Western&Graduate&PostdoctoralStudies

8-18-2020 10:00 AM

Evaluation and bias correction of remotely sensed precipitation products across Canada

Saber Moazamigoodarzi, *The University of Western Ontario*

Supervisor: Najafi, Mohammad Reza, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Civil and Environmental Engineering
© Saber Moazamigoodarzi 2020

Abstract

Reliable estimation of precipitation, as the most important variable in hydrological modelling, is crucial for water resources management. Rain gauges that provide precipitation measurements at point scale have inherent limitations and difficulties in remote regions and complex terrains due to accessibility, gauge undercatch, among others. Alternatively, satellite and radar precipitation data can estimate precipitation at high spatial and temporal resolution by utilizing several types of space and ground-borne sensors. However, due to the indirect estimates of precipitation by remotely sensed products, their measurements are subject to systematic biases and are required to be evaluated and bias-adjusted before using at a specific area.

This study investigates the performance of multiple high-resolution remotely sensed precipitation estimates at hourly and daily time scales over Canada for 2014-2018. Four products of the recently released Integrated Multi-satEllite Retrievals for Global precipitation measurement (IMERG-V06) and the Multi-Radar Multi-Sensor (MRMS) Precipitation Rate data for different seasons are analyzed. Evaluations are based on a suite of metrics to assess different characteristics of precipitation estimates using quality-controlled hourly gauge data considered as the truth. The results suggest that Calibrated precipitation (PrCal) outperforms the other IMERG products and estimates the amount of precipitation relatively well particularly over the Prairies and during fall and summer. Over the western and eastern coastal regions, IMERG tends to overestimate precipitation intensities by around 25%. The discrepancy between satellite and ground-based data is higher for more intense precipitation events. Further analyses indicate that while MRMS tends to overestimate the amount of precipitation, it outperforms the IMERG products based on several metrics, especially in detecting the occurrence of precipitation over the eastern coastal regions. Overall, the study of IMERG V06 and MRMS precipitation estimates at a relatively high temporal resolution indicates that both products have the potential to complement ground-based observations over Canada.

Further, a regression quantile mapping method is developed to adjust the systematic spatial and temporal biases of IMERG PrCal across Canada. For this purpose, several climatic and topographic explanatory variables are resampled and applied in the regression-based model to extend satellite bias correction over the ungauged pixels. The proposed method shows promising results by reducing RBias (by ~32%) and increasing correlation values (by ~ 0.31). The bias-corrected precipitation product (for 2014-2018) can be applied by researchers and

various stakeholders, across Canada, for the analysis of extreme precipitation events, water resources management, design of infrastructure, among others.

Finally, the application of daily IMERG data in streamflow simulation is demonstrated by using the original data to drive the calibrated Raven rainfall-runoff model over the Bathewana watershed in southern Ontario for 2001-2015. By comparing with the observed flow, the obtained results indicate that IMERG tends to underestimate the streamflow, however, it is able to preserve its temporal variation reasonably well. Overall, results suggest that further improvements of IMERG data should be considered by its algorithm developers to enhance the quality of this product in cold weather conditions.

Keywords

Lay Summary

Precipitation is the most important component in hydrological applications. Therefore, reliable measurement of precipitation is crucial for having more accurate monitoring of water resources supplies and forecasting extreme weather events such as floods. However, due to the high spatiotemporal variability of precipitation, its accurate estimation is a challenging task especially over complex terrain where the ground-based rain gauges are either sparse or nonexistent. Recently for dealing with the limitations of ground-based stations availability, remotely sensed algorithms that use satellite and radar data have been developed to estimate precipitation. Nevertheless, the remote sensing-based data need to be evaluated before using due to the indirect nature of their estimates. The most well-known and recently released satellite-based precipitation products named Integrated Multi-satEllite Retrievals for Global precipitation measurement (IMERG-V06) and the Multi-Radar Multi-Sensor (MRMS) data are evaluated in this study to investigate the performance of such a high spatiotemporal resolution precipitation data over Canada. Although the findings of this study indicate the promising value of satellite and radar precipitation over most parts of the country, it still shows bias in some regions. Therefore, a Regression-based Quantile Mapping (RQM) method is developed to correct the biases associated with the IMERG data spatially and temporally over the entire country. The proposed framework can significantly improve the IMERG data in different regions during the study period (2014-2018) and provide a high quality of precipitation data over Canada. In addition to statistical evaluations and bias correction, the ability of IMERG precipitation in daily streamflow simulation is assessed by forcing it in a calibrated hydrological model. For this purpose, the Raven model calibrated by using the ground-based rainfall data over the Batchawana as a small watershed ($1280 \ \text{km}^2$) located in the southern part of Ontario, Canada is selected. Due to the error of input IMERG precipitation as well as the uncertainty of the calibrated Raven model, the output simulated streamflow is not promising. However, simulated streamflow by forcing IMERG data can capture the trend of observed discharge reasonably. Overall this study provides insights into remotely-sensed data over Canada and helps to have a high spatiotemporal resolution of precipitations.

Acknowledgments

Table of Contents

List of Tables

List of Figures

Chapter 1

1.Thesis Overview

1.1. Background

Precipitation is the key input variable to hydrological models and plays an important role in water resource planning including flood and drought analysis, monitoring, and forecasting. However, its accurate estimation is challenging particularly in areas with sparse observations and regions with complex terrains. Although direct precipitation measurements using rain gauges are considered as the most accurate observations (Petersen et al., 2005, Singh and Najafi, 2020), there are limitations associated with the at-point representation of an entire domain, and inability to capture precipitation variability at high resolution (Villarini et al. 2008). Therefore, there is considerable interest to use indirect Remotely Sensed Estimate (RSE) data such as radar and satellite products as they provide fine-scale representations of the amount, frequency, and distribution of precipitation (Sungmin et al., 2017; Sun et al. 2018; Wen et al. 2018).

Flooding is the most common natural disaster in Canada and among the costliest according to Public Safety Canada. Many historical flood events in major river systems and populated areas across Canada are associated with heavy rain and subsequent excessive runoff (Lemmen et al. 2016). Therefore, reliable precipitation data with high spatial and temporal resolution are essential for flood risk mitigation and water resources management particularly in mountainous regions with limited accessibility. Nonetheless, Canada like many other countries suffers from a lack of dense network of gauges especially in remote areas such as the Arctic and high elevation zones (Mekis et al., 2018). The existing point coverage of gauge measurements may not represent the highly variable spatial distribution of precipitation properly (Martinaitis et al., 2015). The RSE products can address these limitations by providing high spatial and temporal coverage of precipitation that can be used to detect storm events, assess flooding, and develop mitigation measures.

It is worth mentioning that, RSEs due to the indirect nature of precipitation estimates are subject to the systematic biases which can propagate into hydrological models and lead to uncertainty in streamflow prediction. Therefore, these datasets are required to be corrected before using at each specific region.

Over Canada, only a few studies have analyzed RS data and analysis of most recent products at high spatial and temporal resolution is lacking. In addition, although several methods have been developed to adjust satellite biases, the challenges of improving this data spatially and temporally still remain.

1.2. Research Objectives

The overall objective of this research is to develop a reliable precipitation product based on remotely sensed precipitation estimates at a high spatiotemporal resolution over Canada. Therefore, the first objective is to evaluate the biases associated with two widely used and most well-known RSE datasets including Integrated Multi-satEllite Retrievals for Global precipitation measurement (IMERG-V06) and the Multi-Radar Multi-Sensor (MRMS) products across Canada using ground station rain gauges. The second objective is to correct the bias of IMERG data as a product covering the entire country of Canada and useful in areas with sparse ground observations. For this purpose, in addition to rain gauge data, several other reanalysis covariates are extracted and utilized. Finally, the third objective considers the application of satellite rainfall data in the hydrology field by using a calibrated rainfall-runoff hydrological model over a small watershed located in southern Ontario. In this study, evaluations of the RSE products are performed using available rain gauge records across Canada, which has a diverse hydroclimate due to its extensive geographical features, latitudinal extent, and variations in topography

1.3. Research Questions

The following are the research questions addressed in the study.
   (1) What are the characteristics of RSE biases in comparison with ground station data?

(2) To what extent are satellite and radar precipitation data reliable over different parts of Canada?

(3) How to correct the bias of satellite gridded data by using a sparse network of ground truth rain gauges?

(4) How to correct the bias of satellite gridded data over the ungauged area?

(5) To what extent can satellite precipitation data predict streamflow in small basins?

1.4. Summary of Chapters

This thesis consists of six chapters. After an overview in chapter one, the background literature is presented in chapter 2. It provides a summary of previous studies regarding the evaluation as well as the application of different satellite and radar precipitation products over several regions around the world. In this chapter in addition to introducing different satellite and radar precipitation products, the robustness of these data in different climatic conditions is discussed and their applications in hydrological modeling are summarized. Having more information concerning RSE precipitation products and their potential performances provides hydrological and meteorological modelers with higher quality data which can be useful in more accurate predictions of future rainfall, streamflow, and floods.

Chapter 3 discusses the evaluation of four products of the most recently released satellite IMERG version 6 as well as a combined radar data named MRMS over their entire coverage in Canada. Here, the systematic bias of both datasets is extracted at hourly time scale for different climatic regions. Several metrics are calculated and the distribution of fundamental bias indices are provided spatially and temporally over the study area. The performance of different products and their variations due to the inherent uncertainty in different climatic and topographic conditions are quantified and discussed in detail. It is worthwhile to mention that, for the first time MRMS as the most high quality integrated radar precipitation data is assessed over its domain covering the southern part of Canada. This gridded precipitation QPEs data help to have a better understanding of extreme rainfall events as gauges and satellite-borne systems, providing a gap in knowledge of extreme precipitation (Lengfeld et al. 2020).

Chapter 4 as another main chapter of the thesis, proposes a framework for spatiotemporal bias correction of satellite IMERG daily data over the entire country of Canada considering areas where no ground observation exists. The method is based on both quantile mapping and regression between precipitation as the dependent variable and several independent predictors at each pixel resolution. By extracting, preprocessing, and providing required data including reanalysis climatic covariates, ground observations and satellite estimates during five years from 2014 to 2018, the methodology is developed and applied first over gauged pixels. After, the parameters computed as the output of the quantile regression model at each gauged pixel are interpolated and distributed over all other ungauged pixels and imply into the covariates to extract the bias-corrected satellite data at those ungauged pixels. A combination of various statistical techniques is utilized in this chapter to make a more reliable method of bias correction.

As the last main section, chapter 5 provides a hydrological evaluation of IMERG data in a small basin in southern Ontario. By using IMERG precipitation data to drive the Raven hydrological model, the performance of this satellite product in hydrological simulation is assessed. Raven is a robust and flexible hydrological modeling framework that can be as simple as a single watershed lumped model with only a handful of state variables to a full semi-distributed system model with physically-based infiltration, snowmelt, and routing. In this study, the calibrated semi-distributed Raven model using ground station rain gauge is used and just the input precipitation data is replaced by original satellite data to characterize the uncertainty associated with IMERG rainfall product in the hydrological application.

Finally, chapter 6 provides the concluding remarks and proposes future research in the field of remotely sensed precipitation data.

1.5. References

Lemmen, D.S., Warren, F.J., James, T.S., Mercer Clarke, C.S.L., 2016. Canada's Marine Coasts in a Changing Climate; Government of Canada, Ottawa, ON, 274p.

Lengfeld K., Kirstetter P.E., Fowler H.J., Yu J., Becker A., Flamig Z., Gourley J., 2020. Use of radar data for characterizing extreme precipitation at fine scales and short durations. Environmental Research Letters, in press https://doi.org/10.1088/1748-9326/ab98b4.

Martinaitis, S.M., Cocks, S.B., Qi, Y., Kaney, B.T., 2015. Understanding Winter Precipitation Impacts on Automated Gauge Observations within a Real-Time System, Journal of OF Hydrometeorology, DOI: 10.1175/JHM-D-15-0020.1

Mekis, E., Donaldson, N., Reid, J., Zucconi, A., Hoover, J., Li, Q., Nitu R., Melo S., 2018. An Overview of Surface-Based Precipitation Observations at Environment and Climate Change Canada, Atmosphere-Ocean, 56:2, 71-95, DOI:10.1080/07055900.2018.1433627.

Petersen, W.A., Christian, H.J., Rutledge, S.A., 2005. TRMM observations of the global relationship between ice water content and lightning. Geophys. Res. Lett. 32, 2471–2494.

Singh H, Najafi M.R., 2020. Evaluation of gridded climate datasets over Canada using univariate and bivariate approaches: Implications for hydrological modelling. Journal of Hydrology, 1;584:124673.

