estimation, the overall use of data-driven methods is still limited (Tao et al., 2016b). It is challenging to design the structure, objective function, and training process of DNNs to fit the needs of such problems.

In this chapter, I explore the application of the deep learning techniques to real-valued precipitation amount estimation with bispectral information (IR and WV channels), followed by the R/NR classification described in Chapter 4. Specifically, I aim to accomplish the following tasks: (1) manage the highly skewed distribution of precipitation data by properly modifying DNN's objective function; (2) demonstrate the effectiveness of the overall two-staged model by comparing its performance with an operational product, PERSIANN-CCS; and (3) evaluate and analyze the case study results in the summer and winter seasons.

The remainder of this section is organized as follows. Section 5.2 describes the detailed methodology and model setup for the non-zero precipitation amount estimation part of the two-stage model. In Section 5.3, the model is verified by comparisons between the outputs

of this model and the estimates from PERSIANN-CCS product using multiple common verification measures for precipitation estimation products. Finally, the main conclusions are summarized in Section 5.4.

5.2 Methodology and Model Setup

5.2.1 Objective Function

MSE is the most commonly used objective function in supervised learning problems. It is a reasonable performance measure for many real valued problems. However, the imbalance of the precipitation makes MSE not suitable to work as the sole optimization objective. Optimized by MSE only, the models tend to avoid predicting large rainfall while the one of the main uses of satellite-based precipitation estimation is to track large rainfall events, as discussed in Section 2.1.

Despite the efforts of incorporating data-driven approaches in previous studies, overviewed in Section 2.1, idealized assumptions are imposed to ensure that large values can be predicted by the model. In this section, I add a KL divergence term to the MSE in the loss function, aiming to minimize the estimation error while preserving the distribution of the rainfall amount. Specifically, in the fine-tuning step of the training process, a mini-batch gradient descent algorithm is applied. For each mini-batch, I estimate the KL divergence

between the distributions of the estimates and the target values. First, the density functions are estimated by using a normal kernel function to smooth the samples. The bandwidth used for the kernel smoothing window is 5 mm/hr. Then the KL divergence can be estimated as

$$D_{KL}(P||Q) \approx \sum_x p(x) \log p(x) - \sum_x p(x) \log q(x)$$

where P and Q are the distributions of the target values and of the estimated values, respectively; p and q are corresponding density functions; and x denotes a sequence of values across the range of the precipitation value (0 to 100 mm/hr). Then the loss function is a weighted average of the MSE and the KL divergence

$$L = \frac{1}{2n} \left|\left|\boldsymbol{y} - \hat{\boldsymbol{y}}\right|\right|^2 + \alpha D_{KL}(P||\hat{P})$$

where y and $\hat{\mathrm{y}}$ are vectors of target values and their estimates; α is the weight of the KL divergence term while fixing the weight of the average MSE term to be 1; $\mathrm{D_{KL}(P||\widehat{P})}$ is the KL divergence from the distribution of precipitation amount estimates $\widehat{\mathrm{P}}$ to the true distribution of target values.

When applying backpropagation, the mini-batch gradient descent is applied since the process of estimating p and q requires multiple data points. In this research, the mini-batch size selected is 1000. Then the derivative of the KL divergence can be expressed as

$$D_{KL}{}'(P||Q) \approx \sum_x p'(x)\,(\log\frac{p(x)}{q(x)} + 1) - \sum_x q'(x)\frac{p(x)}{q(x)}$$

where $p'(x)$ and $q'(x)$ can be easily estimated by using a finite difference estimate of the derivative.

5.2.2 Experimental Design

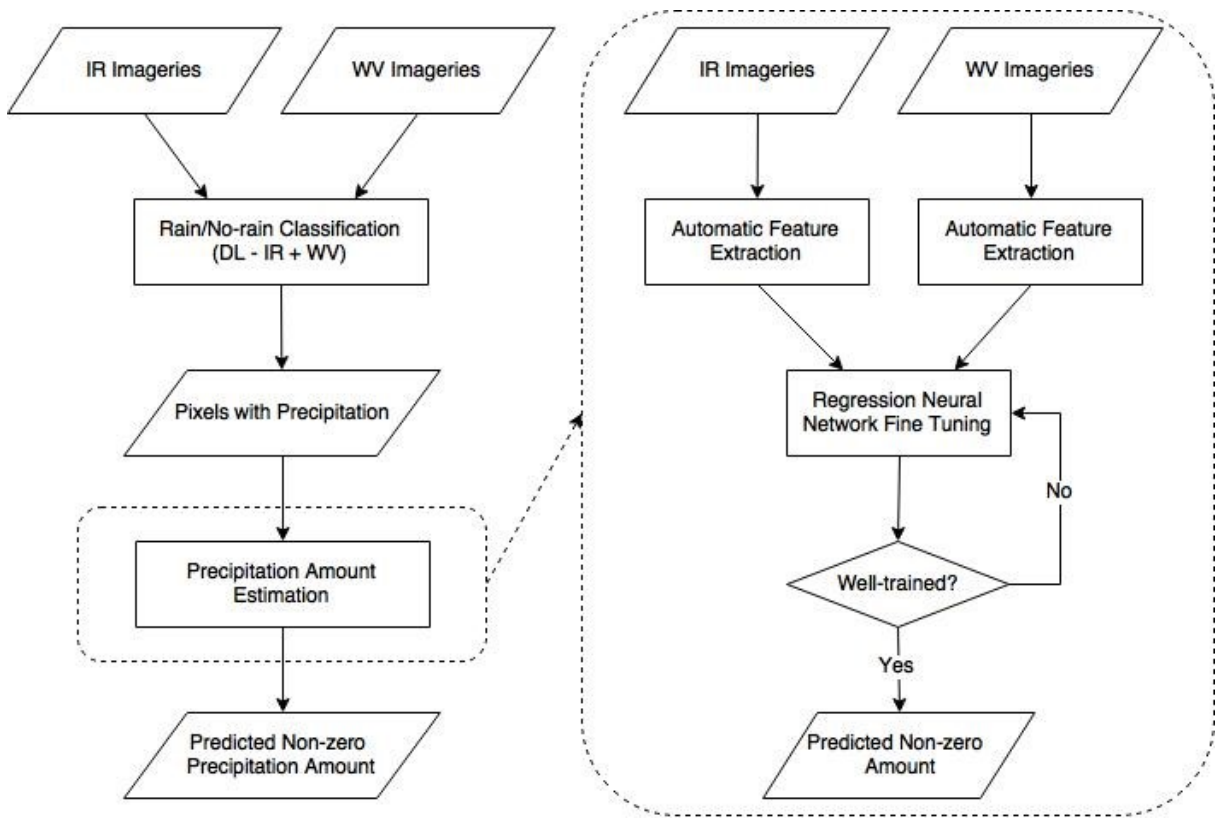

Figure 5-1 Overview of the precipitation estimation process of the two-stage model: the inputs to the precipitation amount estimation model are IR and WV imageries of the pixels classified as with precipitation in the R/NR model

The overview of the experimental design is presented in Figure 5-1. Figure 5-1 shows the whole process of the two-stage model (left part) and the detailed training process for the

second step of the model (right part). As the second step of the two-stage model, I build a non-zero precipitation amount estimation model, referred as DL-Regression in the remainder of this dissertation, using bispectral information only dealing with pixels classified as with precipitation in the first stage (DL- IR + WV). The structure of the DL – Regression model is similar to DL – IR + WV but in a design of regression instead of classification, as shown in Figure 5-2.

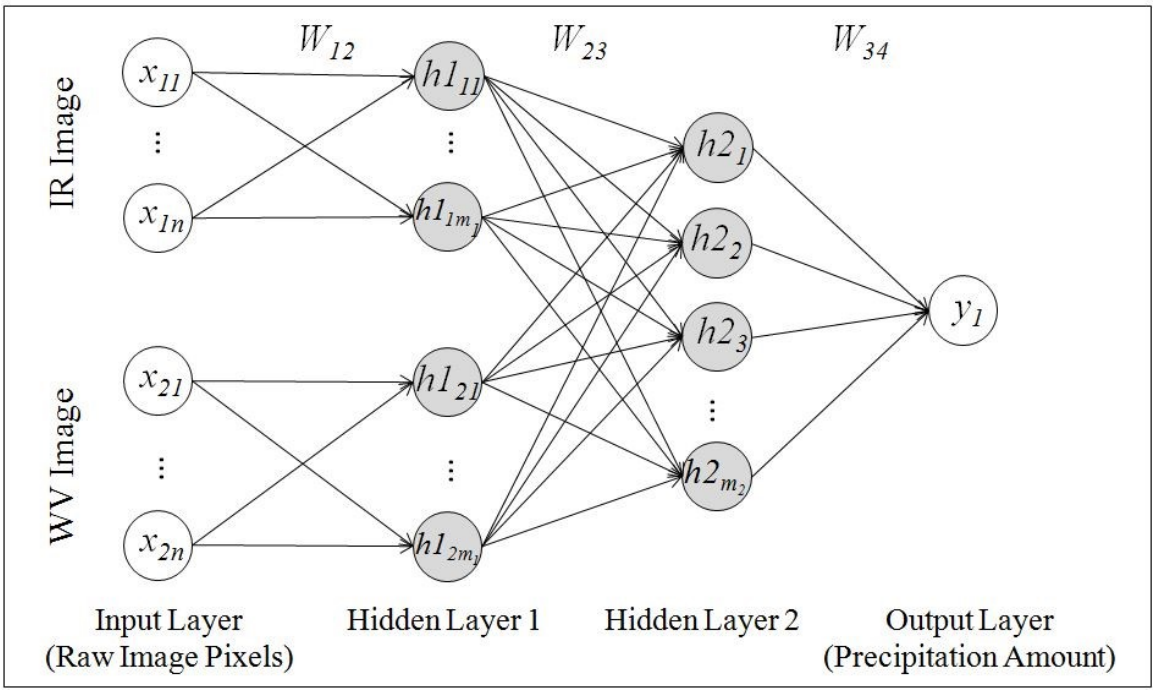

Figure 5-2 Structure of a four-layer neural network with bispectral imageries: the input layer has two images that connect to half of the hidden nodes in the first hidden layer separately; the next two layers are the hidden layers, which can be interpreted as extracted features; the last layer is the output layer, which contains the precipitation amount for the centered pixel

Here, each input patch is a 15 × 15 pixel IR or WV patch, which covers approximately 120 km × 120 km. The training data pool is created by a 15 × 15 pixel window moving around the study region with overlap. The number of hidden nodes for the first hidden layer for the DL-Regression is 2,000, with 1,000 nodes connected to each input patch. The second layer contains 1,000 hidden nodes. The output layer contains one real-valued node, predicting the actual amount of precipitation for the centered pixel of the 15 × 15 pixel input patch. These hyper-parameters are selected based on the previous study parameter selection results and experiments with different number combinations (Tao et al., 2016a; Tao et al., 2016b; Vincent et al., 2010). One additional hyper-parameter needed is the weights of the KL divergence term and the MSE term. For this parameter, I considered the coefficient of KL divergence in the pool of 0.1, 1, 10, 100, and 1000 while maintaining the coefficient of MSE as 1. Additionally, it was experimentally selected to be 100:1.

5.3 Results and Discussion

In this section, I evaluate the performance of the DL-Regression model in the verification periods (summer 2013 and winter 2013–2014) in comparison with the data of PERSIANN-CCS. The validation measurements I used to evaluate precipitation amount estimation performance are average bias, average MSE, COR, and the performance gain of these measurements with respect to PERSIANN-CCS. The definition of performance gain is

presented in Table 5-1. Notice that the DL-Regression model is the second step of the two-stage model. Thus, the precipitation estimation amount is a result from the combined actions of the two steps. In other words, the evaluation is towards the overall two-stage model, instead of the sole second step. In addition, a few case studies are provided as examples to analyze the precipitation estimation performance between the two-stage model and PERSIANN-CCS.

*Table 5-1 Definition of performance gain for MSE, Bias, and COR**

Verification Measures	**Formulas**	**Range and Desirable Value**
Performance Gain for MSE	$PG = \frac{Val_{ref} - Val_{model}}{Val_{ref}} \times 100\%$	Range: -∞ to 1 Desirable Value: 1
Performance Gain for Bias	$PG = \frac{\lvert Val_{ref}\rvert - \lvert Val_{model}\rvert}{\lvert Val_{ref}\rvert} \times 100\%$	Range: -∞ to 1 Desirable Value: 1
Performance Gain for COR	$PG = \frac{Val_{model} - Val_{ref}}{Val_{ref}} \times 100\%$	Range: -∞ to +∞ Desirable Value: +∞

* MSE, Bias, and COR are defined in Table 2-2; Val_{model} denotes the value of a measurement for a model; Val_{ref} denotes the value of a measurement for a reference.

5.3.1 Precipitation Amount Estimation

Table 5-2 provides the overall performances of PERSIANN-CCS and the two-stage model over the verification periods. The two-stage model shows obvious improvement in all three measurements compared to PERSIANN-CCS. Specifically, the two-stage model reduces average bias and average MSE by 40.58% and 51.84%, respectively (0.041 compared to 0.069 and 0.641 compared to 1.331, respectively). Meanwhile, the two-stage model has a decent increase in COR compared to PERSIANN-CCS (23.49%; 0.410 compared to 0.332), which is indirectly optimized in the model. The performance of the two-stage model demonstrates the effectiveness of the design of the two stages and DNN structures of the both DL - IR + WV (described in Section 4.2) and DL-Regression. In addition, the WV channel data also contribute to improving the performance.

Table 5-2 Summary of precipitation estimation performance over the verification periods (including both summer 2013 and winter 2013–2014)

		PERSIANN-CCS	**Two-stage Model**
Average Bias	Value	0.069	0.041
	Performance Gain	-	40.58%
Average MSE	Value	1.392	0.641
	Performance Gain	-	53.95%
COR	Value	0.332	0.410
	Performance Gain	-	23.49%

Figure 5-3 presents maps of the average bias of PERSIANN-CCS and the two-stage model precipitation over the study region averaged on the warm and cold verification periods, respectively. The white color indicates very small bias and shows that the two-stage model outperforms the PERSIANN-CCS model significantly, especially in the summer season.

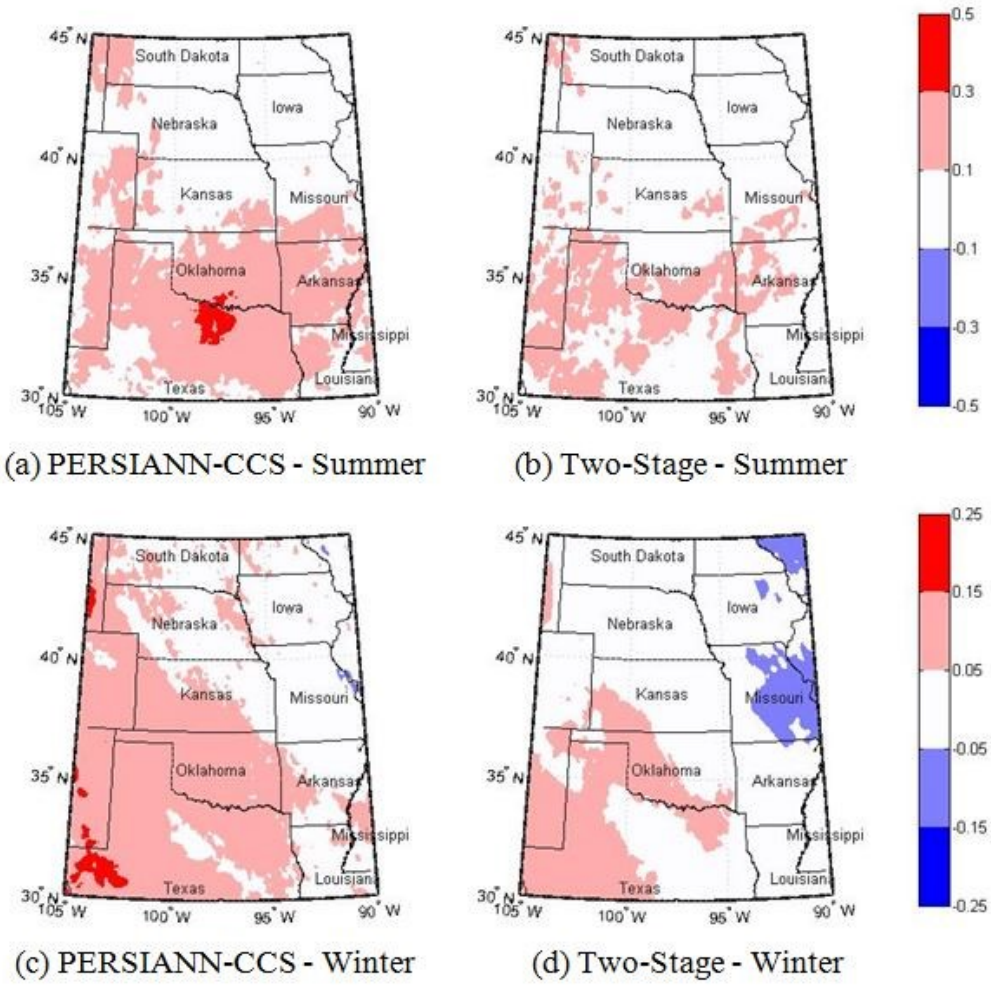

Figure 5-3 Average Bias (mm/hr) of PERSIANN-CCS [(a), (c)] and the two-stage model [(b), (d)] output over the central United States (30–45°N, 90–105°W): (a), (b) summer (June–August 2013); (c), (d) winter (December 2013–February 2014)

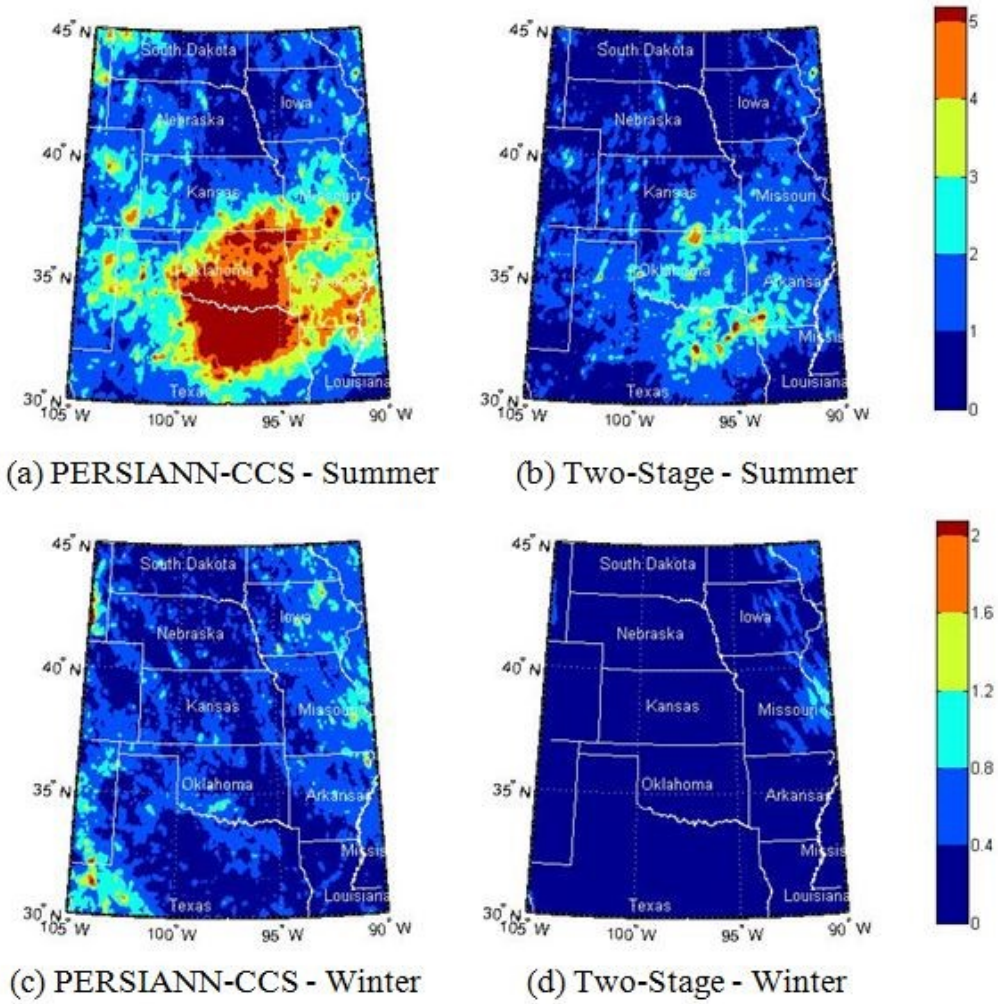

Figure 5-4 Average MSE ([mm/hr]^2) of PERSIANN-CCS [(a), (c)] and the two-stage model [(b), (d)] output over the central United States (30–45°N, 90–105°W): (a), (b) summer (June–August 2013); (c), (d) winter (December 2013–February 2014)

Specifically, the overestimation produced by the PERSIANN-CCS product is mostly removed in the two-stage model for the summer period (Figure 5-3 (a, b)). For winter period, the two-stage model reduces the overestimation in PERSIANN-CSS over the southwestern area

but slightly increases the underestimation in the northeastern area in the study region. Specific calculations are displayed in Table 5-3.

Similar results can be seen in Figure 5-4, which shows the average MSE of PERSIANN-CCS and the two-stage model precipitation over the study region averaged on the warm and cold verification periods, respectively. The warm colors indicate strong differences compared to Stage IV observations, while the cold colors indicate small differences. Compared to PERSIANN-CCS, Figure 5-4 shows that errors in both summer and winter periods strongly decrease in the two-stage model. For the summer period, the heavy error over Oklahoma and Texas are strongly reduced in the two-stage model (Figure 5-4(a, b)). For the winter period, the MSEs are smaller than 0.4 $[mm/hr]^2$ for most pixels for the two-stage model (Figure 5-4(d)), and specific calculations are displayed in Table 5-3.

Table 5-3 provides detailed values for of average bias, average MSE, and COR for PERSIANN-CCS and the two-stage model over different seasons in the verification periods. The average biases are 0.063 mm/day and 0.018 mm/day for the two-stage model, compared to 0.091 mm/day and 0.054 mm/day before bias correction for summer and winter (30.77% and 66.26% performance gain), respectively. There are over 50% improvement in average MSE for both seasons (51.63% and 66.29% for summer and winter, respectively). Compared to PERSIANN-CCS, the COR for the two-stage model is 66.40% higher in the winter season and 18.58% higher in the summer season. In addition,

the overall performance gain is higher for the winter season (66.26%, 66.29%, and 66.40% for average bias, average MSE, and COR, respectively) than the summer season (30.77%, 51.63%, and 18.58% for average bias, average MSE, and COR, respectively). This is similar to the results showed in Section 4.3.

Table 5-3 Summary of precipitation estimation performance for summer 2013 and winter 2013–2014, respectively

		PERSIANN-CCS	Two-stage Model
Summer Season (June–August 2013)			
Average Bias	Value	0.091	0.063
	Performance Gain	-	30.77%
Average MSE	Value	2.338	1.131
	Performance Gain	-	51.63%
COR	Value	0.350	0.415
	Performance Gain	-	18.58%
Winter Season (December 2013–February 2014)			
Average Bias	Value	0.054	0.018
	Performance Gain	-	66.26%
Average MSE	Value	0.445	0.150
	Performance Gain	-	66.29%
COR	Value	0.211	0.351
	Performance Gain	-	66.40%

5.3.2 Case Studies

To demonstrate the performance of the two-stage model for specific events, Figure 5-5 presents snapshots from an event on Jun. 1st, 2013 13:00 UTC. The two-stage model (Figure 5-5(b)) captures well the narrow shape of the rainfall over Kansas and Missouri and the separate small region with precipitation on the border of Missouri and Arkansas. In contrast, PERSIANN-CCS (Figure 5-5 (a)) presents a fatter shape of the rainfall and misses the precipitation on the east.