Sungmin O., Foelsche U., Kirchengast G., Fuchsberger J., Tan J., Petersen W.A., 2018. Evaluation of GPM IMERG Early, Late, and Final rainfall estimates using WegenerNet gauge data in southeastern Austria, Hydrol. Earth Syst. Sci., 21, 6559–6572, 2017 https://doi.org/10.5194/hess-21-6559-2017

Sun Q., Miao C., Duan Q., Ashouri H., Sorooshian S., Hsu K.L., 2018. A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons. Reviews of Geophysics, 56, 79–107. https://doi.org/10.1002/ 2017RG000574

Villarini, G., Mandapaka, P.V., Krajewski, W.F., Moore, R.J., 2008. Rainfall and sampling uncertainties: A rain gauge perspective. Journal of Geophysical Research. VOL. 113, D11102. doi:10.1029/2007JD009214.

Wen Y., Behrangi A., Chen H., Lambrigtsen B., 2018. How well were the early 2017 California Atmospheric River precipitation events captured by satellite products and ground-based radars? Quarterly Journal of the Royal Meteorological Society, 2018;144 (Suppl. 1):344–359. https://doi.org/10.1002/qj.3253

Chapter 2

2. Background Literature

2.1. Satellite Precipitation Products

Quantitative precipitation information is primarily derived from rain gauges, weather radar, and satellite measurements. The rain gauge directly measures the precipitation and is still deemed as the most accurate means for observing rainfall, but it has accuracy limitations due to measuring rainfall at a point scale (Zhang et al., 2018). Another challenge is related to distinguishing between different precipitation phases (liquid and solid) particularly by conventional ground stations. Therefore, considering the limitations associated with rain gauges, remotely sensed precipitation products that provide higher spatial resolution can be utilized. In general, there are two types of remotely sensed data for estimating precipitation; weather radar provides precipitation occurrence and amount at the relatively high spatiotemporal resolution, however, global weather radar coverage is poor with sparse distribution over most parts of the world, especially in developing countries and ocean areas (Zhang et al., 2019). With advanced infrared (IR) and microwave (MW) instruments, satellite observations make up for these deficiencies by providing coverage that is more spatially homogeneous and temporally complete for vast areas of the globe (Sun et al., 2018). Currently, satellites can provide precipitation estimates globally by using three different categories of sensors including visible/IR (VIS/IR) sensors on geostationary (GEO) and low Earth orbit (LEO) satellites, passive MW (PMW) and active MW sensors on LEO satellites (Sun et al., 2018). A GEO satellite is an earth-orbiting satellite, placed at an altitude of approximately 35,800 kilometers directly over the equator that revolves in the same direction the earth rotates. On the other hand, an LEO satellite is an Earth-centered orbit with an altitude of 2,000 km or less (approximately one-third of the radius of Earth) with at least 11.25 periods per day (an orbital period of 128 minutes or less) (Sampaio et al., 2014). It is also noted that IR sensors can estimate precipitation based on the cloud top temperature and its link with the probability and intensity of rainfall at the ground, however, PMW radiometer provides a more direct measurement of precipitation as the PMW radiation can sense through clouds and is sensitive to precipitation-sized droplets (Sun et al., 2018). As IR sensors onboard GEO satellites provide higher temporal resolutions (30 min), and PMW sensors onboard LEO satellites provide more

accurate data, usually combining both sensors would increase the spatiotemporal accuracy and coverage of precipitation data (Sorooshian et al. 2002).

National Aeronautics and Space Administration (NASA) and the Japan Aerospace and Exploration Agency (JAXA), the Tropical Rainfall Measurement Mission (TRMM) and Global Precipitation Measurement (GPM) mission have built unprecedented international cooperation in space asset sharing and scientific collaboration to advance precipitation estimation from space for research and applications. A consortium of international partners provides consistent precipitation estimates from a constellation of satellites combined active/passive sensor measurements. The GPM "Core" satellite is carried by NASA and JAXA, however, it is an international satellite mission, specifically designed to unify and advance precipitation measurements from research and operational microwave sensors for delivering next-generation global precipitation data products.

Through bilateral agreements with either NASA or JAXA, GPM achieves global coverage with a high sampling frequency by relying on both existing satellite programs and new mission opportunities from its partners. Each constellation member may have its unique scientific or operational objectives but contributes microwave measurements to GPM for the generation and dissemination of uniform global precipitation products for worldwide user communities.

In addition to the DPR and GMI (GPM Microwave Imager (GMI) and Dual-frequency Precipitation Radar (DPR) instruments) on the GPM Core Observatory, the GPM constellation satellites have the following groups of conical-scanning microwave imagers:

Special Sensor Microwave Imager/Sounder (SSMIS) instruments on U.S. Defense Meteorological Satellite Program (DMSP) satellites,

The Advanced Microwave Scanning Radiometer-2 (AMSR-2) on JAXA's Global Change Observation Mission - Water 1 (GCOM-W1) satellite,

The Multi-Frequency Microwave Scanning Radiometer (MADRAS) and the multi-channel microwave humidity sounder (SAPHIR) on the Megha-Tropiques satellite provided by the Centre National D'Etudies Spatiales (CNES) of France and the Indian Space Research Organisation (ISRO),

The Microwave Humidity Sounder (MHS) instrument on the National Oceanic and Atmospheric Administration (NOAA)-19 satellite,

MHS instruments on the MetOp series of satellites launched by the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT),

The Advanced Technology Microwave Sounder (ATMS) instruments on the National Polar-orbiting Operational Environmental Satellite System (NPOESS) Preparatory Project (NPP),

ATMS instruments on the upcoming NOAA-NASA Joint Polar Satellite System (JPSS) satellites,

A microwave imager planned for the Defense Weather Satellite System (DWSS), (http://pmm.nasa.gov/GPM/constellation-partners)

Some of the widely used satellite precipitation products (SPPs) are now operationally available including Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN; Sorooshian et al. 2000; AghaKouchak et al. 2012; Hossain and Huffman 2008), the National Oceanic and Atmospheric Administration (NOAA)'s Climate Prediction Center (CPC) morphing technique (CMORPH; Joyce et al. 2004; Stampoulis and Anagnostou 2012; Gumindoga et al., 2019), the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA; Huffman et al. 2007; Villarini and Krajewski 2007; Tian et al. 2007; Habib et al. 2009; Yong et al. 2015; Mei et al. 2016; Moazami et al. 2014, 2016), and the Global Satellite Mapping of Precipitation (GSMaP; Kubota et al. 2007 and 2009). These satellite products have some limitations associated with the spatial and temporal resolution of precipitation due to the number of IR and MW based sensors utilized. To provide more accurate precipitation estimates at fine spatiotemporal scales, NASA and JAXA launched the GPM mission in February 2014, called the Integrated Multi-satellitE Retrievals for GPM (IMERG) with a temporal resolution of 30 min and spatial resolution of 0.1° as a successor to TRMM (Hou et al. 2014; Liu, Z 2016; Tang et al. 2016). The DPR, the first of its kind, was incorporated in the GPM core observatory to improve the reliability of IMERG compared to other SPPs (Anjum et al., 2019).

The record for IMERG V05 (previous version) begins in March 2014 with coverage between ±60° latitudes, while IMERG V06 (last version) extends this record back to June 2000 (eventually to January 1998) with global coverage (90°N/S). IMERG has three runs: Early, Late, and Final to accommodate different user requirements for latency and accuracy. This study uses the gauge-adjusted estimates from the Final runs of IMERG V06B, the latest version of V06. A quasi-Lagrangian interpolation (known as "morphing") is applied to the 0.1° gridded

PMW estimates to fill in gaps in the field using motion vectors computed from total precipitable water vapor from numerical models in V06. The morphed precipitation is further supplemented, via a Kalman filter approach following Joyce and Xie (2011), with microwave-calibrated IR precipitation estimates using the PERSIANN-Cloud Cluster System algorithm (PERSIANN-CCS; Hong et al., 2004; Nguyen et al., 2018). IMERG masks PMW and morphed estimates over frozen surfaces, resulting in the use of IR precipitation within 60°N/S and missing values at high latitudes. The merged satellite estimates are then calibrated by the monthly surface gauge analyses from the Global Precipitation Climatology Centre (GPCC; Schneider et al. 2014, 2015) following the approach employed by Huffman et al. (2007) for the TRMM-TMPA (Tan et al., 2019).

Several currently available satellite precipitation datasets are summarized in Table 1. In this table, Global Precipitation Climatology Project (GPCP) is a monthly precipitation analysis that merges gauge observations with LEO satellite MW data and GEO satellite IR data and is one of the most popular products used in climate studies (Adler et al., 2003).

Table 1. Summary of Major Satellite-Related Precipitation Products Currently Available (Sun et al., 2018)

Dataset	Spatial Res.	Temporal Res.	Coverage	Period	Data source	Reference
GPCP	2.5°	Monthly	Global	1979-Present	GPI, OPI, SSM/I scattering, SSM/I emission, TOVS	(Adler et al., 2003)
TRMM-3B43	0.25°	Monthly	50°S–50°N	1998-Present	TMI, TRMM Combined Instrument, SSM/I, SSMIS, AMSR-E, AMSU-B, MHS, and GEO IR	(Huffman et al., 2007)
TRMM 3B42	0.25°	3 h/Daily	50°S–50°N	1998-Present	TMI, TRMM Combined Instrument, SSM/I, SSMIS, AMSR-E, AMSU-B, MHS, and GEO IR	(Huffman et al., 2007)
GSMaP	0.1°	1 h/Daily	60°S–60°N	2002-2012	TMI, AMSR-E, AMSR-E, SSM/I, multifunctional transport satellites (MTSAT), Meteosat-7/8, GOES 11/12	(Ushio et al., 2009)
PERSIANN-CCS	0.04°	30 min/3, 6 h	60°S–60°N	2003-Present	Meteosat, GOES, GMS, SSM/I, polar/near polar precipitation radar, TMI, AMSR	(Sorooshian et al., 2000)
PERSIANN-CDR	0.25°	3, 6 h/Daily	60°S–60°N	1983-Present	GOES 8, GOES 10, GMS-5, Metsat-6, and Metsat-7, TRMM, NOAA 15, 16, 17, DMSP F13, F14, F15	(Ashouri et al., 2015)
CMORPH	0.25°/8 km	30 min/ 3 h/Daily	60°S–60°N	2002-Present	TMI, SSM/I, AMSR-E, AMSU-B, Meteosat, GOES, MTSAT	(Joyce et al., 2004)
GPM	0.1°	30 min/ 3 h/Daily	90°S–90°N	2000-Present	GMI, AMSR-2, SSMIS, Madaras, MHS, Advanced Technology Microwave Sounder	(Hou et al., 2008, 2014)

2.2. Radar Precipitation Data

Satellites use cloud information for estimating precipitation, however, radar can measure the precipitation by sending radio waves and receiving their reflectance back to the transmitting point. On the other hand, rain gauges record the precipitation reaching the ground surface.

Similar to SPPs, radar precipitation data provide real-time estimates of rain and snow rates at relatively fine spatial and temporal scales, however, indirect measurements based on radar reflectivity can cause errors (Rodriguez et al. 2019). Radar precipitation estimates are influenced by ground clutter (e.g. dust, bugs, birds, and particulates) and other non-meteorological echoes, beam blockage, and bright banding in the melting layer (occurring due to the higher reflectivities associated with snow that is melting as it is falling aloft) (Martinaitis et al. 2017). The Canadian Weather Radar Network (CWRN) consists of 31 weather radars, 29 of which are owned and operated by ECCC and two by Department of National Defence (DND). Coverage is nominal to a range of 256 km for non-Doppler data and 120 km for Doppler data, but some areas within that nominal range, where the radar beam is blocked by topography, buildings, etc., may be unavailable (Mekis et al. 2018).

Multi-radar integration can mitigate such deficiencies in the single-radar framework. The integration of several radars with a set of sensors provides more accurate diagnoses of atmospheric physical processes than using radar data alone (Zhang et al., 2016). For this purpose, the National Centers for Environmental Prediction (NCEP) implemented the Multi-Radar Multi-Sensor (MRMS) system, which integrates multiple overlapping radars with other in situ and remote sensing (satellite) observations and Numerical Weather Prediction (NWP) model output. MRMS currently uses 176 operational radars across the conterminous United States (CONUS) (146 S-band dual-polarization Weather Surveillance Radar-1988 Doppler (WSR-88D) radars) and southern Canada (30 C-band single-polarization weather radars) at very high spatial (1 km) and temporal (2 min) resolution (Zhang et al. 2016). MRMS provides four types of Quantitative Precipitation Estimation (QPE) products: 1) radar-based QPE (radar-only) with a vertical profile of reflectivity (VPR) correction, 2) gauge-based QPE (gauge-only), 3) local gauge and VPR bias-corrected QPE, and 4) gauge-and-precipitation-climatology-

merged QPE (Mountain Mapper QPE designed for the mountainous areas in the western US and Canada but is generated for the entire MRMS domain) (Zhang et al. 2016). MRMS system, as a high-quality QPE product, has been recently used over several parts of CONUS and shows the mitigation of radar beam overshoot primarily via the use of multiple radar inputs for a given grid point (Cocks et al. 2017). The accuracy of the MRMS products over CONUS has led researchers to use them as reference data for evaluating other QPEs like IMERG and TRMM (Gebregiorgis et al. 2018). Note that, although MRMS provides the spatial coverage over southern (south of 55°N) part of Canada, to date there has not been reported any evaluation of it across the country.