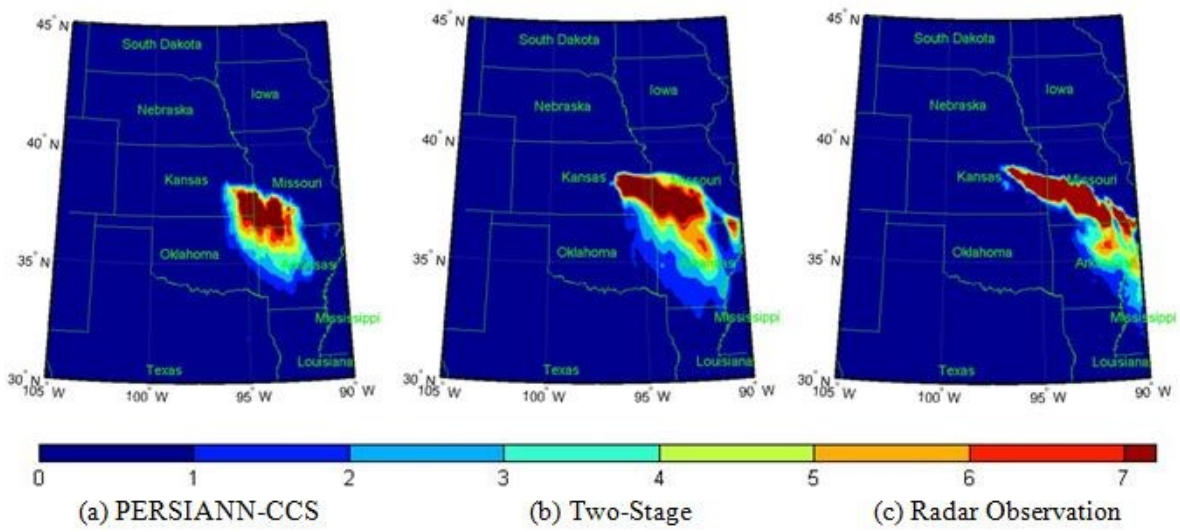

Figure 5-5 Snapshots (mm/hr) of PERSIANN-CCS, the two-stage model, and the radar observation over the central United States for June 1st, 2013, 13:00 UTC. (a) PERSIANN-CCS; (b) the two-stage model; and (c) Radar Observation

This shift of the rainfall shape might be due to its critical assumption of low brightness temperature leading to high rain rates. It also demonstrates that the two-stage model is

capable of better capturing the shape of the precipitation and its peak from the cloud information with the flexibility in the model. In addition, it can extract more useful precipitation-related features and capture some rainfall that PERSIANN-CCS fails to extract.

5.4 Conclusion

This chapter described the second step of the two-stage model: non-zero precipitation amount estimation. I explore the application of DNN for precipitation amount regression with bispectral information (IR and WV channels). The basic structure of the model is similar to the R/NR binary classification model. However, due to the challenging imbalance and high skew in the precipitation data, I incorporate a KL divergence term, in addition to the MSE term, in the objective function. Thus, the model can produce low point-wise estimation error while preserving the distribution of the rainfall amounts. After several combinations of hyper-parameters were tested, the weights of the KL divergence term and the MSE term were set to 100:1 to best balance the needs of different objectives.

In the comparison with PERSIANN-CCS, the two-stage model is 40.58% lower in average bias, 53.95% lower in average MSE, and 23.49% higher in COR in the overall verification periods. Specifically, the improvement is greater for the winter season than the summer

season. In addition, the case studies suggest that the two-stage model can better capture the shape and the peak of the rainfall event from the satellite information.

Combining with the results showed in Section 4.3, I demonstrated the effectiveness of the two-stage deep learning framework for precipitation estimation from bispectral satellite information in both R/NR classification and real-value amount estimation. The contribution of the improvement includes the additional information provided by WV channel, the capability of deep learning techniques to extract useful features from the raw satellite inputs, and the proper design of neural networks to separately tackle precipitation identification and the amount estimation.

To serve as an operational product, it is essential to validate the model at a large scale to demonstrate its capability to eventually serve at a global coverage. With our experiments in the central United States in Chapter 4 and Chapter 5, the next step is to enlarge this verification coverage. In the next chapter, I apply the developed model to a large coverage of the United States to evaluate the performance of the two-stage model in the areas outside its calibration region.

Chapter 6

Large-Scale Application

6.1 Introduction

As discussed in Section 1.1, one of the most important properties of satellite-based precipitation estimation products is their global application. This is essential for weather and climate analysis at a large scale or focusing on remote regions. Therefore, in developing satellite-based precipitation estimation models, the evaluation of large-scale application is necessary before promoting such models to near-real-time operational products (Tao et al., 2017). In this Chapter, I focus on the evaluation of the developed two-stage model at a large scale to further demonstrate the potential of the model to serve as an operational product.

In Section 4.3 and Section 5.3, I experimentally demonstrated the effectiveness of the design of the two-stage model, structure of the DNNs, and the value of the WV channel data in addition to the IR channel. The study region, as introduced in Section 2.4, is the central United States. In other words, both training and test data come from pixels within this region but with different time periods.

In this section, to demonstrate the two-stage model's capability to generalize to large-scale application, I use the calibrated two-stage model from Chapter 4 and Chapter 5 and apply it directly to a larger region with ground observation for verification. The study region selected is the stable coverage of Stage IV data (30–45°N, 85–115°W), as shown in Figure 6-1.

Figure 6-1 Map of the large-scale application region

The remainder of this section is organized as follows. Section 6.2 presents a comparison between the output of the two-stage model and PERSIANN-CCS, using Stage IV data as reference. Finally, the main conclusions are summarized in Section 6.3.

6.2 Results and Discussion

In this section, I evaluate the performance of the two-stage model in the verification periods (summer 2013 and winter 2013-2014) over a large-scale coverage of United States, in comparison with the data of PERSIANN-CCS. The verification includes both precipitation identification and real-valued amount estimation. The validation measurements I used to evaluate precipitation identification performance are POD, FAR, and CSI. The estimation amount measurements used are average bias, average MSE, and COR. The definitions of the measurements are presented in Section 2.1. For all measurements, the performance gains with respect to PERSIANN-CCS are also presented.

6.2.1 Rain/No-Rain Classification

Table 6-1 summarizes the overall R/NR classification performances of PERSIANN-CCS and the two-stage model over the verification periods of the large coverage of the United States (30–45°N, 85–115°W). The two-stage model shows significant performance gain (32.56% in CSI) in precipitation identification compared to PERSIANN-CCS. More specifically, the two-stage model has 23.30% performance improvement in POD compared to PERSIANN-CCS (0.418 compared to 0.339) while the performance improvement of FAR is 19.32% (0.528 compared to 0.630). Compared to the results presented in Table 4-2, the overall improvement is not as significant as the verification on the main study region (32.56% compared to 43.66% in CSI). This is understandable since now I apply the model

to areas it was not trained on, which introduces more errors and uncertainty. However, the overall promising performance demonstrates the capability of the two-stage model in delineating precipitation regions correctly, even in the model's unfamiliar areas.

Table 6-1 Summary of R/NR classification performance of PERSIANN-CCS and the two-stage model over the verification periods (including both summer 2013 and winter 2013–2014) of the large coverage of the United States (30–45°N, 85–115°W)

		PERSIANN-CCS	**Two-Stage Model**
POD	Value	0.339	0.418
	Performance Gain	-	23.30%
FAR	Value	0.630	0.528
	Performance Gain	-	19.32%
CSI	Value	0.215	0.285
	Performance Gain	-	32.56%

Table 6-2 provides specific details on performances of the models over different seasons in the verification periods and over the large coverage of the United States. . In the summer and winter seasons, over 20% increases in POD can be detected in the two-stage model compared to PERSIANN-CCS (22.57% and 21.02% performance gains, respectively), while FAR remains similar to PERSIANN-CCS for the summer season (2.09% performance gain) but has a significant improvement for the winter season (40.85% performance gain). Similar to the pattern found in Table 4-3, compared to PERSIANN-CCS, the overall improvements shown in CSI are higher for the winter (45.45% performance gain) than the

summer (16.67% performance gain). These performances are highly consistent with the model's performances in the main study region, which proves the model's capability of extending to unfamiliar regions. In addition, the performance gains over the large coverage of the United States are less than the performance gains for the central United States for both seasons, which is consistent with the results presented in Table 6-1.

Table 6-2 Summary of precipitation estimation performance of PERSIANN-CCS and the two-stage model over the large coverage of the United States for summer 2013 and winter 2013–2014, respectively

		PERSIANN-CCS	Two-Stage Model
Summer Season (June–August 2013)			
POD	Value	0.381	0.467
	Performance Gain	-	22.57%
FAR	Value	0.537	0.526
	Performance Gain	-	2.09%
CSI	Value	0.264	0.308
	Performance Gain	-	16.67%
Winter Season (December 2013–February 2014)			
POD	Value	0.295	0.357
	Performance Gain	-	21.02%
FAR	Value	0.662	0.470
	Performance Gain	-	40.85%
CSI	Value	0.187	0.272
	Performance Gain	-	45.45%

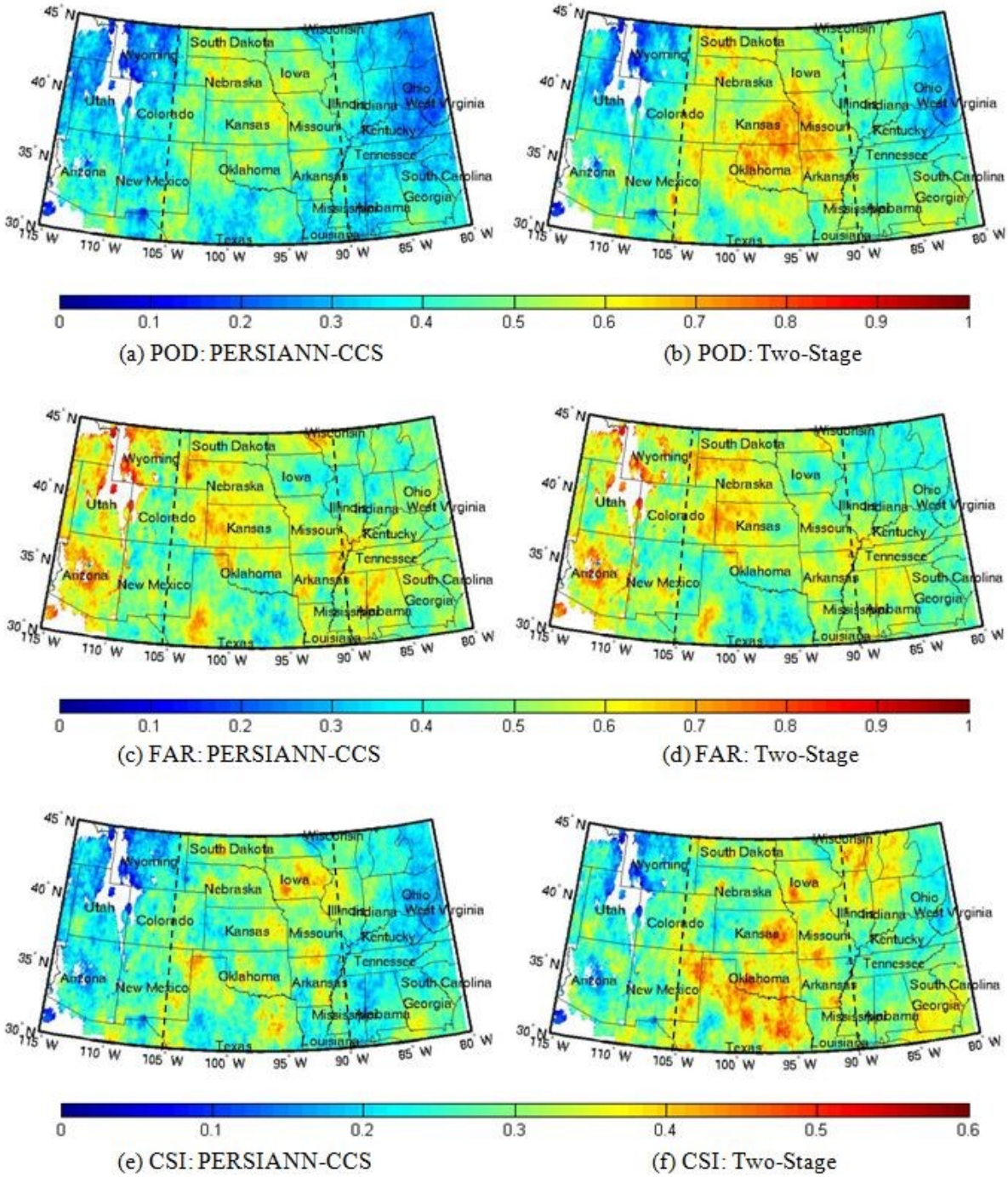

Figure 6-2 POD, FAR, and CSI of PERSIANN-CCS and the two-stage model over the large coverage of the United States for summer 2013 (June–August). (a), (b): POD; (c), (d) FAR; (e), (f) CSI. The white color means that less than 50 precipitation pixels in the location are observed within the corresponding periods