2.3. Statistical Evaluation of QPEs

For having reliable hydrological simulation as well as better estimation of water resources conditions, accurate estimates of precipitation as the key variable in water application is required. In the case of a sparse and unreliable network of ground rain gauges, particularly over complex terrain and remote areas, satellite and radar QPEs product are widely used in hydrological applications (Jiang et al. 2019). However, QPEs, due to their indirect nature of precipitation estimations are subject to error and uncertainty and need to be evaluated before utilizing in water resources models (Moazami et al. 2014). Several studies regarding the statistical and hydrological evaluation of IMERG and MRMS as the objective products of this research have been conducted in recent years.

Huang et al., 2018 evaluated the performance of IMERG in depicting the spatial-temporal characteristics of precipitation variations over Taiwan at multiple (including annual, seasonal, intraseasonal, diurnal, and semidiurnal) timescales. The results obtained from this study showed that, quantitatively, IMERG underestimated the magnitude of precipitation over most of Taiwan for all the examined timescales; spatially, the bias in variability was larger over the mountainous areas than over the plain areas; temporally, the bias in variability was larger in the warm seasons than in the cold seasons. Despite these differences, IMERG was able to keep the variation of precipitation, especially the peak values qualitatively. Sungmin et al. 2017, compared IMERG version 3 Early, Late, and Final (IMERG-E, IMERG-L, and IMERG-F) half-hourly rainfall estimates with gauge-based gridded rainfall data from the WegenerNet Feldbach region (WEGN) high-density climate station network in southeastern Austria. Results

showed that IMERG-F rainfall estimates are in the best overall agreement with the WEGN data, followed by IMERG-L and IMERG-E estimates, particularly for the hot season. Both studies indicated the better performance of IMERG during warm seasons. On the other hand,

Gebregiorgis et al., 2018 during the assessment over CONUS, illustrated that IMERG is more consistent with the reference data for all seasons, except for a slight underestimation over Florida and the southeast coastal region of the CONUS during fall. They also concluded that the inclusion of IR data into IMERG algorithms results in overestimation in all seasons except winter. Although several studies have indicated that IMERG outperforms other SREs globally (Gebregiorgis et al. 2018, Zhang et al. 2019) due to the marked improvement of IR precipitation retrieval by implementing CMORPH-KF (Kalman-filter) and PERSIANN-CCS, evaluation of IMERG estimates over several regions around the globe has reported different level of uncertainties associated with this product that may limit its direct use in practical applications (Sungmin et al. 2017; Tang et al. 2017; Wang et al. 2018; Asong et al. 2017; Tan et al. 2019). Thus, the accuracy of IMERG precipitation products needs to be assessed against in situ observations.

Tan et al., 2016 evaluated IMERG against a dense network of gauges in the mid-Atlantic region of the United States. In their approach, ancillary variables leveraged in IMERG to attribute the errors to the individual instruments or techniques within the algorithm. They concluded that as a whole, IMERG exhibits some misses and false alarms for rain detection, while its rain-rate estimates tend to overestimate drizzle and underestimate heavy rain with considerable random error. They also mentioned that the most reliable IMERG estimates come from passive microwave satellites, but infrared estimations perform poorly. On the other hand, Asong et al., 2017 evaluated GPM version 03 IMERG Final Run product against ground-based reference measurements (at the 6-hourly, daily, and monthly time scales) over different terrestrial ecozones of southern Canada within a 23-month period from 12 March 2014 to 31 January 2016. They concluded that IMERG and ground-based observations show similar regional variations of mean daily precipitation, while IMERG tends to overestimate median to heavy precipitation amounts over the Pacific Maritime ecozone. These two performed assessments of IMERG in two different regions indicate contrary findings regarding over/underestimates for heavy rainfall which makes the necessity of evaluating the performance of this data in each specific area before using in hydrological applications. By reviewing the kinds of literature

concerning the evaluation of satellite precipitation data over different regions around the globe, it can be realized that most studies have limitations of using the number of statistical metrics, time scale (mostly considered daily or monthly scales), diversity of study region. For filling these gaps and having a comprehensive evaluation of the most recent high-resolution satellite and radar precipitation products, this study aims to assess hourly IMERG V06 and MRMS at hourly time scale over diverse climatic and topographic zones (cold, mountainous, coastal, and plain regions) by using a complete list of statistical and categorical metrics.

2.4. Bias Correction of Satellite Precipitation Products

Due to inherent error and uncertainty associated with QPEs, this study aims to develop a framework for bias correction of satellite IMERG precipitation products. Therefore, in this section some recently published studies regarding SPPs bias correction are summarized. By using ground-based precipitation measurement obtained from CPC daily gridded as a true data over the southwestern United States, Boushaki et al., 2009 used a merging methodology to adjust the bias of PERSIANN-CCS in hourly temporal and 0.04-degree spatial resolution. They first adjusted the bias at daily temporal and 0.25 degree spatial according to the original CPC data resolution, then downscaled 0.25 degree to 0.04 degree and redistributed the daily bias proportionally to the hourly rainfall estimates. The results indicated that the method can improve satellite estimates on a daily scale effectively, however, on the sub-daily scale a limited improvement was noticed. This simple scaling method has limitations as it needs simultaneous reference data and cannot take advantage of historical data (Yang et al. 2016). As a powerful method in reducing the systematic bias of regional climate model precipitation estimates, Quantile Mapping (QM) has shown the best skill in several studies (Cannon et al. 2015, Ajaaj et al. 2016, Yang et al. 2016, Ringard et al. 2017). Ringard et al. 2017, used the QM to correct the daily TRMM-TMPA-3B42V7 data and found that this technique can reduce the bias up to 70% for rainfall intensities less than 25 mm/d, but it performs weakly to correct the higher intensities. Yang et al. 2016, proposed a coupled nonparametric QM and Gaussian weighting (GW) interpolation scheme to adjust biases of PERSIANN-CCS over Chile.

The bias correction approaches developed in these studies can be categorized in two main groups. The first group uses the rain gauge data directly to correct the bias of SPPs in a specific time scale (daily, monthly, etc.) to remove the mean bias value during the considered time,

which does not capture the inter-time variability or may not remove the higher-order moments. The second group uses QM for correcting the biases based on the differences between quantiles of SPPs and ground truth data. This technique can effectively capture the evolution of the mean and the variability of precipitation while matching all statistical moments (Ajaaj et al., 2016); however, there is no adjustment to the temporal structure of precipitation it is unable to capture wet and dry spell lengths. To address these limitations, we correct the associated biases in the IMERG PrCal product spatially and temporally through a novel statistical approach. The detailed explanation regarding the method developed in this study are provided in Chapter 4.

2.5. Hydrological Evaluation of QPEs

The high spatial and temporal resolution of IMERG satellite precipitation has motivated hydrologists to apply this data in hydrological models. Indeed, droughts and floods can be monitored by high-resolution satellite-based products (AghaKouchak et al., 2015; AghaKouchak & Nakhjiri, 2012; Wu et al., 2014; Yilmaz et al., 2010). However, as mentioned before, SPPs contain uncertainties in retrieving precipitation characteristics and so the reliability of hydrologic predictions based on satellite-derived precipitation data need to be evaluated.

Several studies have been conducted by using a variety of lumped, semi-distributed and distributed hydrological models over different basins. Yuan et al., 2018 evaluated IMERG Final Run version 05 precipitation and TRMM-TMPA-3b42v7 products in daily and 3-hourly streamflow simulations by utilizing the grid-based Xinanjiang (GXAJ) hydrological model as a lumped, conceptual hydrological model calibrated with the gauge-based precipitation over Yellow River source region (YRSR), a mountainous Alpine region in northwestern China. IMERG with the Nash Sutcliffe Efficiency (NSE) coefficient of 0.810 demonstrated a good performance compared with the gauge-based simulation with NSE of 0.807, while for 3B42V7 data the NSE is 0.792. The disadvantage of this study was using a lumped model where the parameters were estimates to be spatially uniform, which may not sufficiently represent the hydrological features in the study area with complex climate and terrain.

In the hourly scale of hydrological assessment of IMERG, Li et al. 2017 applied the Coupled Routing and Excess Storage (CREST) distributed hydrological model over the Ganjiang River basin as the seventh-largest branch of the Yangtze River in the Jiangxi province of China. This study concluded that the hourly IMERG product can be used to simulate streamflow well based on the parameters calibrated by gauge (NSE= 0.7) and radar (NSE = 0.72) over Ganjiang River basin, although the parameters calibrated by IMERG is unusable (NSE = -2). The main advantage of this model was enforcing it with high spatiotemporal quality of radar QPE data that led to more reliable of IMERG precipitation products; however, the CREST model is complex and needs several physical and theoretical parameters for calibration (for more detail about this model see Wang et al., 2011).

Falck et al., 2015 investigated the applicability of error correction to satellite-based precipitation products in streamflow simulations over the 19 sub-basins of the Tocantins–Araguaia basin in the center-north region of Brazil. Four satellite products including TRMM-TMPA-3B42RT (Real-Time version of TMPA), CMORPH, GSMaP, and NOAA Hydroestimator (HYDRO-E) were evaluated. In order to analyze the uncertainty of simulated streamflow, they used Ensemble streamflow simulations of a distributed hydrological model developed by the Brazilian National Institute (MHD-INPE) (a grid-based model, Rodriguez and Tomasella, 2015; Mohor et al., 2015) by enforcing satellite rainfall products corrected using a two-dimensional stochastic satellite precipitation data (SREM2D, Hossain and Anagnostou, 2006). The findings of this study showed that SREM2D is able to correct for errors in the satellite precipitation data by pushing the modeled streamflow ensemble closer to the reference river discharge when compared to the simulations forced with uncorrected rainfall input. Ensemble streamflow error statistics (MAE and RMSE) depicted a decreasing trend as a function of the catchment area for all satellite products. As the main advantage of this study, Streamflow ensemble simulations reduced the error in CMORPH-, HYDRO-E-, and GSMaP-forced simulations when compared to the corresponding reference statistics for basins larger than 25,000 km2, showing that SREM2D was able to correct the error in the forcing rainfall in terms of both MAE and RMSE. Nevertheless, No remarkable difference among the different satellite products was observed at smaller basin scales, where SREM2D seems to consistently increase uncertainty in terms of MAE and RMSE to the simulated streamflow concerning the reference values.

Over the mid-size Illinois River basin, Behrangi et al., 2011 evaluated TMPA-RT, TMPA-V6, CMORPH, PERSIANN, and PERSIANN-adj) as forcing data in SACramento Soil Moisture Accounting (SACSMA) model for streamflow simulations at 6-h and monthly time scales. The calibration of the hydrological model is conducted for each satellite product separately. The SACSMA model was considered as a lumped model and the Shuffled Complex Evolution-Univ. of Arizona (SCE-UA; Duan et al., 1992) algorithm in conjunction with the Multi-step Automatic Calibration Scheme (MACS; Hogue et al., 2000) was used to calibrate the model parameters. The SCE-UA is a robust and efficient optimization algorithm for calibration of complex conceptual hydrologic models. The results indicated that satellite products are able to capture the streamflow pattern at both 6-h and monthly time series reasonably; however, they overestimated/underestimated both precipitation and simulated streamflow over warm/cold months significantly.

In order to well capturing of streamflow spatially over large basins, applying a distributed hydrological model with bias-adjusted SPPs input data, can be more useful. Sun et al., 2016, evaluated bias-corrected CMORPH (CMORPH-CRT), CMORPH satellite–gauge merged product (CMORPH-BLD), combined CMORPH RAW data with the daily precipitation from 2400 ground weather stations over Mainland China (CMORPH-CMA), and TMPA-3B42V7 into the distributed Variable Infiltration Capacity (VIC) model over the Huaihe River basin to simulate both long-term streamflow and extreme flood events. Among the three CMORPH-based QPEs, CMORPH-CMA matched the best with the observed, followed by CMORPH-BLD. They also found that over western China with sparse gauges, CMORPH-CMA is more reasonable than the gauge-based precipitation product. Therefore, CMORPH CMA could serve as an alternative high-quality QPE in China to evaluate the global satellite QPEs.

The abovementioned studies, all proved the key role of SPPs into well-calibrated distributed hydrological models in reliable streamflow prediction. Therefore, in this study, we evaluate the performance of IMERG PrCal product in streamflow simulation by enforcing it into the Raven hydrological model over a small basin.

2.5. References

Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., & Ping-Ping, X. (2003). The Version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). Journal of Hydrometeorology, 4, 1147–1167.

AghaKouchak, A., & Nakhjiri, N. 2012. A near real-time satellite-based global drought climate data record. Environmental Research Letters, 7(4), 044037. https://doi.org/10.1088/1748-9326/7/4/044037.

AghaKouchak, A., Farahmand, A., Melton, F. S., Teixeira, J., Anderson, M. C., Wardlow, B. D., & Hain, C. R. 2015. Remote sensing of drought: Progress, challenges and opportunities. Reviews of Geophysics, 53, 452–480. https://doi.org/10.1002/2014RG000456.

Behrangi, A., Khakbaz, B., Jaw, T.C., AghaKouchak, A., Hsu, K., Sorooshian, S., 2011. Hydrologic evaluation of satellite precipitation products over a mid-size basin. J. Hydrol. 397, 225–237.

Boushaki, F. I., K. Hsu, S. Sorooshian, G. Park, S. Mahani, and W. Shi (2009), Bias adjustment of satellite precipitation estimation using ground-based measurement: A case study evaluation over the southwestern United States, J. Hydrometeorol., 10(5), 1231–1242, doi:10.1175/2009jhm1099.1.

Duan, Q.Y., Sorooshian, S., Gupta, V., 1992. Effective and efficient global optimization for conceptual rainfall–runoff models. Water Resources Research 28, 1015–1031.

Falck A.S., Maggioni V., Tomasella J., Vila D.A., Diniz F.L.R., 2015. Propagation of satellite precipitation uncertainties through a distributed hydrologic model: A case study in the Tocantins−Araguaia basin in Brazil, Journal of Hydrology, http://dx.doi.org/10.1016/j.jhydrol.2015.05.042.

Gebregiorgis, A. S., Kirstetter, P.-E., Hong, Y. E., Gourley, J. J., Huffman, G. J., Petersen,W. A., Xue, X., Schwaller,M. R., 2018. To what extent is the day 1 GPM IMERG satellite precipitation estimate improved as compared to TRMM TMPART? Journal of Geophysical Research: Atmospheres, 123, 1694–1707. https:// doi.org/10.1002/2017JD027606.

Hogue, T.S., Sorooshian, S., Gupta, H., Holz, A., Braatz, D., 2000. A multistep automatic calibration scheme for river forecasting models. Journal of Hydrometeorology 1, 524–542.

Hossain, F., Anagnostou, E.N., 2006a. Assessment of a multi-dimensional satellite rainfall error model for ensemble generation of satellite rainfall data. Geosci. Remote Sens. Lett. (GRSL) 3 (3), 419–423.

Liguang Jiang, Peter Bauer-Gottwein, 2019, How do GPM IMERG precipitation estimates perform as hydrological model forcing? Evaluation for 300 catchments across Mainland China, Journal of Hydrology, https://doi.org/10.1016/j.jhydrol.2019.03.042.

Moazami S., Golian S., Kavianpour M.R., Hong Y., 2014, Uncertainty analysis of bias from satellite rainfall estimates using copula method, Atmospheric Research.

Mohor, G.S., Rodriguez, D.A., Tomasella, J., Siqueira Junior, J.L., 2015. Exploratory analyses for the assessment of climate change impacts on the energy production in an amazon run-of-river hydropower plant. J. Hydrol. Reg. Stud. http://dx.doi.org/10.1016/j.ejrh.2015.04.003.

Na Li, Guoqiang Tang, Ping Zhao, Yang Hong, Yabin Gou, Kai Yang, 2017, Statistical assessment and hydrological utility of the latestmulti-satellite precipitation analysis IMERG in Ganjiang River basin, Atmospheric Research, http://dx.doi.org/10.1016/.

Rodriguez, D.A., Tomasella, J., 2015. On the ability of large-scale hydrological models to simulate land use and land cover change impacts in Amazonian Basins. Hydrol. Sci. J. http://dx.doi.org/10.1080/02626667.2015.1051979.

Sampaio, Jarbas; Wnuk, Edwin; Vilhena de Moraes, Rodolpho; Fernandes, Sandro, 2014, "Resonant Orbital Dynamics in LEO Region: Space Debris in Focus". Mathematical Problems in Engineering, objects. doi:10.1155/2014/929810

Sorooshian, S., Gao, X., Hsu, K., Maddox, R. A., Hong, Y., Gupta, H. V., & Imam, B.,2002. Diurnal variability of tropical rainfall retrieved from combined GOES and TRMM satellite information. Journal of Climate, 15, 983–1001.

Sun R.,Yuan H., Liu X., Jiang X., 2016. Evaluation of the latest satellite–gauge precipitation products and their hydrologic applications over the Huaihe River basin, Journal of Hydrology, http://dx.doi.org/10.1016/j.jhydrol.2016.02.054.

Wan-Ru Huang, Ya-Hui Chang, Pin-Yi Liu, 2018, Assessment of IMERG precipitation over Taiwan at multiple timescales, Atmospheric Research, https://doi.org/10.1016/.

Wang, J., Hong, Y., Li, L., Gourley, J.J., Khan, S.I., Yilmaz, K.K., Adler, R.F., Policelli, F.S., Habib, S., Irwn, D., Limaye, A.S., Korme, T., Okello, L., 2011. The coupled routing and excess storage (CREST) distributed hydrological model. Hydrol. Sci. J. 56 (1), 84–98.

Wu, H., Adler, R. F., Tian, Y., Huffman, G. J., Li, H., & Wang, J. (2014). Real-time global flood estimation using satellite-based precipitation and a coupled land surface and routing model. Water Resources Research, 50(3), 2693–2717.

Yang, Z., Hsu, K., Sorooshian, S., Xu, X., Braithwaite, D., & Verbist, K. M. J. (2016). Bias adjustment of satellite-based precipitation estimation using gauge observations: A case study in Chile. Journal of Geophysical Research. https://doi.org/10.1002/2015JD024540

Yilmaz, K. K., Adler, R. F., Tian, Y., Hong, Y., & Pierce, H. F. (2010). Evaluation of a satellite-based global flood monitoring system. International Journal of Remote Sensing, 31(14), 3763–3782.

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., Grams, H., Wang, Y., CoCKs, S., Martinaitis, S., Arthur, A., Cooper, K., Brogden, J., Kitzmiller D., 2016. Multi-Radar Multi-Sensor (MRMS) Quantitative Precipitation Estimation: Initial Operating Capabilities", Bulletin of the American Meteorological Society. DOI:10.1175/BAMS-D-14-00174.1.

Zhang, C., Chen, X., Shao, H., Chen, S., Liu, T., Chen, C., Ding, Q., Du, H., 2018. Evaluation and intercomparison of high-resolution satellite precipitation estimates – GPM, TRMM, and CMORPH in the Tianshan Mountain Area. Remote Sens. 2018 (10), 1543. https://doi.org/10.3390/rs10101543.

Chapter 3

3. Evaluation of remotely sensed precipitation products across Canada

3.1. Introduction

Evaluating RSEs has been an important task to improve the quality of these datasets and their use in several hydrological and meteorological applications (Gebregiorgis at al., 2017). The superior advantages associated with RSEs to rain gauge data such as continuous high spatial resolution, coverage over remote/complex areas, and easy accessibility make them an appropriate resource for estimating precipitation, particularly over ungauged regions. The specific data of RSEs are provided by either satellite or radar which uses IR- and MW-based radiowaves to estimate the precipitation amount and frequency. Due to the number of IR and MW based sensors utilized, RSEs indicate different levels of spatial and temporal uncertainties which are not constant over different climatic and topographic regions. As previously mentioned in Chapter 2, IMERG among other global SPPs by using an enhanced technique of combining more number of IR- and MW-based sensors as well as CMORPH algorithms, provides more accurate precipitation estimates at fine spatiotemporal scales. Despite the quality of IMERG, different levels of uncertainty have been reported for this data in several recent publications (Sungmin et al. 2017; Tang et al. 2017; Wang et al. 2018; Asong et al. 2017; Tan et al. 2019). Thus, the accuracy of IMERG precipitation products needs to be assessed against in situ observations at each specific region.

In addition to SPPs, radar precipitation data provides fine spatial and temporal estimates that can be used in remote areas and also as real-time input data in the hydrological models for flood prediction. The limitation concerned with the radar data is its influence by ground clutter (e.g. dust, bugs, birds, and particulates) and other non-meteorological echoes, beam blockage, and bright banding in the melting layer (Martinaitis et al. 2017). Although MRMS by integrating multiple radars has reduced the uncertainties, it still shows errors especially for cold seasons and over mountainous regions. Also, the coverage of MRMS is limited to the CONUS and the southern part of Canada. By considering the advantages and limitations regarding satellite and radar precipitation datasets as well as the lack of a dense ground observed network

across Canada, this chapter aims to perform the first comprehensive analysis of MRMS over the southern part of Canada (south of 55°N) and the latest version of IMERG (V06) precipitation products at relatively high temporal resolution (hourly and daily) with complete coverage (90°S to 90°N) across the country. The performance of both RSEs in representing different characteristics of precipitation is evaluated using a set of 15 statistical and categorical metrics. The main objectives of this chapter can be categorized as follows:

(1) Evaluate the performance of different IMERG V06 products (PrCal, PrUncal, PrIR, PrHQ) at hourly temporal resolution over Canada.

(2) Evaluate the performance of MRMS data at hourly temporal scale over the southern parts of Canada as the coverage of radar data is limited to this area.

(3) Identify climatic and topographic conditions as well as different cold/warm seasons, across Canada, in which these products provide reliable precipitation estiamtes

3.2. Study Area

The study area in this research is Canada as the second-largest country in the world with an area of 9.9 million km2. Canada is surrounded by the Oceans of Pacific in the west, Atlantic in the east, and the Arctic in the north. The country has a diverse hydroclimate due to its extensive geographical features, latitudinal extent, and topographic variations. Polar and Arctic climate is dominant in the northern parts and regions on the west experience temperate climate with heavy precipitation associated with air currents from the Pacific while the east coast has less rainfall. The presence of the Great Lakes can moderate the weather in southern parts of Ontario and Quebec with hot, humid summers and cold, snowy winters (Asong et al. 2017, Singh et al. 2019). By having these diverse climatic conditions as well as extreme weather events experienced over most parts of the country each year (Singh and Najafi, 2020), the necessity of reliable precipitation measurements is of great importance. However, Canada like several other countries suffer from a low density of in-situ precipitation stations and also unevenly distributed data over the country. I addition, other problems with ground-based data such as a gap in the measurements, the quality control process, and susceptibility of the error have been reported (Mekis et al., 2018, Singh and Najafi, 2020). These limitations lead to low spatial and

temporal resolutions of precipitation measurements across the country. Figure 1 shows the hourly rain gauges distributed as well as seven study zones over Canada. The reason for taking the seven zones is the different time that each zone has to better match the hourly local time rain gauges with the corresponding UTC-based (Coordinated Universal Time) satellite data.
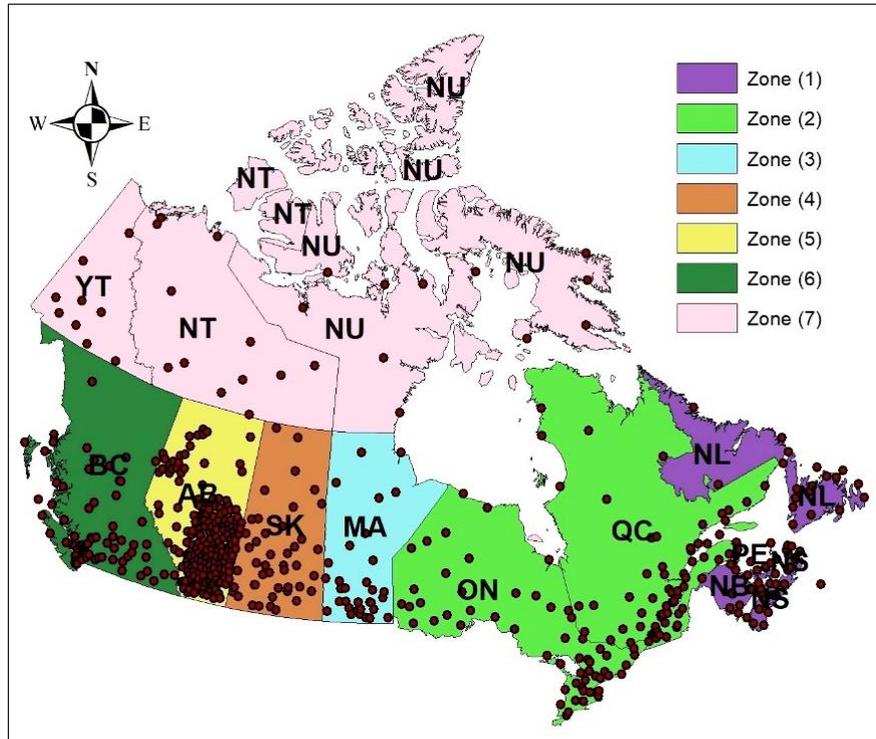
Figure 1. The location of hourly rain gauges across the study area. AB: *Alberta*, BC: *British Columbia*, MA: *Manitoba*, NB: *New Brunswick*, NL: *Newfoundland and Labrador*, NT: *Northwest Territories*, NS: *Nova Scotia*, NU: *Nunavut*, ON: *Ontario*, PE: *Prince Edward Island*, QC: *Quebec*, SK: *Saskatchewan*, YT: *Yukon* Territories

3.3. Data

3.3.1. Ground-Based Observations

We assess the amount of precipitation total at daily and hourly time scales. The hourly ground-based precipitation records are available from the automatic station network operated by Environment and Climate Change Canada (ECCC). The network consists of 585 fully automated stations, including both Surface Weather and Reference Climate Stations (RCS). Parameters that are typically observed at these locations are air temperature, humidity, precipitation accumulation, precipitation intensity, snow depth, air pressure, and wind speed

and direction. Three types of instruments are used for automatic measurement of total precipitation amount in Canada, namely weighing gauges, Tipping Bucket Rain Gauges (TBRGs), and optical sensors (Mekis et al. 2018). The total hourly precipitation is estimated as the sum of the four 15-minute precipitation amounts for minutes 00 through 60, inclusive. Precipitation amounts are stored in mm with a resolution of 0.1 mm. The quarter-hour total precipitation amounts are derived over 15-minute intervals (00-15, 15-30, 30-45, 45-60) by taking the difference of the gauge weight between the end and start of each period (Technical Documentation Digital Archive of Canadian Climatological Data, ECCC). ECCC operates several quality control checks to correct the existing errors in the RCS hourly weather stations after 2014, however before this year quality checks were not implemented at the ingest stage (Technical Documentation Digital Archive of Canadian Climatological Data, ECCC, 2018).