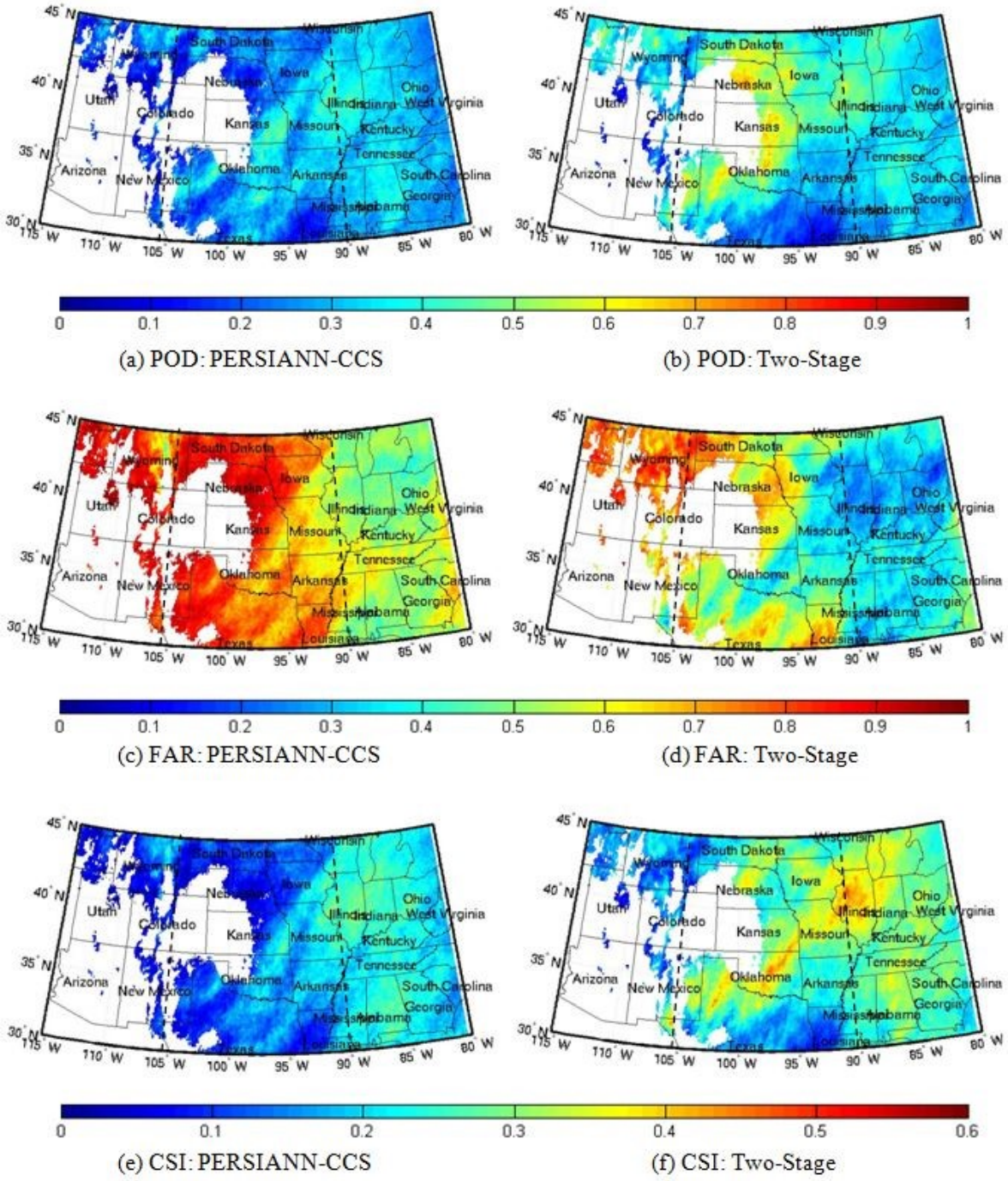

Figure 6-3 POD, FAR, and CSI of PERSIANN-CCS and the two-stage model over the large coverage of the United States for winter 2013–2014 (December 2013–February 2014). (a), (b): POD; (c), (d) FAR; (e), (f) CSI. The white color means that less than 50 precipitation pixels in the location are observed within the corresponding periods

Figures 6-2 and 6-3 present the maps of POD, FAR, and CSI of the PERSIANN-CCS and the two-stage model over the large coverage in the summer and winter verification periods, respectively. The region between the black dotted lines is the main study region, where the model was trained in. The warm colors indicate high measurement values, while cold colors indicate low measurement values. The white color indicates that less than 50 precipitation pixels in the location are observed within the corresponding periods. High values are desirable for POD and CSI, while low values are desirable for FAR. Figure 6-2(a, b) shows that the two-stage model uniformly outperforms PERSIANN-CCS over the middle to east part. For FAR (Figure 6-2(c, d)), the performance of the two models are almost the same, which is consistent with the FAR values presented in Table 6-2 (0.537 and 0.526 for PERSIANN-CCS and the two-stage model, respectively). Compared to PERSIANN-CCS, the two-stage model shows an obvious improvement in CSI (Figure 6-2(e, f)), especially in the middle to east part. More specifically, 84.4% of the pixels show better performances in CSI with the two-stage model, compared to PERSIANN-CCS, as presented in Table 6-3. The improved performances found in the northeast and southeast particularly demonstrate the model's capability of extending to unfamiliar regions. However, in the west part, no significant improvement can be visually distinguished. This might be because the model was not trained in this region and thus fails to capture the forms of rainfall in this region. Similarly, Figure 6-3 presents the results for winter season. In CSI, an obvious improvement can be found over the mid-east part of the map, centered with Illinois. There

are also identifiable improvements in the northwest part of the map, where enough precipitation pixels are in the verification period. Specifically, 80.7% of the pixels show better performance in CSI with the two-stage model, compared to PERSIANN-CCS, as presented in Table 6-3. Overall, performance improvements are significant and consistent geographically for both seasons.

Table 6-3 Percentage of pixels with better performance in the two-stage model, compared to PERSIANN-CCS model over the large coverage of the United States for summer 2013 and winter 2013–2014, respectively

	Summer	**Winter**
POD	89.4%	75.8%
FAR	62.9%	77.3%
CSI	84.4%	80.7%

6.2.2 Precipitation Amount Estimation

Table 6-4 provides the overall performances of PERSIANN-CCS and the two-stage model over the verification periods of the large coverage of the United States (30–45°N, 85–115°W). The two-stage model has significant improvement in all three measurements compared to PERSIANN-CCS. The two-stage model reduces average bias and average MSE by 23.40% and 44.52%, respectively (0.036 compared to 0.047 and 0.562 compared to 1.013, respectively). At the same time, the two-stage model has 27.21% higher COR compared to PERSIANN-CCS (0.374 compared to 0.294). Compared to the results

presented in Table 5-2, the improvement in average bias and average MSE are not as significant as over the main study region (23.40% compared to 40.58% and 44.92% compared to 53.95%, respectively), which is expected. However, the performance gain of COR in the larger coverage is slightly higher than the main study region (27.21% compared to 23.49%), which states that the two–stage model is capable of consistently capturing the pattern of the precipitation amount. Additionally, the overall performance demonstrates the capability of the two-stage model in accurately estimating precipitation amount, even in the model's unfamiliar areas.

Table 6-4 Summary of precipitation estimation performance of PERSIANN-CCS and the two-stage model over the large coverage of the United States (including both summer 2013 and winter 2013–2014)

		PERSIANN-CCS	**Two-Stage Model**
Average Bias	Value	0.047	0.036
	Performance Gain	-	23.40%
Average MSE	Value	1.013	0.562
	Performance Gain	-	44.52%
COR	Value	0.294	0.374
	Performance Gain	-	27.21%

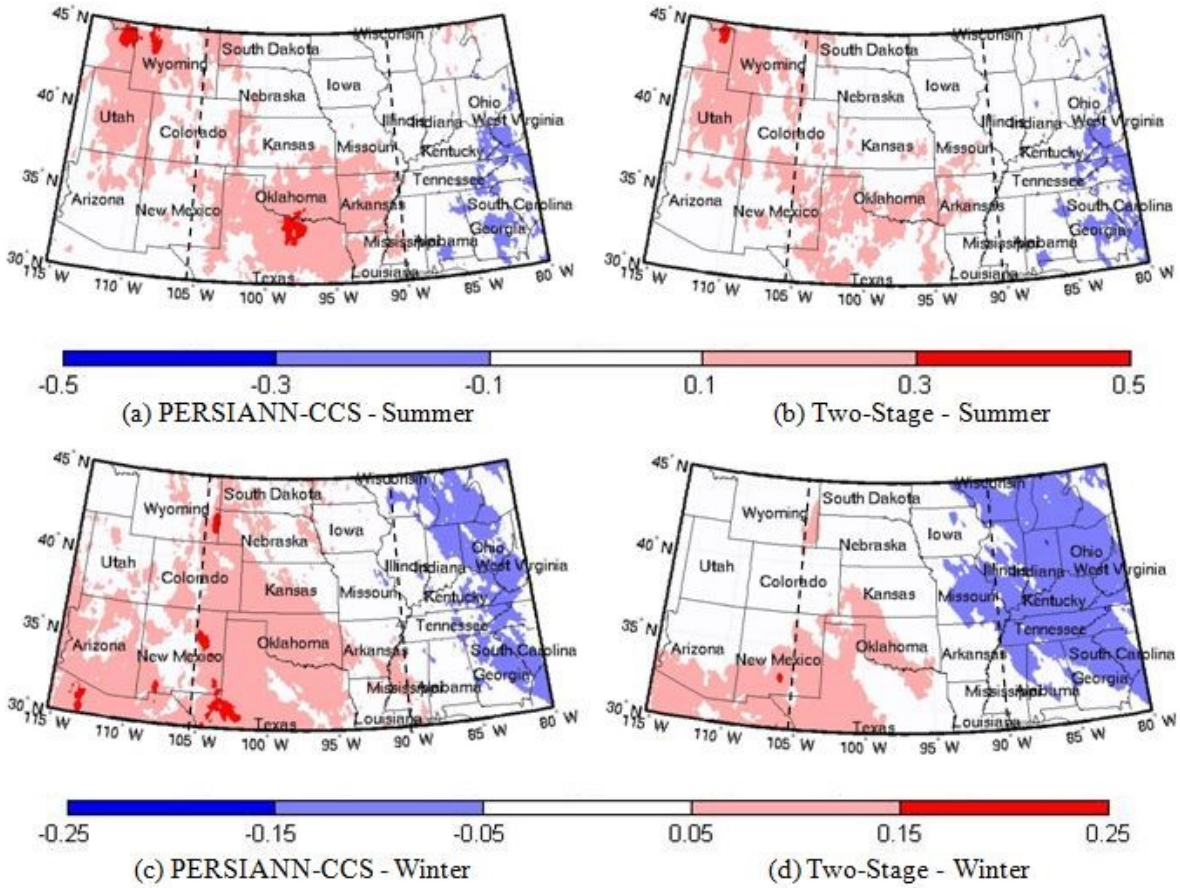

Figure 6-4 Average Bias (mm/hr) of PERSIANN-CCS [(a), (c)] and the two-stage model [(b), (d)] output over the large coverage of the United States (30–45°N, 85–115°W): (a), (b) summer (June–August 2013); (c), (d) winter (December 2013–February 2014)

Figure 6-4 presents maps of the average bias of PERSIANN-CCS and the two-stage model precipitation over the large coverage of the United States averaged on the warm and cold verification periods. The region between the black dotted lines is the main study region, where the model was trained in. In the summer (Figure 6-4(a, b)), the two-stage model outperforms PERSIANN-CCS and reduces the overestimation in the northwest and the

south areas, while both models show similar patterns of underestimation in the southeast area. In the winter (Figure 6-4(c, d)), the two-stage model reduces the overestimation in the middle to west area but introduces more underestimation in the east area of the map compared to PERSIANN-CCS. The Specific performance measures for both seasons are displayed in Table 6-4. Moreover, as presented in Table 6-6, compared to PERSIANN-CCS, 59.3% and 62.8% of the pixels show better performance in average bias with the two-stage model in the summer and winter seasons, respectively.