In this study, we select 530 hourly station records (assigned as HLY01 total precipitation in digital archive) with less than 10% missing data in each month over the five years starting from 2014 to the end of 2018. For daily evaluations of IMERG, 325 quality controlled daily station records provided by ECCC (download link: https://climate-change.canada.ca/climate-data/#/daily-climate-data) with less than 10% missing data in each month are collected. This data has been assigned as DLY02 in digital archive and has received some level of QC. Since the source of daily data provided by ECCC is different from the hourly data, we aggregated the corresponding hourly stations with daily ones to check the consistency between two data sets and found that more than 70 percent of hourly stations are in agreement with the daily data. As shown in Figure 1, rain gauges are not evenly distributed across Canada and the density is higher in southern parts of the country.

The reliability of automatic precipitation instruments for solid precipitation measurement can be undermined due to the blockage of the orifice by snow capping the gauge or accumulating on the side of the orifice walls, wind undercatch of snow due to the formation of updrafts over the gauge orifice, the unknown role of turbulence on gauge catch, and the large variability in gauge catch efficiency for a given gauge and wind speed (Rasmussen et al. 2012). Because of the highly variable nature of snow depth and the unreliability of measurements, direct snowfall observation is no longer derived from ECCC automatic stations since December 2013 until further improvements are developed for this parameter (Merkis et al. 2018).

3.3.2. IMERG Satellite Data

In this study, the recently released (June 2019) version (V06B) of IMERG mission Final Run with high spatial (0.1°) and temporal (30 min) resolution is analyzed for 2014 to 2018. The IMERG algorithm is intended to inter-calibrate, merge, and interpolate "all" satellite microwave precipitation estimates, together with microwave-calibrated infrared satellite estimates, and monthly precipitation gauge records (Huffman et al. 2019a). The IMERG V06 data are available globally from -90° S to 90° N latitude with three Early (~4 hours after observation time), Late (~14 hours after observation time), and Final (~3.5 months after the observation time) runs to accommodate different user requirements for latency and accuracy (Tan et al. 2019). The post-real-time Final Run uses the Global Precipitation Climatology Center (GPCC) monthly precipitation gauge analysis and the European Centre for Medium-Range Weather Forecasts (ECMWF) ancillary data for calibration. Therefore, this product is expected to provide the most reliable estimates that are suited for research works (Huffman et al. 2019a). In order to create the final half-hourly calibrated IMERG precipitation estimates, the ratio between the monthly accumulation of half-hourly multi-satellite-only fields and the monthly satellite-gauge field (satellite calibrated with monthly gauges) is computed. Next, each half-hourly field of multi-satellite-only precipitation estimates in the month is multiplied by the ratio field (Huffman et al. 2019a). It should be stated that, ~90% of gauges considered in this study have not been used for IMERG calibration (Mekis et al. 2018). In addition, the applied ratio does not remove biases at sub-monthly scales (i.e. hourly and daily).

The four different precipitation fields of IMERG data are categorized as Calibrated precipitation (precipitationCal), which represents records after the final post-processing step described above, Uncalibrated precipitation (precipitationUncal), which is recorded data before the final post-processing step (precipitationCal and precipitationUncal fields are identical for the Early and Late Runs, as there are no additional corrections applied), Infrared (IR) geostationary satellite precipitation data (IRprecipitation), and precipitation extracted from merging High-Quality Passive Microwave (PMW) sensors (HQprecipitation), which only includes microwave data and has significant gaps. PrecipitationCal is considered as the most reliable IMERG precipitation estimate (Huffman et al. 2019b).

IMERG V06 has some major improvements over previous versions. First, to drive the morphing scheme it uses total precipitable water vapor from reanalysis data. However, previous versions of IMERG adopt geostationary infrared (GEO-IR) data to calculate the motion vectors of precipitation systems, which leads to the mismatch between IR-based cloud-top motions and surface precipitation motions. Second, passive microwave estimates are morphed at high latitudes to reduce spatial gaps (Huffman et al. 2019b; Tang et al. 2020). Third, the latest version of the Goddard Profiling Algorithm (GPROF2017) ingested in IMERG V06 retrieves total hydrometeor mass in the atmospheric column, except for the conical-scan imager PMW retrievals, which only considers total solid hydrometeor mass over land and coast and then implicitly correlates it to surface precipitation in any phase including rain, drizzle, snow, and hail (Huffman et al. 2019b). Further, IMERG V06 includes a new data field called the probability of liquid precipitation, which provides different phases of the precipitation (i.e., liquid, solid, or mixed). In this study, we evaluate the total precipitation amounts derived from IMERG V06 retrieval products. Analysis of different phases of precipitation will be considered in future studies.

3.3.4. MRMS Product

The radar-based precipitation data utilized in this study are derived from a product of MRMS named Surface Precipitation Rate (SPR). SPR uses a quality-controlled reflectivity product called the Seamless Hybrid Scan Reflectivity (SHSR) mosaic and Surface Precipitation Type (SPT) field to compute instantaneous rain rates in mm h^{-1} (Zhang et al. 2016). SHSR is first derived from single radar polar grids, and then mosaicked onto the MRMS national Cartesian grid (Grams et al 2014). The MRMS domain extends from 20°N to 55°N latitude and from 130°W to 60°W longitude with a horizontal resolution of 0.01°×0.01° (Figure 1 in Zhang et al. 2016). MRMS ingests 3D volume scan data from 146 S-band dual-polarization WSR-88D in the US and 30 C-band single-polarization weather radars operated by ECCC in Canada. The gauge quality-controlled data of MRMS is integrated with atmospheric environmental data (such as surface and wet-bulb temperatures, wind and relative humidity extracted from NWP model), lightning, and rain gauge observations to generate a suite of severe weather and QPE products (Zhang et al. 2016).

In this study, the total precipitation amount extracted from MRMS SPR product as radar-only QPE (not bias corrected by local gauges) is used to avoid errors associated with the limitations of rain gauge measurements and interpolation method applied in the local gauge bias-corrected product. The misrepresentation of ground-based winter precipitation can influence local bias correction values (Martinaitis et al. 2015). The SPR data is available from November 1st, 2014 till present at a temporal resolution of 2-minutes. We perform the evaluations for the time period of 2015 until the end of 2018.

There are several factors that can increase the uncertainties in radar precipitation estimations particularly during cold season. Radar variables are indirect measurements of precipitation rates (R), therefore empirical relationships between radar reflectivity (Z) and (R) are developed to derive radar QPE. Different empirical relationships are required for different precipitation phases and regimes. An automated surface precipitation classification is employed in MRMS such that appropriate relationships may be applied. Some major uncertainties of radar QPE products are associated with improper calibration and limited operational Z–R and Z–S (liquid equivalent snowfall rate) relationships due to differing snowfall properties. Also, highly variable falling speeds of snow can introduce spatial and temporal uncertainties in winter precipitation estimation. This can cause significant elapsed time between radar detection aloft and ground measurement (Martinaitis et al. 2015). In addition, the VPR correction applied in the SHSR field for mitigating radar errors does not work when the surface temperature is below 0°C and hence is not useful in snow detection. Further, the correction usually works better on flat land than on complex terrain, where orographic forcing modulates precipitation distributions (Zhang et al. 2016). Similar to IMERG, we analyze the ability of MRMS in detecting the total amount of precipitation across the Canadian domain.

3.4. Methodology

This section describes the evaluation procedure of multiple IMERG V06B satellite precipitation products and MRMS at different temporal scales across Canada. The IMERG data are compared against rain gauge records at two temporal scales: 1) hourly- that analyses precipitation estimates from four products including precipitationCal (hereafter PrCal), precipitationUncal (PrUncal), HQprecipitation (PrHQ), and IRprecipitation (PrIR) obtained from different IR and PMW sensors and 2) daily- that is performed for the widely-used PrCal

27

dataset. For both (hourly and daily) analyses, the half-hourly IMERG data are aggregated and matched with the local hourly/daily gauge records. The UTC-based (Coordinated Universal Time) satellite data is processed to be consistent with local records considering the seven different time zones corresponding to different rain gauges across Canada as well as daylight saving times for almost half of the year in many parts of the country.

Evaluation of the MRMS product (with a 2min/1km temporal and spatial resolution) is performed at hourly timescale using gauges that cover up to 55° N latitude. All analyses are performed by evaluating the gridded RSEs at locations where there are rain gauges available. A direct comparison between gauge points and their corresponding satellite/radar pixels is conducted at each given time separately. Therefore, no transformation and interpolation from the points to areal precipitation data are made to prevent the uncertainty associated with the spatial estimates of precipitation, especially for the areas with fewer gauges. To have a more reliable assessment of the IMERG and MRMS precipitation estimates the corresponding biases are characterized for different seasons, winter (DJF), spring (MAM), summer (JJA), and fall (SON). Several continuous and categorical evaluation metrics are used to assess the ability of RSE to detect rainfall occurrence and amount.

3.4.1. Continuous Verification Metrics

Continuous indices are used to measure the accuracy of the estimated precipitation magnitudes from IMERG and MRMS data. The widely used metrics including root mean square error (RMSE), mean absolute error (MAE), relative bias (Rbias), and Pearson correlation coefficient (CC) are applied. In addition, four statistical indices namely hit bias (Hbias), miss bias (Mbias), false bias (Fbias), and correct negative bias (CNbias) are considered to quantify the error characteristics of RSE associated with detectability performance. The equations and a brief description of these metrics are listed in Table 2. Rbias describes the systematic biases of RSEs, MAE is used to represent the overall errors of the QPEs without considering their directions, RMSE is used to measure the average error magnitude, which gives greater weights to the larger errors relative to MAE, and CC characterizes the degrees of consistencies in temporal variabilities. Hbias, Mbias, Fbias, and CNbias display the systematic biases of RSEs associated with hit, miss, false, and non- events, respectively. Hit events refer to hourly/daily records where both RSE and ground-based precipitation values are more than 0.1 mm h^{-1}. Miss events correspond to gauge records more than 0.1 mm h^{-1}, while RSEs are less than 0.1 mm h^{-1}.

Contrary to the miss events, false events are associated with higher than 0.1 mm h^{-1} of rain detected by satellite/radar while no precipitation has been recorded by rain gauges. Finally, non-events represent the conditions when both satellite/radar and gauge records show precipitation values less than 0.1 mm h^{-1}. The range of R, H, M, F, and CN bias is between -∞ and +∞ with the optimal value of 0. MAE and RMSE vary between 0 to +∞, and CC ranges from -1 to +1. Larger errors are associated with larger Rbias, MAE, RMSE, Hbias, Mbias, Fbias, and CNbias values.