Figure 6-5 presents the average MSE of PERSIANN-CCS and the two-stage model precipitation over the large coverage of the United States averaged on the warm and cold verification periods. The region between the white dotted lines is the main study region, where the model was trained in. The warm colors indicate strong differences compared to Stage IV observations, while the cold colors indicate small differences. Figure 6-5(a, b) shows the significant decreases in MSE throughout the whole enlarged coverage for the two-stage model in summer season compared to PERSIANN-CCS. Besides the main study region, the central United States, significant improvements can be found in northwest and northeast regions. Specifically, MSEs are less than 3 $[mm/hr]^2$ for most pixels for the two-stage model (Figure 6-5(b)). Similarly, Figure 6-5(c, d) shows significant improvements of the two–stage model in the winter season, especially in the middle to west area, where the most pixels' MSE are less than 0.4 $[mm/hr]^2$ for the two-stage model

(Figure 6-5(d)). In the east, improvement can also be identified by the reduced area with MSEs larger than 0.8 [mm/hr]^2. In summary, a consistent and significant improvement can be found in MSE of the two-stage model in the enlarged coverage, compared to PERSIANN-CCS, which specific calculations are displayed in Table 6-5. Moreover, as presented in Table 6-6, compared to PERSIANN-CCS, nearly all (96.0% and 97.9% for the summer and winter seasons, respectively) pixels show better performance in average bias with the two-stage model.

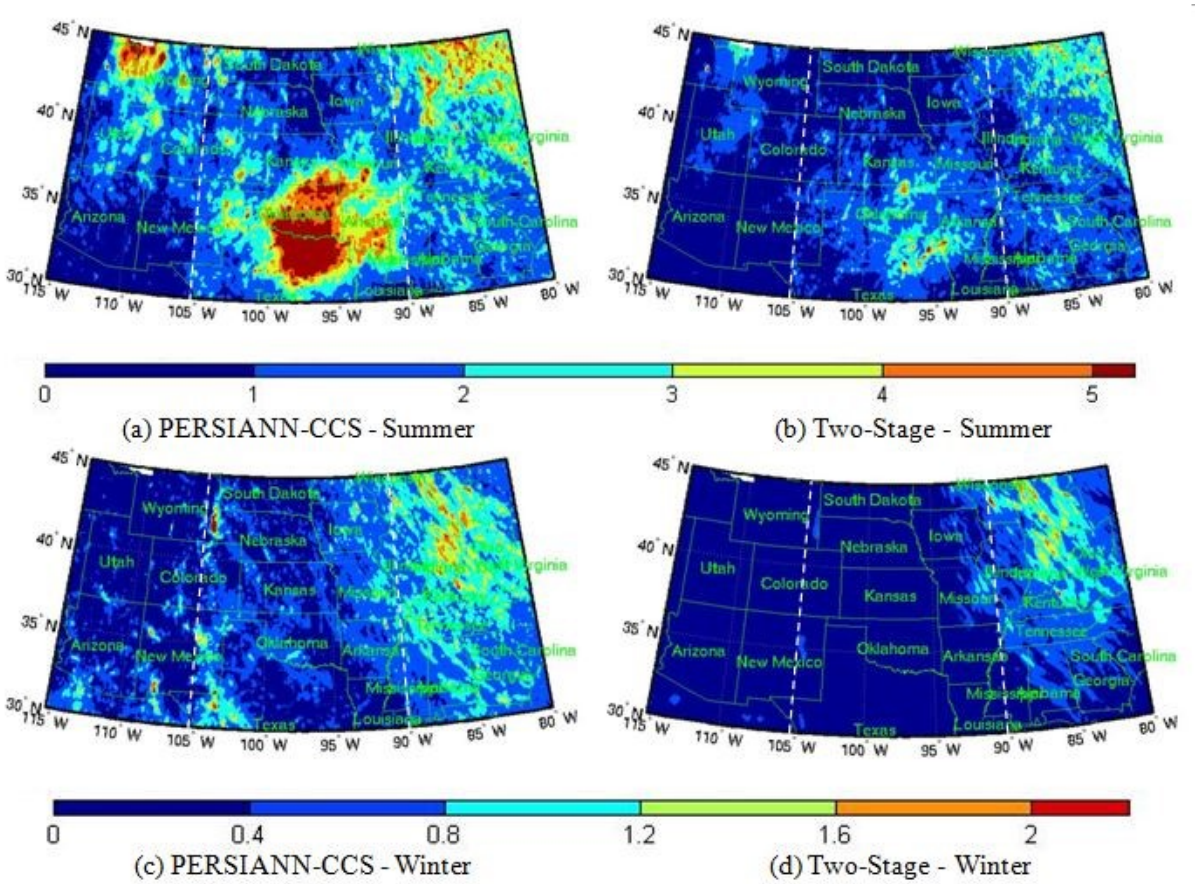

Figure 6-5 Average MSE ([mm/hr]^2) of PERSIANN-CCS [(a), (c)] and the two-stage model [(b), (d)] output over the large coverage of the United States (30–45°N, 85–115°W): (a), (b) summer (June–August 2013); (c), (d) winter (December 2013–February 2014)

Table 6-5 Summary of precipitation estimation performance of PERSIANN-CCS and the two-stage model over the large coverage of the United States for summer 2013 and winter 2013–2014, respectively

		PERSIANN-CCS	Two-stage Model
Summer Season (June–August 2013)			
Average Bias	Value	0.040	0.036
	Performance Gain	-	10.00%
Average MSE	Value	1.468	0.884
	Performance Gain	-	39.80%
COR	Value	0.319	0.388
	Performance Gain	-	5.96%
Winter Season (December 2013–February 2014)			
Average Bias	Value	0.055	0.036
	Performance Gain	-	34.55%
Average MSE	Value	0.560	0.240
	Performance Gain	-	57.14%
COR	Value	0.218	0.311
	Performance Gain	-	42.66%

Table 6-5 provides detailed values for of average bias, average MSE, and COR for PERSIANN-CCS and the two-stage model in different seasons over the large coverage of the United States. There are 10.00% and 34.55% increases in the average biases for the two-stage model for summer and winter seasons compared to PERSIANN-CCS (0.036 compared to 0.040 and 0.036 compared to 0.055, respectively). More obvious

improvement can be found in MSEs, where the two-stage model has 39.80% and 57.14% decreases in summer and winter season, respectively. In addition, compared to PERSIANN-CCS, the COR for the two-stage model is 57.14% higher in the winter season and 5.96% higher in the summer season. Similar to the results presented in Table 5-3 for the main study region, the overall performance gain is higher for winter season than summer season, which shows the consistency of the model when applied to unfamiliar regions.

Table 6-6 Percentage of pixels with better performance in the two-stage model, compared to PERSIANN-CCS model over the large coverage of the United States for summer 2013 and winter 2013–2014, respectively

	Summer	**Winter**
Average Bias	59.3%	62.8%
Average MSE	96.0%	97.9%

6.3 Conclusion

This chapter presented the results of the application of the developed two-stage deep learning framework for precipitation estimation from bispectral satellite, introduced in Chapter 4 and Chapter 5, over a large area. The objective was to evaluate its reliability and consistency for areas outside the original region where the two-stage model was developed and trained. Specifically, I directly apply the previously well-trained two-stage model to a

larger coverage of the United States (30–45°N, 85–115°W), where reliable Stage IV data are available. This large-scale evaluation is particularly important for the future operational global application of the model.

The experiments show significant improvements for the two-stage model in both R/NR classification and precipitation amount estimation compared to PERSIANN-CCS at the enlarged coverage of the United States. In R/NR classification, the performance gain in CSI for the two-stage model is 31.50% compared to PERSIANN-CCS. In precipitation amount estimation, the two-stage model is 23.40% lower in average bias, 44.52% lower in average MSE, and 27.21% higher in COR. In addition, the improvement is greater for the winter season than the summer season. These results are highly consistent with the findings presented in Chapter 4 and Chapter 5 for the main study region, though the performance gains are generally slightly lower.

The experiment results demonstrate the effectiveness of the two-stage deep learning framework for precipitation estimation from bispectral satellite information over unfamiliar areas in R/NR classification and real-valued amount estimation. The model can capture the relationship between the satellite information and the precipitation regardless of geographic locations, which is an important prerequisite for a model to serve as a satellite-based precipitation estimation product with global coverage.

The current application is an initial step for the large-scale application. Further experiments are extremely important for the preparation of the model to serve as an operation product. For example, it is necessary to calibrate the parameters with samples from a larger study region, such as the continental United States, to better capture the variability in the forms of precipitation. The model should then evaluate multiple locations, within the United States and abroad, to better evaluate the global performance of the model.

Chapter 7

Summary and Conclusions

7.1 Dissertation Summary

Compared to traditional ground-based precipitation measurements, near-real-time satellite-based precipitation estimation products have the advantage of providing instant global precipitation estimation at high-spatiotemporal resolution. However, the undesired accuracy and stability of satellite-based precipitation estimation products prevent them from serving as the main measurements for weather, climate, and hydrologic applications. In developing satellite-based precipitation estimation products, it is necessary to incorporate satellite data with sufficient precipitation-related information and use effective methodologies to extract such information and eventually provide accurate precipitation estimates.

In this dissertation, a state-of-the-art deep learning framework for precipitation estimation from bispectral satellite information, the IR and WV channels, was developed. I first verified the effectiveness of deep learning techniques in extracting useful features from the

satellite information by applying them toward operational satellite products as a bias reduction model without adding extra information. I then designed a two-stage deep learning framework for direct precipitation estimation from the bispectral satellite information. The first stage is binary precipitation identification, serving to optimize objectives for delineating precipitation regions and eliminate the massive amount of NR pixels. The second stage is non-zero precipitation amount estimation, serving to estimate the point-wise precipitation amount accurately while preserving its highly skewed distribution. For both stages, I evaluated the models with various common performance measurements and compared their performances with PERSIANN-CCS, an operational satellite-based precipitation estimation product. In addition, I explored the two-stage model's performance at large scale to demonstrate its potential to be further developed to an operational product with global coverage.

7.2 Conclusions

At a high level, the main contributions of this dissertation are as follows:

1) Introduced state-of-the-art deep learning techniques to the field of satellite-based precipitation estimation. Moreover, it is concluded that these

data-driven methodologies can benefit many fields of weather forecasting, climate variability, hydrology, and water resource management.

2) Developed a deep learning precipitation estimation model that has the potential to serve as a near-real-time operational product at global coverage with high-spatiotemporal resolution.

3) Provided an effective framework, the two-stage structure, for precipitation estimation using data-driven methods that can handle the highly skewed distribution of precipitation and fulfill the needs of multiple objectives, which offered many possibilities for future works.

Specifically, the following objectives outlined in Chapter 1 are addressed throughout the dissertation:

1) Demonstrate the effectiveness of deep learning algorithms for precipitation-related information extraction from satellite infrared imageries by developing a bias reduction system on satellite-based precipitation estimation products.

In Chapter 3, the bias reduction model for satellite-based precipitation estimation products without adding extra data sources was presented. The experiments show the capability of deep learning techniques to extract additional useful information from the

raw input data and help reduce bias in the precipitation estimation. In this process, advantage was taken of the R/NR prediction made by the product to reduce the large amount of NR pixels and focused on reducing false alarms and correcting biases in pixels with precipitation.

The main contribution of this study is to provide preliminary verification of the effectiveness of DNNs in the application of satellite-based precipitation estimation. It was also demonstrated that IR imageries still contain valuable information for precipitation retrieval. Additionally, a bias reduction model that may post-process operational products for better estimation was provided.

2) Design a deep learning framework that extracts useful features from bispectral satellite information to produce a consistent and high-quality precipitation estimation product.

In Chapter 4 and Chapter 5, details of a two-stage deep learning framework for precipitation estimation from bispectral satellite information (IR and WV channels) were introduced. This is the primary work of this dissertation. The first stage of the model is an R/NR binary classification process. Pixels classified with precipitation are then fed into the second stage of the model, which estimates the real-valued precipitation amount. At both stages, DNNs serve as the basic structure to extract

precipitation-related information from IR and WV channels and connect it to precipitation identification and amount estimation.

This two-stage prototype is designed to tackle the challenge of the imbalance of precipitation data and to fulfill the needs of multiple objectives for precipitation estimation, including accurate R/NR classification, minimal error in point-wise estimates, and similarly distributed of the skewed precipitation observations. This prototype also provides many opportunities to apply other advanced machine learning models and further improve the performance of the estimation.

3) Evaluate the effectiveness of the methodology and the value of additional satellite information on both binary precipitation identification and precipitation amount estimation by comparing its performance with operational products.

In Chapter 4 and Chapter 5, the results of experiments designed to evaluate the two-stage model were presented. Common performance measurements for both binary precipitation identification and precipitation amount estimation were produced, and all measurements were compared with PERSIANN-CCS, a commonly used operational product. In addition, multiple case studies to explore the model's actual performances for specific events were presented.

In R/NR classification, a 43.66% performance gain can be detected in CSI over the verification periods for the two-stage model compared to PERSIANN-CCS. In real-valued precipitation amount estimation, the two-stage model is 40.58% lower in average bias, 53.95% lower in average MSE, and 23.49% higher in COR over the verification periods. These values prove the effectiveness of the two-stage deep learning framework for precipitation estimation from bispectral satellite information in both R/NR classification and real-value amount estimation.

4) Evaluate the performance of the developed DNN model at a large-scale application and thus demonstrate its potential of future use at a near-global scale.

As a necessary step towards an operational product with global coverage, the results of preliminary large-scale applications covering an expanded region of the United States (30–45°N, 85–115°W) were presented in Chapter 6. The results still show significant improvements for the two-stage model in both R/NR classification and precipitation amount estimation compared to PERSIANN-CCS.

In R/NR classification, a 32.56% performance gain can be detected in CSI for the two-stage model compared to PERSIANN-CCS. In real-valued precipitation amount estimation, the two-stage model is 23.40% lower in average bias, 44.52% lower in average MSE, and 27.21% higher in COR. These values are consistent with the findings for the main study region, though the performance gains are generally slightly lower.

They demonstrated the effectiveness of the two-stage model, in both R/NR classification and real-valued amount estimation, outside the original region where the model was developed and trained,

7.3 Future Directions

As mentioned in Section 7.1, the two-stage deep learning framework developed in this dissertation offers many potential directions for future research. Improvements can come from research in three areas: additional data sources, further model development, and operational application.

7.3.1 Additional Data Sources

As demonstrated in Chapter 4, additional satellite channels, such as WV channel, can provide additional precipitation-related information and help improve the precipitation estimation performance. There are other satellite channels, such as the VIS channel, that has been proven to be effective in improving the accuracy of precipitation estimation (Behrangi et al., 2010; Capacci and Conway, 2005; Zinner et al., 2008). Moreover, there are many other satellite datasets available that provide valuable information for precipitation estimation, such as Global Precipitation Measurement (GPM), Moderate resolution Imaging

Spectroradiometer (MODIS), and CloudSat (Hou et al., 2008; Jensen, 2007; Weisz et al., 2007).

The additional datasets should not be limited to satellite datasets when other data sources are also beneficial. For example, Numerical Weather Prediction (NWP) models provide predictions on many related factors, including short-term precipitation forecast, 10-meter vertical velocity, relative humidity, and convective available potential energy (Ghil et al., 1981; Lorenc, 1986; Shapiro and Thorpe, 2004; Thielen et al., 2008). In addition, surface physical characteristics, such as digital elevation of land surface, can be very helpful to better capture the rainfall in the mountainous areas (Gesch et al., 1999; Zhang and Montgomery, 1994).

On the other hand, there may be some improvements by using the ground measurements to serve in calibrating the parameters. For example, the hourly satellite snapshots and hourly cumulative radar observations may not be the best pairs to draw a relationship. One way to solve this mismatch is upgrading the ground-based observation data to a higher temporal resolution for experiments, such as the new 5-min Next-Generation Radar (NEXRAD) data (Heiss et al., 1990; Zhang et al., 2011). Meanwhile, mountainous areas may need more careful considerations in the choices of ground measurements for the model to be accurate (Daly et al., 1994; Germann et al., 2009).

7.3.2 Model Development

Throughout this dissertation, SDAEs, which are described in Section 2.3, served as the basic model structures for both stages of the two-stage model. With the rapid development in the field of deep learning, many other effective neural network structures were developed that might fit our problem even better. Convolutional neural networks (CNNs), a type of neural network considering connectivity pattern between its neurons, are widely used in image and video recognition (Donahue et al., 2014; Krizhevsky et al., 2012; LeCun et al., 1995; Simonyan and Zisserman, 2014). Its capability of extracting high-level patterns from images is suitable for dealing with satellite imageries. Recurrent neural networks (RNNs), taking dynamic temporal behavior into account, is also a design with great potential for this problem since the temporal information is critical in the formation of precipitation (Barbounis et al., 2006; Funahashi and Nakamura, 1993; Lukoševičius and Jaeger, 2009).

Additionally, there is room for further investigation and development in the design and calibration process of the neural networks. For instance, more experiments on parameter tuning might lead to an improved model, because neural networks are often combined with many hyper-parameters. In particular, activation functions can be further designed and tested, which can lead to improved accuracy in other problems in computer vision (Chen and Manning, 2014; Hornik, 1991; Maas et al., 2013).

In addition, probabilistic graphic models (PGMs) also have the potential to serve as base models for the two-stage framework, at least for the classification stage. Specifically, conditional random fields (CRFs) provide structured predictions instead of independently treating each pixel (Ihler et al., 2005; Ping et al., 2014; Sudderth et al., 2010), and this might be a great fit in delineating the precipitation regions.

Another fundamental method to improve precipitation estimation is to design the integrated model's performance measures or objective functions. As discussed in Section 2.1, it is hard to define good performance for a precipitation estimation model due to the imbalance nature of precipitation data and the needs of multiple objectives for precipitation estimation products. However, a properly designed performance measure is essential for comparing different products. One preliminary idea for the measurement is the earth mover's distance (EMD), which measures the distance between two probability distributions over a region D (Levina and Bickel, 2001; Rubner et al., 2000).

7.3.3 Operational Application

As I discussed in Chapter 6, additional experiments are needed to prepare the model for a near-real-time operational precipitation estimation product. Studies should be conducted on a larger region with longer training and verification periods to investigate the stability of the model. For example, one could obtain some ground observations from other

countries and geographic regions to train and evaluate at a global scale, instead of limiting it to the United States.

Moreover, it is necessary to test the model's estimates in hydrological applications, which is one of the main uses of satellite-based precipitation estimation products. The model's performance in hydrological models should be compared with other current operational products. In addition, it is a good idea to investigate the model's performance on extreme events, which is essential for disaster monitoring and preparation.

REFERENCES

Adler, R. F., Kidd, C., Petty, G., Morissey, M., & Goodman, H. M. (2001). Intercomparison of global precipitation products: The third precipitation intercomparison project (PIP–3). *Bulletin of the American Meteorological Society*, *82*(7), 1377-1396.

AghaKouchak, A., & Nakhjiri, N. (2012). A near real-time satellite-based global drought climate data record. *Environmental Research Letters*, *7*(4), 44037-44044.

AghaKouchak, A., Nasrollahi, N., Li, J., Imam, B., & Sorooshian, S. (2011). Geometrical characterization of precipitation patterns. *Journal of Hydrometeorology*, *12*(2), 274-285.

Ajami, N. K., Hornberger, G. M., & Sunding, D. L. (2008). Sustainable water resource management under hydrological uncertainty. *Water Resources Research*, *44*(11).

Anderson, J., Chung, F., Anderson, M., Brekke, L., Easton, D., Ejeta, M., ... & Snyder, R. (2008). Progress on incorporating climate change into management of California's water resources. *Climatic Change*, *87*, 91-108.

Arkin, P. A., & Meisner, B. N. (1987). The relationship between large-scale convective rainfall and cold cloud over the western hemisphere during 1982-84. *Monthly Weather Review*, *115*(1), 51-74.

Arking, A., & Childs, J. D. (1985). Retrieval of cloud cover parameters from multispectral satellite images. *Journal of Climate and Applied Meteorology*, *24*(4), 322-333.

Ba, M. B., & Gruber, A. (2001). GOES multispectral rainfall algorithm (GMSRA). *Journal of Applied Meteorology*, *40*(8), 1500-1514.

Barbounis, T. G., Theocharis, J. B., Alexiadis, M. C., & Dokopoulos, P. S. (2006). Long-term wind speed and power forecasting using local recurrent neural network models. *IEEE Transactions on Energy Conversion, 21*(1), 273-284.

Behrangi, A., Hsu, K. L., Imam, B., Sorooshian, S., & Kuligowski, R. J. (2009a). Evaluating the utility of multispectral information in delineating the areal extent of precipitation. *Journal of Hydrometeorology*, *10*(3), 684-700.

Behrangi, A., Hsu, K. L., Imam, B., Sorooshian, S., Huffman, G. J., & Kuligowski, R. J. (2009b). PERSIANN-MSA: A precipitation estimation method from satellite-based multispectral analysis. *Journal of Hydrometeorology*, *10*(6), 1414-1429.

Behrangi, A., Imam, B., Hsu, K., Sorooshian, S., Bellerby, T. J., & Huffman, G. J. (2010). REFAME: Rain estimation using forward-adjusted advection of microwave estimates. *Journal of Hydrometeorology*, *11*(6), 1305-1321.

Bellerby, T. J., & Sun, J. (2005). Probabilistic and ensemble representations of the uncertainty in an IR/microwave satellite precipitation product. *Journal of Hydrometeorology*, *6*(6), 1032-1044.

Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and trends® in Machine Learning*, *2*(1), 1-127.

Bengio, Y., Lamblin, P., Popovici, D., & Larochelle, H. (2007). Greedy layer-wise training of deep networks. *Advances in neural information processing systems*, *19*, 153.

Boushaki, F. I., Hsu, K. L., Sorooshian, S., Park, G. H., Mahani, S., & Shi, W. (2009). Bias adjustment of satellite precipitation estimation using ground-based measurement: A case study evaluation over the southwestern United States. *Journal of Hydrometeorology*, *10*(5), 1231-1242.

Capacci, D., & Conway, B. J. (2005). Delineation of precipitation areas from MODIS visible and infrared imagery with articial neural networks. *Meteorological Applications*, *12*(4), 291-305.

Chen, D., & Manning, C. D. (2014, October). A Fast and Accurate Dependency Parser using Neural Networks. In *Proceeding of Conference on Empirical Methods in Natural Language Processing* (pp. 740-750).

Cheng, M., Brown, R., & Collier, C. C. (1993). Delineation of precipitation areas using Meteosat infrared and visible data in the region of the United Kingdom. *Journal of Applied Meteorology*, *32*(5), 884-898.

Ciregan, D., Meier, U., & Schmidhuber, J. (2012, June). Multi-column deep neural networks for image classification. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on* (pp. 3642-3649). IEEE.

Daly, C., Neilson, R. P., & Phillips, D. L. (1994). A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of applied meteorology*, 33(2), 140-158.

Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E., & Darrell, T. (2014). DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition. In *Proceedings of the 31st International Conference on Machine Learning (ICML-14)* (pp. 647-655).

Ebert, E. E., Manton, M. J., Arkin, P. A., Allam, R. J., Holpin, C. E., & Gruber, A. (1996). Results from the GPCP algorithm intercomparison programme. *Bulletin of the American Meteorological Society*, *77*(12), 2875-2887.

Funahashi, K. I., & Nakamura, Y. (1993). Approximation of dynamical systems by continuous time recurrent neural networks. *Neural networks, 6*(6), 801-806.

Germann, U., Berenguer, M., Sempere-Torres, D., & Zappa, M. (2009). REAL—Ensemble radar precipitation estimation for hydrology in a mountainous region. *Quarterly Journal of the Royal Meteorological Society, 135*(639), 445-456.

Gesch, D. B., Verdin, K. L., & Greenlee, S. K. (1999). New land surface digital elevation model covers the Earth. *EOS, Transactions American Geophysical Union*, *80*(6), 69-70.

Ghil, M., Cohn, S., Tavantzis, J., Bube, K., & Isaacson, E. (1981). Applications of estimation theory to numerical weather prediction. In *Dynamic meteorology: Data assimilation methods* (pp. 139-224). Springer New York.

Glorot, X., & Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *International Conference on Artificial Intelligence and Statistics* (pp. 249-256).

Glorot, X., Bordes, A., & Bengio, Y. (2011a, April). Deep Sparse Rectifier Neural Networks. In *14th International Conference on Artificial Intelligence and Statistics* (Vol. 15, pp. 315-323).

Glorot, X., Bordes, A., & Bengio, Y. (2011b). Domain adaptation for large-scale sentiment classification: A deep learning approach. In *Proceedings of the 28th international conference on machine learning (ICML-11)* (pp. 513-520)..

Grassotti, C., & Garand, L. (1994). Classification-based rainfall estimation using satellite data and numerical forecast model fields. *Journal of Applied Meteorology*, *33*(2), 159-178.

Griffith, C. G., Woodley, W. L., Grube, P. G., Martin, D. W., Stout, J., & Sikdar, D. N. (1978). Rain estimation from geosynchronous satellite imagery—Visible and infrared studies. *Monthly Weather Review*, *106*(8), 1153-1171.

Habib, E., Haile, A. T., Sazib, N., Zhang, Y., & Rientjes, T. (2014). Effect of bias correction of satellite-rainfall estimates on runoff simulations at the source of the Upper Blue Nile. *Remote Sensing*, *6*(7), 6688-6708.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision* (pp. 1026-1034).