Table 2. List of the continuous verification metrics used to evaluate the error characteristics of IMERG and MRMS products. P_{S_i} is the value of RSE for the ith hourly/daily record, P_{O_i} is the value of ground-based precipitation observation for the ith hourly/daily record, N is the total number of records, $\bar{P}_S$ is the average value of RSEs for N hourly/daily records over each grid cell, and $\bar{P}_O$ is the average value of rain gauge observations for N records over each cell. The superscripts H, M, F, and NO represent precipitation estimates for hit, miss, false, and non-events, respectively

Continuous metrics	Equation	Perfect value	Description
Rbias	$\dfrac{\sum_{i=1}^{N}(P_{S_i} - P_{O_i})}{\sum_{i=1}^{N} P_{O_i}} \times 100\%$	0	Percentage difference between gauge observations and RSEs for all events Rbias > 0: overestimation; Rbias < 0: underestimation
MAE	$\dfrac{1}{N} \times \sum_{i=1}^{N} \left| P_{O_i} - P_{S_i} \right|$	0	Mean absolute error between ground-based observations and RSEs
RMSE	$\sqrt{\dfrac{\sum_{i=I}^{N}(P_{O_i} - P_{S_i})^2}{N}}$	0	Root mean square error between gauge observations and RSEs
CC	$\dfrac{\sum_{i=1}^{N}(P_{O_i} - \overline{P_O})(P_{S_i} - \overline{P_S})}{\sqrt{\sum_{i=1}^{N}(P_{O_i} - \overline{P_O})^2 \sum_{i=1}^{N}(P_{S_i} - \overline{P_S})^2}}$	1	Pearson correlation coefficient that measures linear correlation between gauge observations and RSEs
Hbias	$\dfrac{\sum_{i=1}^{N}(P_{S_i}^{H} - P_{O_i}^{H})}{\sum_{i=1}^{N} P_{O_i}} \times 100\%$	0	Percentage difference between gauge observations and RSEs for hit events Hbias > 0: overestimation; Hbias < 0: underestimation
Mbias	$\dfrac{\sum_{i=1}^{N}(P_{S_i}^{M} - P_{O_i}^{M})}{\sum_{i=1}^{N} P_{O_i}} \times 100\%$	0	Percentage difference between gauge observations and RSEs for miss events by RSEs Mbias > 0: overestimation; Mbias < 0: underestimation
Fbias	$\dfrac{\sum_{i=1}^{N}(P_{S_i}^{F} - P_{O_i}^{F})}{\sum_{i=1}^{N} P_{O_i}} \times 100\%$	0	Percentage difference between gauge observations and RSEs for false events by RSEs Fbias > 0: overestimation; Fbias < 0: underestimation
CNbias	$\dfrac{\sum_{i=1}^{N}(P_{S_i}^{NO} - P_{O_i}^{NO})}{\sum_{i=1}^{N} P_{O_i}} \times 100\%$	0	Percentage difference between gauge observations and RSEs for non-events by RSEs

3.4.2. Categorical Verification Metrics

To measure the ability of RSE data to detect rain/no rain events (based on a threshold of 0.1 mm h^{-1}), seven categorical metrics listed in Table 3 are utilized, which include hit (H), miss (M), false (F), and correct negative (CN) fractions, as well as the probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI).

Table 3. List of categorical metrics used to evaluate the error characteristics of IMERG and MRMS products. n_H is the number of hit events (both the observed and estimated precipitation values are equal to or more than 0.1 mm h^{-1}), n_M represents the number of miss events (observed data show precipitation events > 0.1 mm h^{-1} but the RSE data miss them), n_F denotes the number of false events (non-event based on the observed data, however satellite/radar detects precipitation), n_{NO} indicates the number of non-events (none of the observed and estimated precipitation values are equal to or more than 0.1 mm h^{-1})

Categorical metrics	Formula	Perfect value	Description
Hit fraction	$\dfrac{n_H}{n_H + n_M + n_F + n_{NO}} \times 100\%$	Hit + Correct Negative fractions = 100%	Fraction of correctly detected events by SREs relative to all rainfall events
Miss fraction	$\dfrac{n_M}{n_H + n_M + n_F + n_{NO}} \times 100\%$	0	Fraction of missed rainfall events by SREs relative to all rainfall events
False fraction	$\dfrac{n_F}{n_H + n_M + n_F + n_{NO}} \times 100\%$	0	Fraction between false alarm events by SREs and all rainfall events
Correct Negative fraction	$\dfrac{n_{NO}}{n_H + n_M + n_F + n_{NO}} \times 100\%$	Hit + Correct Negative fractions = 100%	Fraction of correctly detected non-events by SREs relative to all rainfall events
POD	$\dfrac{n_H}{n_H + n_M}$	1	Fraction of gauge observed events that were correctly detected by SREs
FAR	$\dfrac{n_F}{n_H + n_F}$	0	Fraction of detected events by SREs that were not observed by gauges
CSI	$\dfrac{n_H}{n_H + n_M + n_F}$	1	Fraction of gauge observed events that were correctly detected by SREs with no consideration of correct negative events

The optimal performance corresponds to M and F fraction values of 0. Values of the H and CN fractions depend on the number of rain/no-rain events, however, their optimal sum is 100%. POD, FAR, and CSI range between 0 and 1 with the optimal values of 1, 0, and 1, respectively. POD is sensitive to hits, but ignores false alarms, while FAR is sensitive to false alarms and ignores misses. POD and FAR are both very sensitive to the climatological frequency of the event and should be used in conjunction. CSI is sensitive to hits and penalizes both misses and false alarms. It also depends on climatological frequency of events (poor scores for rare events) since some hits can occur purely due to random chance.

3.5. Results

The evaluation results of GPM-IMERG V06 and MRMS SPR products based on ECCC's rain gauge records are presented. First, we assess the performance of IMERG PrCal at daily, and then the four IMERG products and MRMS at hourly timescales.

3.5.1. Evaluation of GPM-IMERG V06 at Daily Timescale

Figure 2 shows the RMSE (a measure of bias) and CC (representing consistencies in temporal variations) values corresponding to the IMERG PrCal product at daily timescale for each season across Canada. Results are shown for the 10 km pixels that include ground-based observations for the five-year period of 2014 to 2018. Overall, the biases (based on RMSE) are lower in the Prairie provinces including Manitoba, Saskatchewan, Alberta (zones-3-5) and eastern parts of British Columbia (zone-6). Although the temporal variations of the IMERG estimated precipitation are more consistent with gauge data records over the east coast, the corresponding magnitudes are less accurate compared to other regions. RMSE and CC values vary across seasons. During fall and summer, the accuracy of IMERG is relatively high in several sites, while winter shows weaker performance with lower correlations between IMERG and ground observations, which was expected as discussed in the Data section.
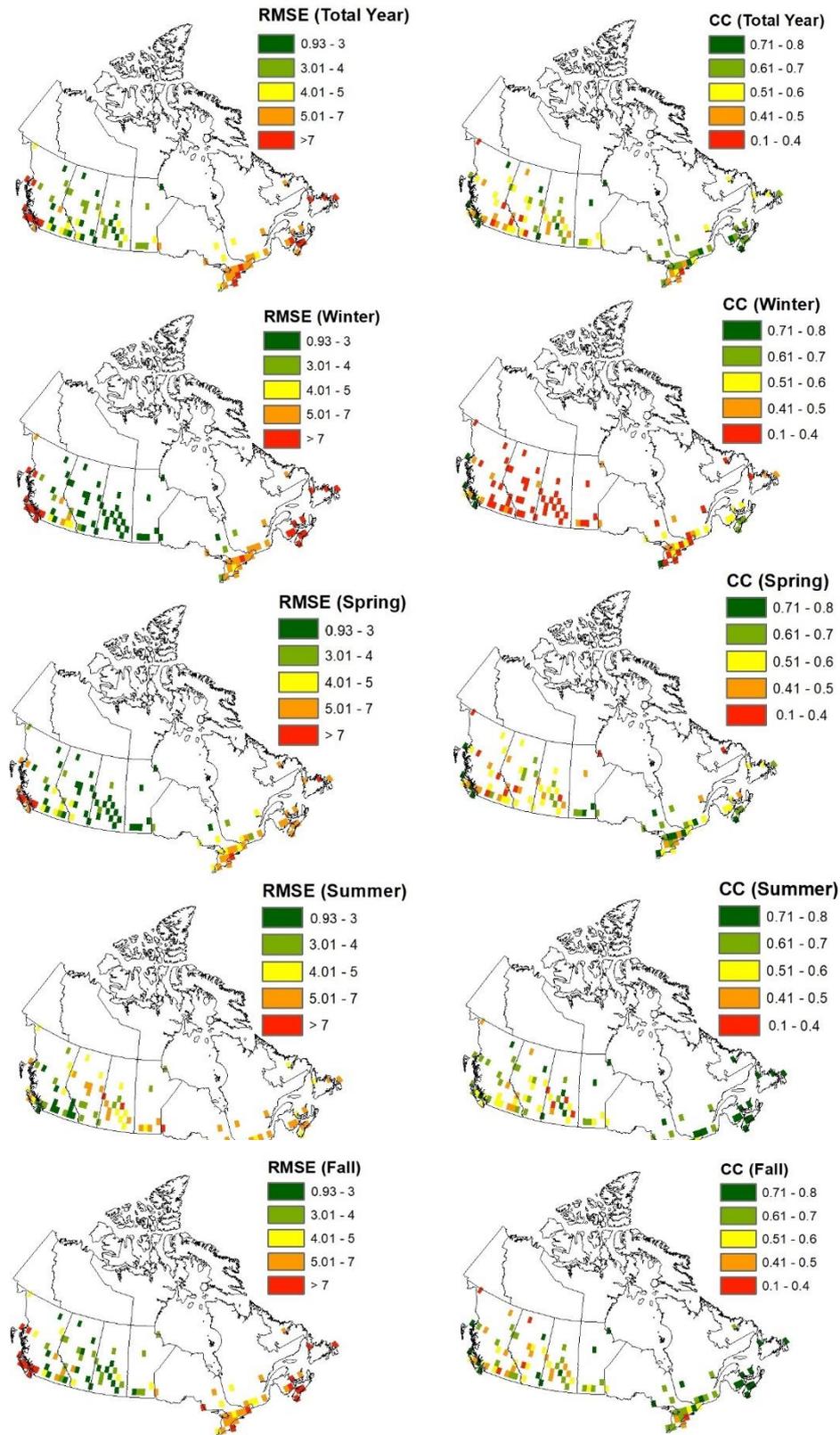
Figure 2. Spatially distributed biases of daily IMERG V06 across 325 sites based on RMSE (mm d-1) (left) and correlation (right) for different seasons, and during the entire five-year period (Total Year)

The overall (five-year) and seasonal performance of daily IMERG precipitation estimates are evaluated based on Rbias, MAE, RMSE, and CC in Figure 3. Box-and-whisker plots show the first (Q1) and third (Q3) quartiles (i.e. Interquartile Range), maximum (Q3 + 1.5×IQR) − minimum (Q1-1.5×IQR) values (whiskers), and the medians of the metrics between all sites across Canada. Positive values of Rbias indicate the tendency of IMERG to overestimate precipitation, which is more considerable in winter (Rbias varies between 10%-50%) compared to the other seasons (5%-25%), as expected. MAE and RMSE (which emphasizes on biases in extremes) values are consistent across seasons with larger variations in winter. The MAE ranges between 1.4-3.1 mm d^{-1} (interquartile range) and RMSE ranges from ~3.5 to ~6.5 mm d^{-1} across all sites. The best agreement between IMERG estimates and the observed data, according to the CC index, is in the fall and summer with average values ranging between 0.5 to 0.7.
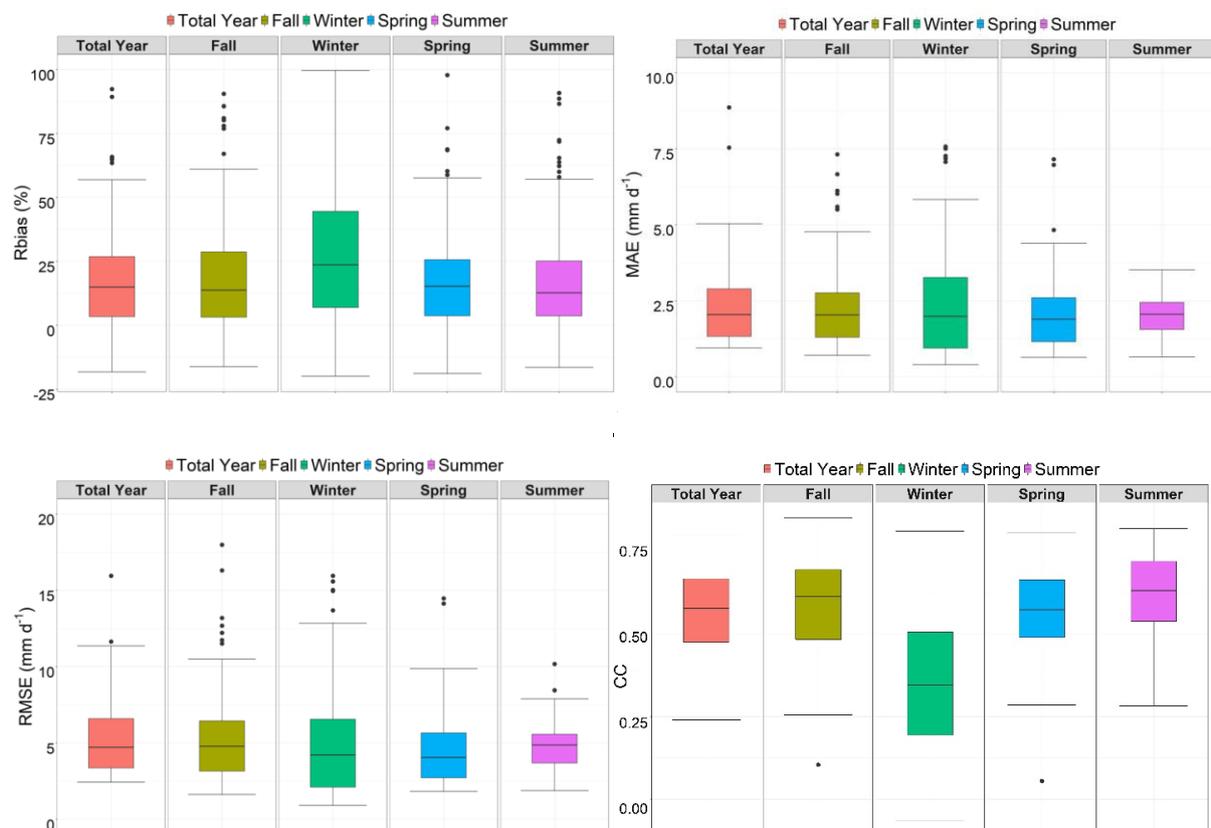
Figure 3. Performance evaluation of IMERG V06 at daily timescale based on Rbias, MAE, RMSE, and CC; Box-and-whisker plots show the first (Q1) and third (Q3) quartiles (i.e. Interquartile Range), maximum (Q3 + 1.5×IQR) − minimum (Q1-1.5×IQR) values (whiskers), and the medians of the metrics between all sites across Canada

The performance of IMERG in detecting the occurrence of precipitation is evaluated using POD, FAR, and CSI (Figure 4). Overall, the results are promising given that the median values of POD and CSI are mostly above 0.75 and 0.5, respectively indicating that the precipitation occurrence is often captured by satellite records. FAR values (~0.3 – 0.5 varying between locations and seasons) imply that the RSE product incorrectly shows the occurrence of precipitation in about 35% (median) of non-events. IMERG performance is best in the summer with POD ~ 0.78 – 0.88 and FAR ~ 0.35 – 0.45, and worst in the winter with POD ~ 0.5 – 0.75 and FAR ~ 0.22 – 0.6.
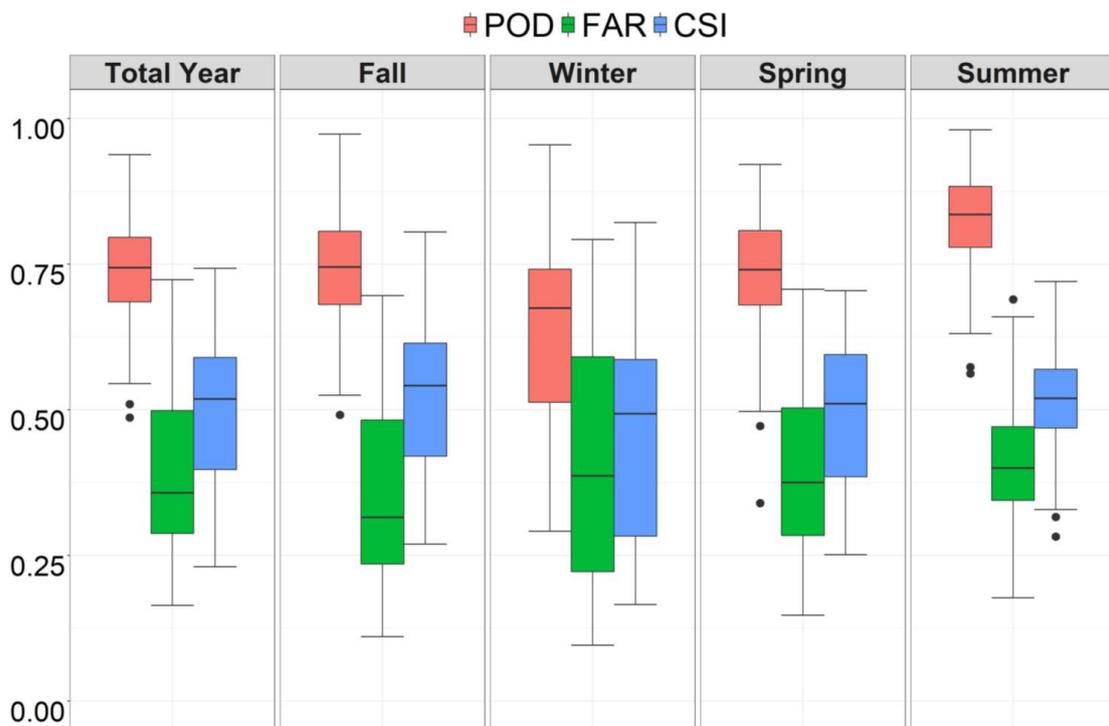
Figure 4. Evaluation of IMERG V06 based on POD, FAR, and CSI

We perform further analysis of IMERG biases using metrics that quantify the misrepresentation of the amounts of precipitation (Figure 5). The IMERG product has the highest hit, miss, and false biases in winter compared to the other seasons indicating its worst performance during the cold season, while it shows better performance in warmer periods (i.e. summer). Positive values of the hit bias indicate that IMERG overestimates the observed precipitation amount by ~10% on average, which is in agreement with the results from other metrics such as Rbias, MAE, and RMSE. ~30% of all days (averaged across all sites) within the five-year period experienced more than 0.1 mm d^{-1} of precipitation, which is correctly detected by the IMERG

V06 PrCal product (Figure 6). In addition, IMERG represents ~40% of no-precipitation days and hence it detects ~70% of all events, accurately. The product, however, does not capture ~11% of the total events (> 0.1 mm h^{-1}) while it provides false detection in ~20% of the total number of events.
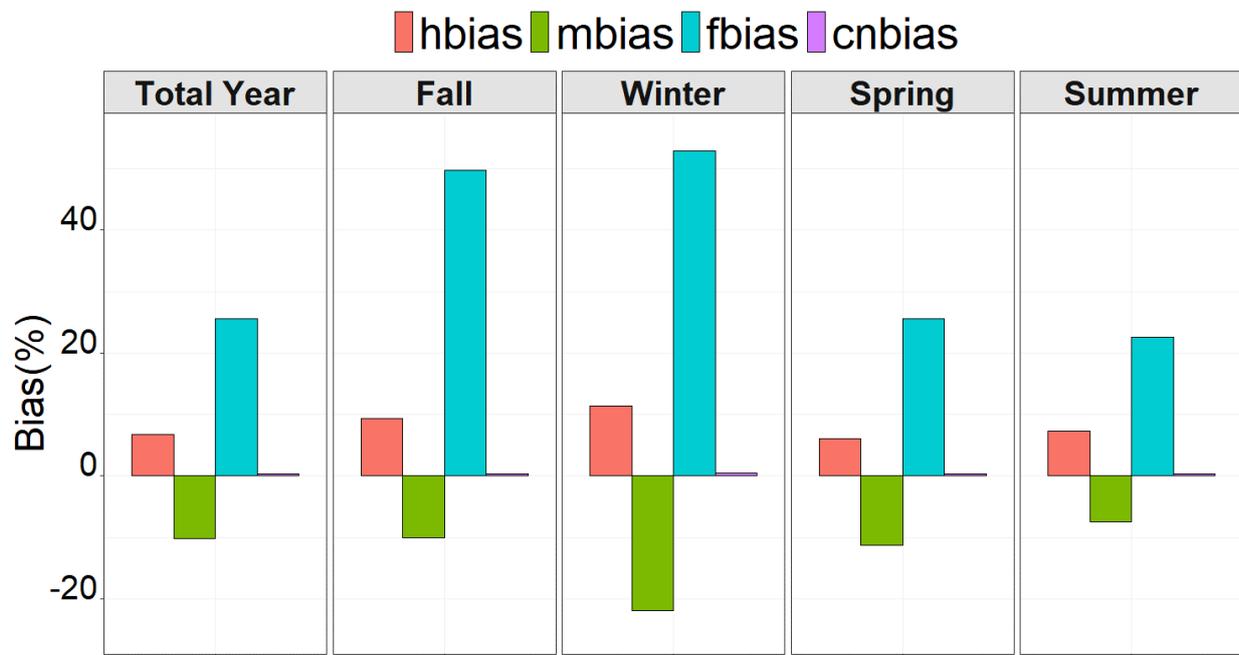
Figure 5. Barplot comparison of daily IMERG and ground-based observations based on hbias, mbias, fbias, and cnbias across Canada for 2014-2018
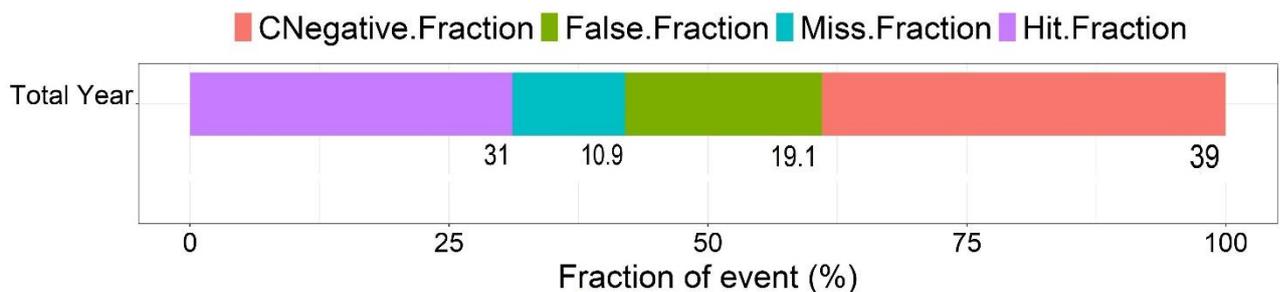
Figure 6. Fractions of daily hit, miss, false and correct negative events for the IMERG PrCal product averaged over the study area

Figure 7 evaluates the performance of IMERG in representing the "true" precipitation at different quantiles. For this purpose, the satellite and ground-based precipitation quantiles (1% to 99%) are found using the five-year daily data record at each site for different seasons, and the average values of each quantile across all sites (represented by red dots in Figure 7) are

taken. Results show that overall, IMERG tends to overestimate light to moderate precipitations (lower quantiles) particularly during summer.
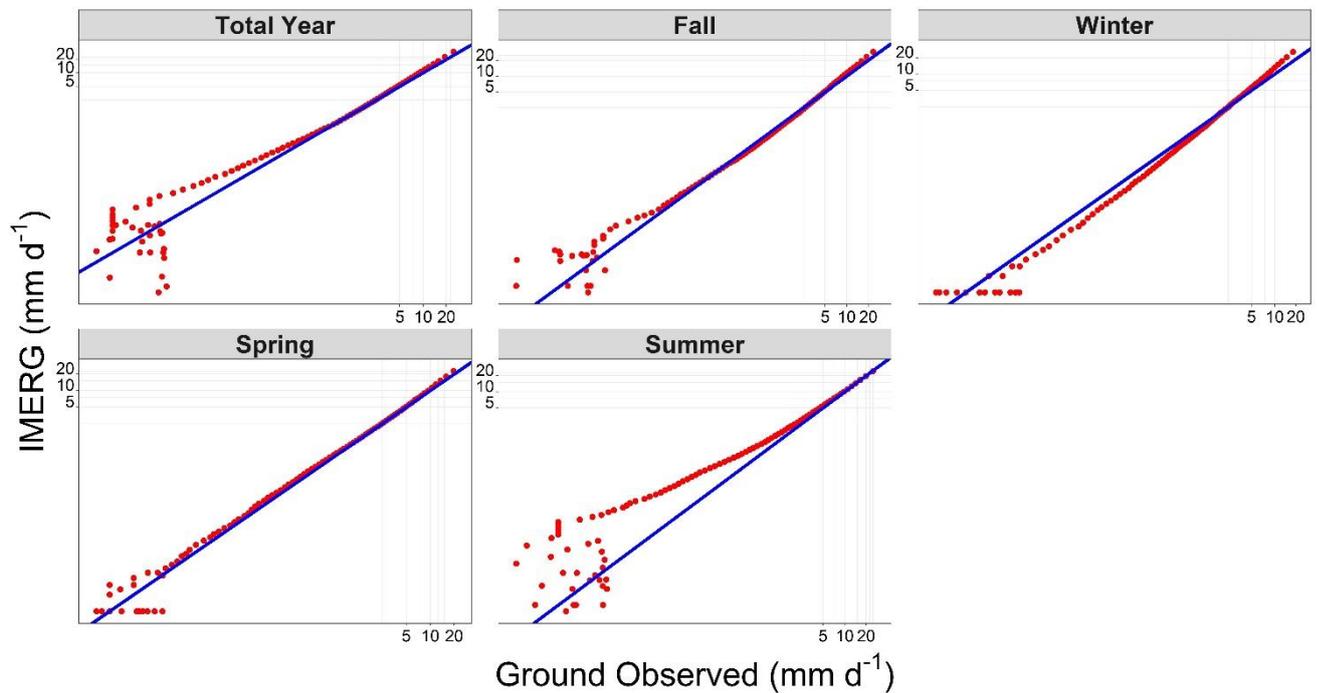
Figure 7. Comparison between the spatially averaged quantiles of the true precipitation and IMERG PrCal values (mm d^{-1}) for 2014-2018 (x and y axes are in log space)

3.5.2. Evaluation of IMERG V06 and MRMS at Hourly Timescale

Investigating intense rainfall events over short durations is critical for flood risk analysis particularly over urban areas. We evaluate precipitation estimates from four IMERG satellite products as well as the MRMS radar data at hourly timescale using ECCC's ground-based observations. This is the first analysis of the IMERG (latest product) and MRMS data over Canada at a relatively high temporal resolution. As mentioned before, MRMS collects the base level data from all radars in a network and processes them at a centralized location to produce high 1-km spatial and 2-min temporal resolution Quantitative Precipitation Estimates (QPEs). This allows for easy integration of multi-sensor data and provides enhanced QPE products (Zhang et al. 2016). Both IMERG and MRMS products are aggregated to hourly timescale to perform the evaluations.