Heiss, W. H., McGrew, D. L., & Sirmans, D. (1990). NEXRAD: next generation weather radar (WSR-88D). *Microwave Journal, 33*(1), 79-89.

Hinton, G. E., & Zemel, R. S. (1994). Autoencoders, Minimum Description Length and Helmholtz Free Energy. In *Advances in Neural Information Processing Systems* (pp. 3-10).

Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. *Neural computation*, *18*(7), 1527-1554.

Hong, Y., Hsu, K. L., Sorooshian, S., & Gao, X. (2004). Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *Journal of Applied Meteorology*, *43*(12), 1834-1853.

Hong, Y., Gochis, D., Cheng, J. T., Hsu, K. L., & Sorooshian, S. (2007). Evaluation of PERSIANN-CCS rainfall measurement using the NAME event rain gauge network. *Journal of Hydrometeorology*, *8*(3), 469-482.

Hornik, K. (1991). Approximation capabilities of multilayer feedforward networks. *Neural networks*, *4*(2), 251-257.

Hossain, M., Rekabdar, B., Louis, S. J., & Dascalu, S. (2015, July). Forecasting the weather of Nevada: A deep learning approach. In *Neural Networks (IJCNN), 2015 International Joint Conference on* (pp. 1-6). IEEE.

Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., ... & Iguchi, T. (2014). The global precipitation measurement mission. *Bulletin of the American Meteorological Society*, *95*(5), 701-722.

Hsu, K. L., Gupta, H. V., Gao, X., & Sorooshian, S. (1999). Estimation of physical variables from multichannel remotely sensed imagery using a neural network: Application to rainfall estimation. *Water Resources Research*, *35*(5), 1605-1618.

Hsu, K. L., Gao, X., Sorooshian, S., & Gupta, H. V. (1997). Precipitation estimation from remotely sensed information using artificial neural networks. *Journal of Applied Meteorology*, *36*(9), 1176-1190.

Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., ... & Stocker, E. F. (2007). The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology*, *8*(1), 38-55.

Ihler, A. T., John III, W. F., & Willsky, A. S. (2005). Loopy belief propagation: Convergence and effects of message errors. *Journal of Machine Learning Research, 6*(May), 905-936.

Inoue, T., & Aonashi, K. (2000). A comparison of cloud and rainfall information from instantaneous visible and infrared scanner and precipitation radar observations over

a frontal zone in East Asia during June 1998. *Journal of Applied Meteorology*, *39*(12), 2292-2301.

Jensen, R. (2007). The digital provide: Information (technology), market performance, and welfare in the South Indian fisheries sector. *The quarterly journal of economics*, *122*(3), 879-924.

Joyce, R. J., Janowiak, J. E., Arkin, P. A., & Xie, P. (2004). CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology*, *5*(3), 487-503.

Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., & Fei-Fei, L. (2014). Large-scale video classification with convolutional neural networks. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition* (pp. 1725-1732).

Kidd, C., Kniveton, D. R., Todd, M. C., & Bellerby, T. J. (2003). Satellite rainfall estimation using combined passive microwave and infrared algorithms. *Journal of Hydrometeorology*, *4*(6), 1088-1104.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).

Kuligowski, R. J. (2002). A self-calibrating real-time GOES rainfall algorithm for short-term rainfall estimates. *Journal of Hydrometeorology*, *3*(2), 112-130.

Kummerow, C., & Giglio, L. (1995). A method for combining passive microwave and infrared rainfall observations. *Journal of Atmospheric and Oceanic Technology*, *12*(1), 33-45.

Kurino, T. (1997). A satellite infrared technique for estimating "deep/shallow" precipitation. *Advances in Space Research*, *19*(3), 511-514.

Le, Q. V., Ngiam, J., Coates, A., Lahiri, A., Prochnow, B., Le, Q. V., & Ng, A. Y. (2011). On optimization methods for deep learning. In *Proceedings of the 28th International Conference on Machine Learning (ICML-11)* (pp. 265-272).

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.

LeCun, Y., Jackel, L. D., Bottou, L., Cortes, C., Denker, J. S., Drucker, H., ... & Vapnik, V. (1995). Learning algorithms for classification: A comparison on handwritten digit recognition. *Neural networks: the statistical mechanics perspective*, *261*, 276.

Lee, H., Ekanadham, C., & Ng, A. Y. (2008). Sparse deep belief net model for visual area V2. In *Advances in neural information processing systems* (pp. 873-880).

Lensky, I. M., & Rosenfeld, D. (2003). A night-rain delineation algorithm for infrared satellite data based on microphysical considerations. *Journal of Applied Meteorology*, *42*(9), 1218-1226.

Levina, E., & Bickel, P. (2001). The earth mover's distance is the mallows distance: Some insights from statistics. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on* (Vol. 2, pp. 251-256). IEEE.

Levizzani, V., & Setvák, M. (1996). Multispectral, high-resolution satellite observations of plumes on top of convective storms. *Journal of the atmospheric sciences*, *53*(3), 361-369.

Li, Z., Li, J., Menzel, W. P., Schmit, T. J., & Ackerman, S. A. (2007). Comparison between current and future environmental satellite imagers on cloud classification using MODIS. *Remote Sensing of Environment*, *108*(3), 311-326.

Lin, Y., & Mitchell, K. E. (2005). 1.2 the NCEP stage II/IV hourly precipitation analyses: Development and applications. In *19th Conf. Hydrology*.

Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, *112*(474), 1177-1194.

Lovejoy, S., & Austin, G. L. (1979). The delineation of rain areas from visible and IR satellite data for GATE and mid-latitudes. *Atmosphere-Ocean*, *17*(1), 77-92.

Lu, X., Tsao, Y., Matsuda, S., & Hori, C. (2013, August). Speech enhancement based on deep denoising autoencoder. In *Interspeech* (pp. 436-440).

Lu, Y., Zhang, L., Wang, B., & Yang, J. (2014, July). Feature ensemble learning based on sparse autoencoders for image classification. In *Neural Networks (IJCNN), 2014 International Joint Conference on* (pp. 1739-1745). IEEE.

Lukoševičius, M., & Jaeger, H. (2009). Reservoir computing approaches to recurrent neural network training. Computer Science Review, 3(3), 127-149.

Maas, A. L., Hannun, A. Y., & Ng, A. Y. (2013, June). Rectifier nonlinearities improve neural network acoustic models. In *Proceedings of the 30th International Conference on Machine Learning (ICML-13)* (Vol. 30, No. 1).

Martin, D. W., Kohrs, R. A., Mosher, F. R., Medaglia, C. M., & Adamo, C. (2008). Over-ocean validation of the global convective diagnostic. *Journal of Applied Meteorology and Climatology*, *47*(2), 525-543.

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. *IEEE transactions on Geoscience and remote sensing*, *42*(5), 1018-1032.

McCollum, J. R., Krajewski, W. F., Ferraro, R. R., & Ba, M. B. (2002). Evaluation of biases of satellite rainfall estimation algorithms over the continental United States. *Journal of Applied Meteorology*, *41*(11), 1065-1080.

Moazami, S., Golian, S., Kavianpour, M. R., & Hong, Y. (2014). Uncertainty analysis of bias from satellite rainfall estimates using copula method. *Atmospheric Research*, *137*, 145-166.

Nasrollahi, N., Hsu, K., & Sorooshian, S. (2013). An artificial neural network model to reduce false alarms in satellite precipitation products using MODIS and CloudSat observations. *Journal of Hydrometeorology*, *14*(6), 1872-1883.

Netzer, Y., Wang, T., Coates, A., Bissacco, A., Wu, B., & Ng, A. Y. (2011, December). Reading digits in natural images with unsupervised feature learning. In *NIPS workshop on deep learning and unsupervised feature learning* (Vol. 2011, No. 2, p. 5).

Nguyen, P., Sellars, S., Thorstensen, A., Tao, Y., Ashouri, H., Braithwaite, D., ... & Sorooshian, S. (2014). Satellites track precipitation of super Typhoon Haiyan. *Eos, Transactions American Geophysical Union*, *95*(16), 133-135.

O'Sullivan, F., Wash, C. H., Stewart, M., & Motell, C. E. (1990). Rain estimation from infrared and visible GOES satellite data. *Journal of Applied Meteorology*, *29*(3), 209-223.

Ping, W., Liu, Q., & Ihler, A. T. (2014, September). Marginal Structured SVM with Hidden Variables. In *Proceedings of the 31th International Conference on Machine Learning (ICML-14)* (pp. 190-198).

Ranzato, M. A., Boureau, Y.-L., & Cun, Y. L. (2008). Sparse feature learning for deep belief networks. In *Advances in neural information processing systems* (pp. 1185-1192).

Pilewskie, P., & Twomey, S. (1987). Cloud phase discrimination by reflectance measurements near 1.6 and 2.2 μm. *Journal of the atmospheric sciences*, *44*(22), 3419-3420.

Rosenfeld, D., & Gutman, G. (1994). Retrieving microphysical properties near the tops of potential rain clouds by multispectral analysis of AVHRR data. *Atmospheric research, 34*(1-4), 259-283.

Rubner, Y., Tomasi, C., & Guibas, L. J. (2000). The earth mover's distance as a metric for image retrieval. *International journal of computer vision*, *40*(2), 99-121.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1988). Learning representations by back-propagating errors. *Cognitive modeling*, *5*(3), 1.

Scofield, R. A., & Kuligowski, R. J. (2003). Status and outlook of operational satellite precipitation algorithms for extreme-precipitation events. *Weather and Forecasting, 18*(6), 1037-1051.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Sorooshian, S., AghaKouchak, A., Arkin, P., Eylander, J., Foufoula-Georgiou, E., Harmon, R., ... & Skofronick-Jackson, G. (2011). Advanced concepts on remote sensing of precipitation at multiple scales. *Bulletin of the American Meteorological Society*, *92*(10), 1353-1357.

Sudderth, E. B., Ihler, A. T., Isard, M., Freeman, W. T., & Willsky, A. S. (2010). Nonparametric belief propagation. *Communications of the ACM*, *53*(10), 95-103.

Sun, Y., Wang, X., & Tang, X. (2014). Deep learning face representation from predicting 10,000 classes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1891-1898).

Shapiro, M. A., & Thorpe, A. J. (2004). THORPEX international science plan. *WMO/TD*, *1246*.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1-9).

Tao, Y., Gao, X., Hsu, K., Sorooshian, S., & Ihler, A. (2016a). A deep neural network modeling framework to reduce bias in satellite precipitation products. *Journal of Hydrometeorology*, *17*(3), 931-945.

Tao, Y., Gao, X., Ihler, A., Hsu, K., & Sorooshian, S. (2016b, July). Deep neural networks for precipitation estimation from remotely sensed information. In *Evolutionary Computation (CEC), 2016 IEEE Congress on* (pp. 1349-1355). IEEE.

Tao, Y., Gao, X., Ihler, A., Sorooshian, S., & Hsu, K. (2017). Precipitation Identification with Bispectral Satellite Information Using Deep Learning Approaches. *Journal of Hydrometeorology*, (2017).

Thielen, J., Schaake, J., Hartman, R., & Buizza, R. (2008). Aims, challenges and progress of the Hydrological Ensemble Prediction Experiment (HEPEX) following the third HEPEX workshop held in Stresa 27 to 29 June 2007. *Atmospheric Science Letters*, *9*(2), 29-35.

Tian, Y., Peters-Lidard, C. D., Eylander, J. B., Joyce, R. J., Huffman, G. J., Adler, R. F., ... & Zeng, J. (2009). Component analysis of errors in satellite-based precipitation estimates. *Journal of Geophysical Research: Atmospheres*, *114*(D24).

Tjemkes, S. A., Van de Berg, L., & Schmetz, J. (1997). Warm water vapour pixels over high clouds as observed by METEOSAT. *Contributions to atmospheric physics*, *70*(1), 15-21.

Tsonis, A. A., & Isaac, G. A. (1985). On a new approach for instantaneous rain area delineation in the midlatitudes using GOES data. *Journal of climate and applied meteorology*, *24*(11), 1208-1218.

Turk, F. J., & Miller, S. D. (2005). Toward improved characterization of remotely sensed precipitation regimes with MODIS/AMSR-E blended data techniques. *IEEE Transactions on Geoscience and Remote Sensing*, *43*(5), 1059-1069.

Vincent, P., Larochelle, H., Bengio, Y., & Manzagol, P. A. (2008, July). Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th international conference on Machine learning* (pp. 1096-1103). ACM.

Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., & Manzagol, P. A. (2010). Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *Journal of Machine Learning Research*, *11*(Dec), 3371-3408.

Wang, N., & Yeung, D. Y. (2013). Learning a deep compact image representation for visual tracking. In *Advances in neural information processing systems* (pp. 809-817).

Weisz, E., Li, J., Menzel, W. P., Heidinger, A. K., Kahn, B. H., & Liu, C. Y. (2007). Comparison of AIRS, MODIS, CloudSat and CALIPSO cloud top height retrievals. *Geophysical Research Letters*, *34*(17).

Weng, F., Zhao, L., Ferraro, R. R., Poe, G., Li, X., & Grody, N. C. (2003). Advanced microwave sounding unit cloud and precipitation algorithms. *Radio Science*, *38*(4).

Xie, J., Xu, L., & Chen, E. (2012). Image denoising and inpainting with deep neural networks. In *Advances in Neural Information Processing Systems* (pp. 341-349).

Zeiler, M. D., Ranzato, M., Monga, R., Mao, M., Yang, K., Le, Q. V., ... & Hinton, G. E. (2013, May). On rectified linear units for speech processing. In *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on* (pp. 3517-3521). IEEE.

Zhang, J., Howard, K., Langston, C., Vasiloff, S., Kaney, B., Arthur, A., ... & Seo, D. J. (2011). National Mosaic and Multi-Sensor QPE (NMQ) system: Description, results, and future plans. *Bulletin of the American Meteorological Society*, *92*(10), 1321-1338.

Zhang, W., & Montgomery, D. R. (1994). Digital elevation model grid size, landscape representation, and hydrologic simulations. *Water resources research*, *30*(4), 1019-1028.

Zhao, J., Gong, M., Liu, J., & Jiao, L. (2014, July). Deep learning to classify difference image for image change detection. In *Neural Networks (IJCNN), 2014 International Joint Conference on* (pp. 411-417). IEEE.

Zhou, G., Sohn, K., & Lee, H. (2012). Online incremental feature learning with denoising autoencoders. *Ann Arbor*, *1001*, 48109.

Zinner, T., Mannstein, H., & Tafferner, A. (2008). Cb-TRAM: Tracking and monitoring severe convection from onset over rapid development to mature phase using multi-channel Meteosat-8 SEVIRI data. *Meteorology and Atmospheric Physics*, *101*(3), 191-210.

Appendix A

Reproduced PERSIANN-CCS

In this section, I provide details on the results of reproduced PERSIANN-CCS for the main study regions and periods and compare it with the operational PRSIANN-CCS to demonstrate its methodology's stability. The methodology is described in section 2.2. And the training and verification data are described in section 2.4. The verification measurements are introduced in section 2.1.

Table A-1 provides performance measurements of the operational PERSIANN-CCS and the reproduced PERSIANN-CCS over the verification periods. For classification, the operational PERSIANN-CCS performs slightly better than the reproduced PERSIANN-CCS (0.213 and 0.202 in CSI for the operational and reproduced PERSIANN-CCS, respectively). For regression, the performances of the average MSEs are similar, which demonstrates the stability of the PERSIANN-CCS algorithm. The reproduced PERSIANN-CCS outperforms the operational one in average bias (0.069 and 0.041 for the operational and reproduced PERSIANN-CCS, respectively) while its performance in COR is worse (0.332 and 0.281 for the operational and reproduced PERSIANN-CCS, respectively). The overall performances

are comparable, which demonstrates the stability of the algorithm for difference observation data.

Table A-1 Summary of performance over the verification periods for the operational and the reproduced PERSIANN-CCS (including both summer 2013 and winter 2013–2014)

		PERSIANN-CCS (Operational)	**PERSIANN-CCS (Reproduced)**
Classification	**POD**	0.391	0.390
	FAR	0.681	0.705
	CSI	0.213	0.202
Regression	**Average Bias**	0.069	0.041
	Average MSE	1.392	1.324
	COR	0.332	0.281

Figure A-1 provides the performance of CSI for the operational and the reproduced PERSIANN-CCS in the validation periods for both winter and summer seasons. As Figure A-1 shown, the geophysical distribution of the performances are very similar to each other for both seasons for the operational and the reproduced PERSIANN-CCS. For the summer season, the performance of the reproduced PERSIANN-CCS is slightly worse than the original PERSIANN-CCS. For the winter season, the performances of the original and the reproduced PERSIANN-CCS are quite similar.

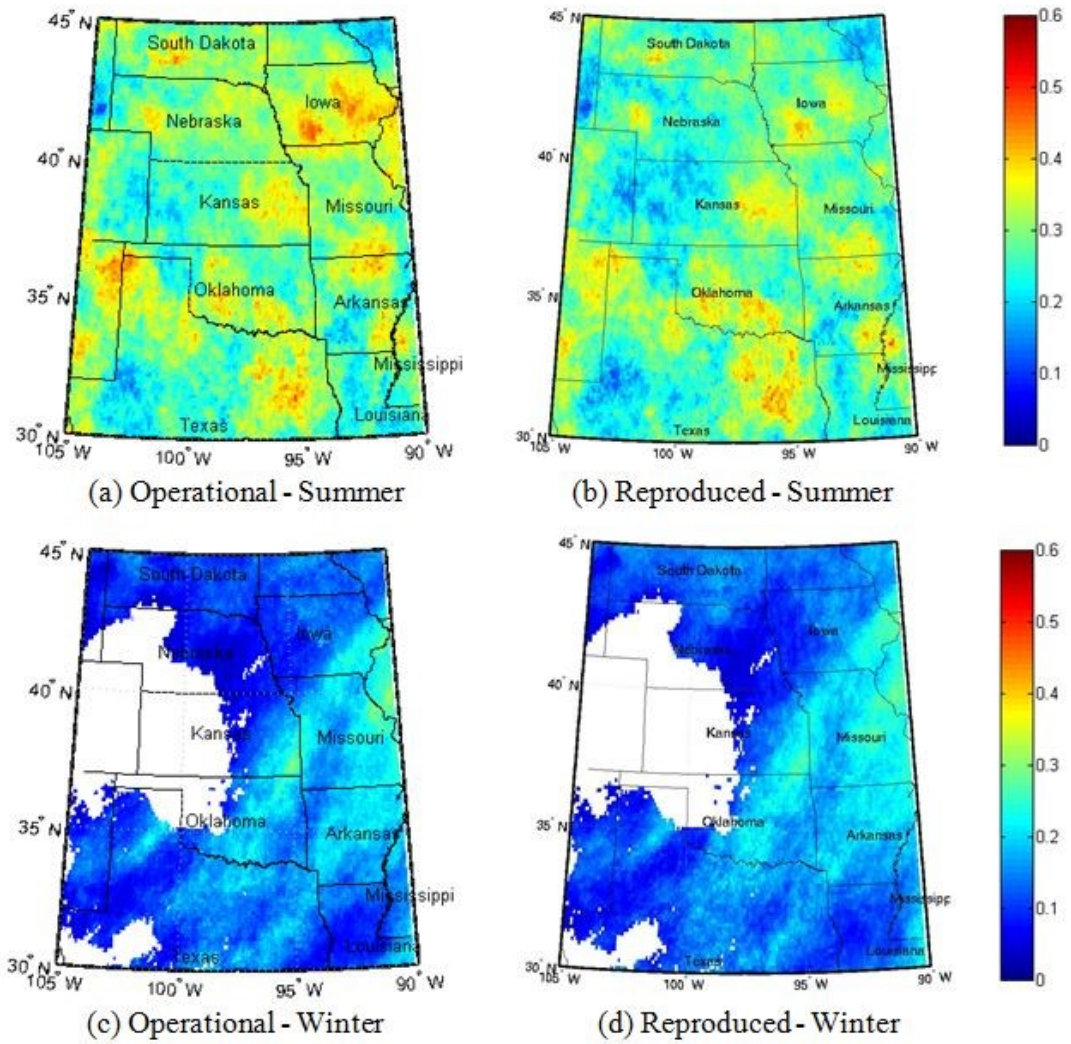

Figure A-1 CSI of the operational PERSIANN-CCS [(a), (c)] and the reproduced PERSIANN-CCS [(b), (d)] over the central United States (30–45°N, 90–105°W): (a), (b) summer (June–August 2013); (c), (d) winter (December 2013–February 2014)

Figure A-2 shows average MSE for the operational and the reproduced PERSIANN-CCS in the validation periods for both winter and summer seasons. As Figure A-2(a, b) shown, the overall geophysical pattern of the performances are quite similar. The reproduced version

performs slightly worse in the southwest corner. On the other hand, for the winter season, the overall performance for the reproduced PERSIANN-CCS is better than the operational version. The spatial consistency validates the stability of the PERSIANN-CCS algorithm.

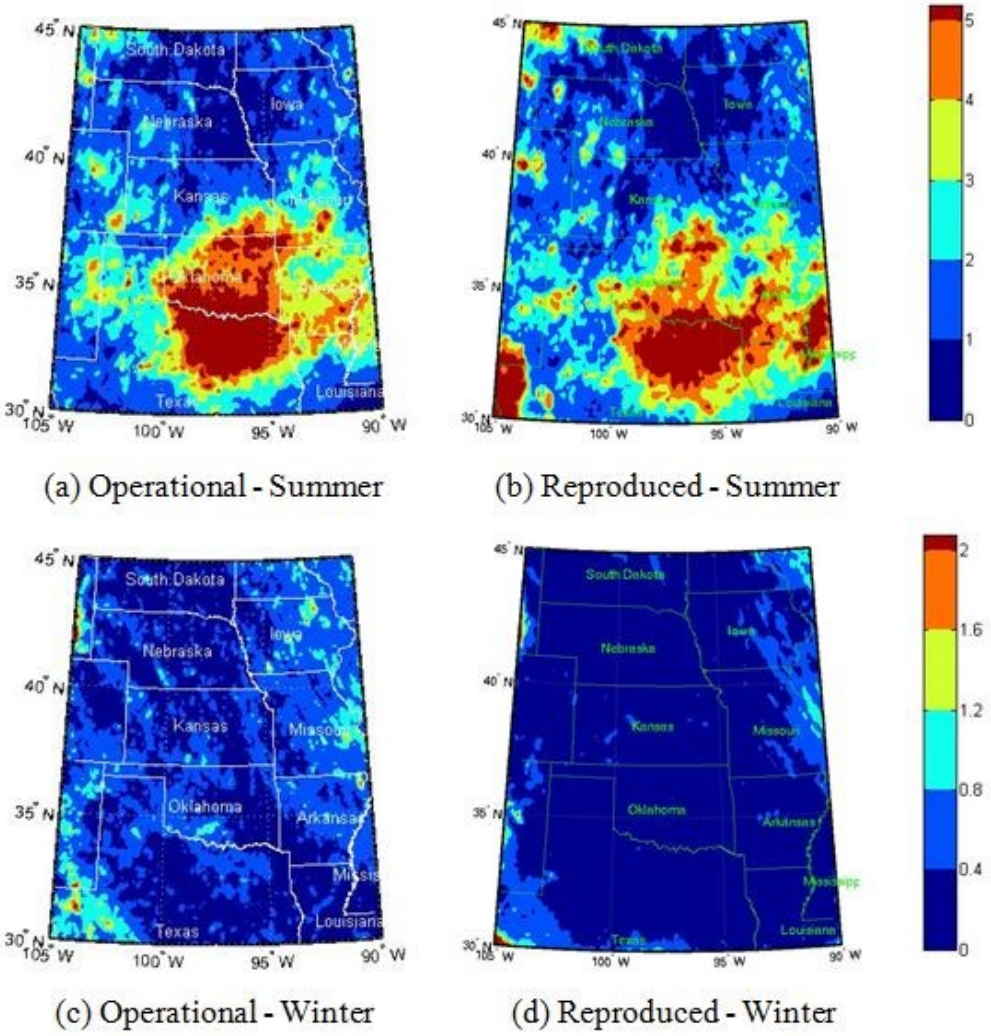

Figure A-2 Average MSE ([mm/hr]^2) of the operational PERSIANN-CCS [(a), (c)] and the reproduced PERSIANN-CCS [(b), (d)] over the central United States (30–45°N, 90–105°W): (a), (b) summer (June–August 2013); (c), (d) winter (December 2013–February 2014)

In this section, I provided the performance of the reproduced PERSIANN-CCS, compared to the operational version. The performances of the two versions are comparable to each other, which demonstrate the algorithm's stability to difference calibration datasets. In the rest of the dissertation, I use the data from the operational PERSIANN-CCS as our baseline performance, due to its accessibility.