The spatial distributions of RMSE and CC corresponding to four different IMERG products over the five-year period of 2014-2018 are shown in Figure 8. Overall, PrCal and PrUncal have more accurate estimates with lower RMSE and higher CC values compared to those of PrIR and PrHQ. All products show better performance in the Prairie provinces (Saskatchewan and Alberta) and regions on the west (British Columbia) with RMSE values ranging between 0.25-0.75 mm h^{-1} for PrCal, PrUncal, and PrIR. IMERG data can represent the temporal variability of the ground-based observations relatively well as suggested by CC values ~ 0.4 and above, particularly in eastern and western coasts and parts of central Canada.

Similar assessments are conducted for the MRMS precipitation rates across the coverage area for radar network, which includes regions that lie within 42°-55° latitude. Figure 9 represents the spatially distributed RMSE and CC values corresponding to the MRMS precipitation product for the entire period and different seasons during 2015-2018. According to both metrics, MRMS shows a satisfactory performance particularly over the regions in the east (including southwest Ontario) and Prairies with RMSE values ranging from 0.2 to 0.5 and CC values from 0.6 to 0.9 in all seasons. In fall and spring, MRMS data are more consistent with gauge records in most parts of the country. However, during winter the performance is relatively weak with RMSE values consistent with those of the other seasons but low correlations. Contrary to the MRMS performance in winter, the assessments show relatively high CC values and high RMSEs during summer. This is partly because of false estimates of the radar in no rain conditions (resulting in high RMSE), while it can detect rain events well resulting in relatively high CC values.
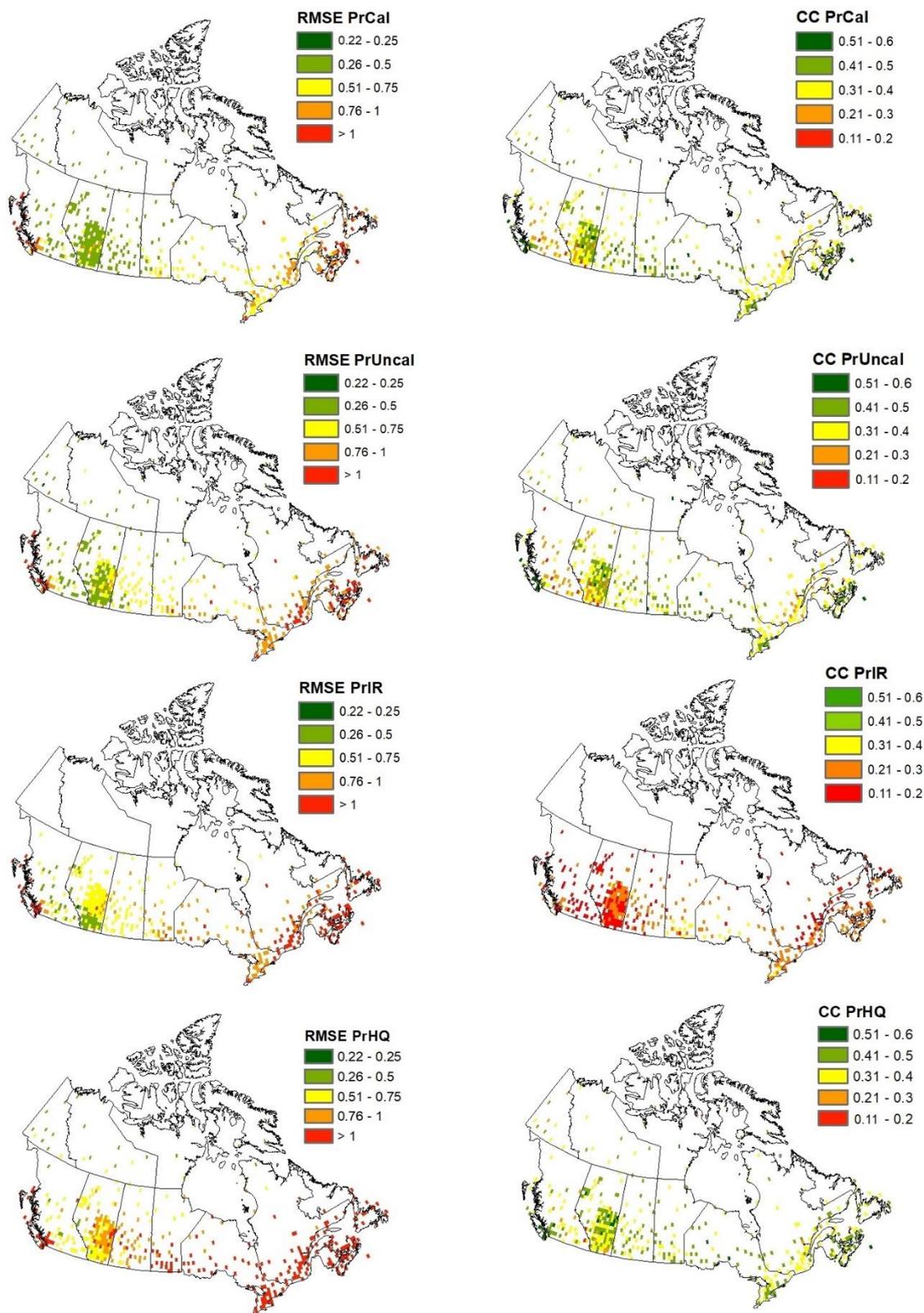
Figure 8. Spatially distributed biases of four hourly IMERG V06 products across 530 sites based on RMSE (mm h^{-1}) (left) and correlation (right). The products include PrCal (calibrated based on monthly gauge data), PrUncal (satellite only data), PrIR (Infrared-based sensor data), and PrHQ (High-Quality Passive Microwave based sensors data)
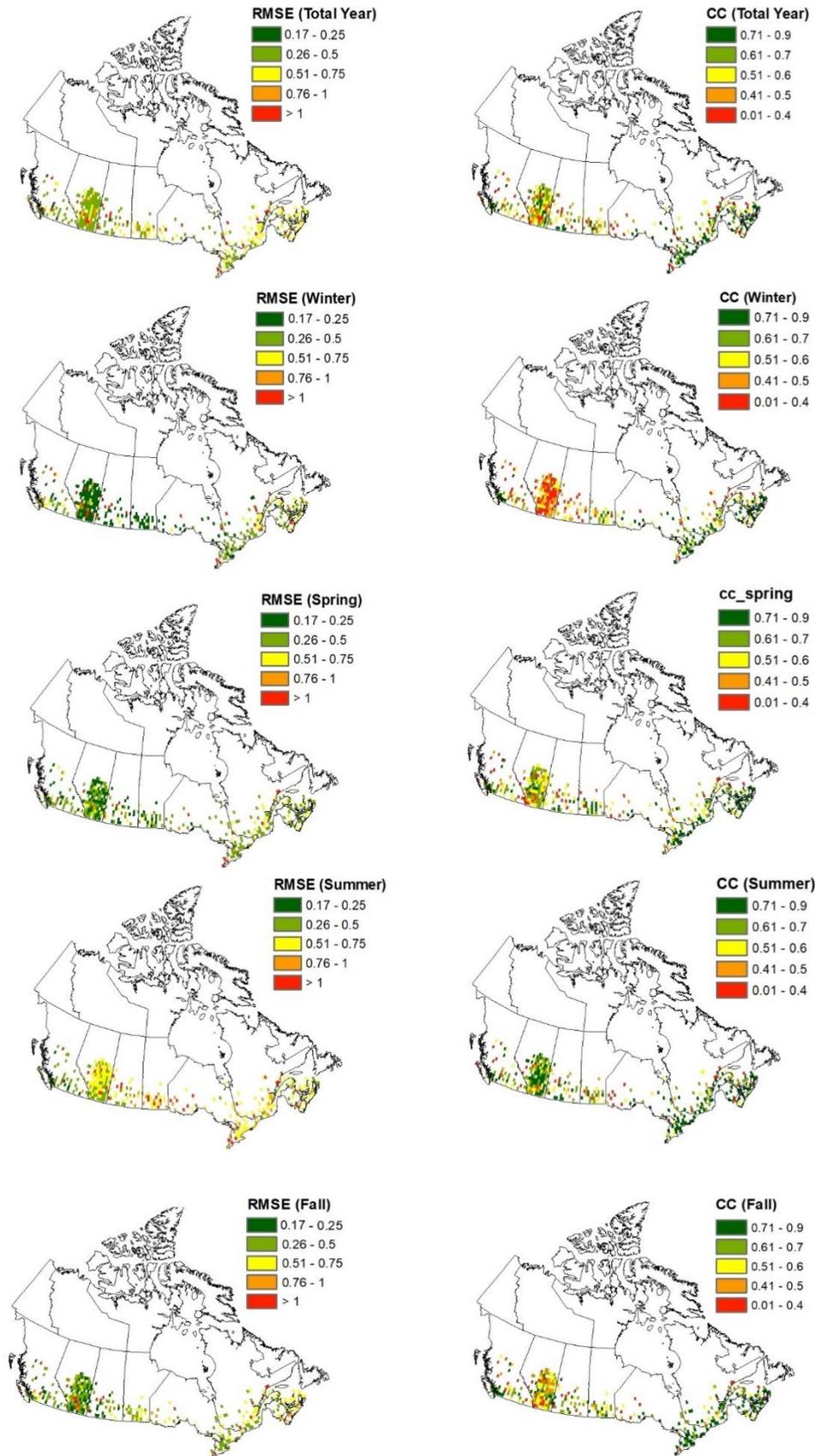
Figure 9. Spatially distributed biases of hourly MRMS data across 505 sites based on RMSE (mm h^{-1}) (left) and correlation (right) for different seasons

Further, we perform a regional evaluation of the four IMERG products as well as the MRMS precipitation estimates at hourly timescale for the seven zones defined in Figure 1. The results based on RBias, MAE, and CC for zone-1 to 7 are shown in Figure 10. MRMS shows the best performance in almost all zones and over the four seasons, although with relatively large uncertainties. Except for zone-1 where MRMS underestimates the precipitation amount, with Rbias ranging between -35% and 0, it shows overestimated values in all other zones. MRMS QPEs are most reliable in fall, spring, and summer in all zones. For example, in zone-1, best estimates are found in spring with median Rbias ~ -10%, MAE ~ 0.12 mm h^{-1}, and CC ~ 0.7, and in zone-2, fall shows the best estimates with Rbias, MAE, and CC of 5%, 0.1 mm h^{-1}, and 0.73, respectively. Winter shows the weakest correlations. Further, MRMS data have relatively strong linear association with the gauge data (CC~ 0.6-0.8) in zone-1, whereas the values of CC vary between 0.5 to 0.75 in other zones, except for winter in zones-4 and 5 in which they range between 0.3 and 0.5.

PrCal outperforms the other IMERG products in most cases, with PrUncal following closely in terms of MAE and CC. However, PrUncal shows larger overestimations compared to PrCal. Rbias values corresponding to PrCal range between 10% and 25% in zones-2-5, however, in zone-1 and zone-6 PrHQ shows better performance in terms of Rbias (0-15% and -5% to 35%, respectively). In addition, PrCal has more accurate estimates for different seasons across all zones, except for fall and spring in zone-1 and winter and summer in zone-2 in which PrHQ shows better results. Based on MAE, PrCal has lower bias and outperforms the other products across all zones and during different seasons. In zones-1 to 5 the variations of MAEs are minor, as represented by relatively short boxplots, indicating less variability in the estimates across sites. In western areas close to the Pacific (zone-6), however, boxplots are wider indicating more variability in biases between sites. MRMS outperforms all IMERG products to a large degree based on the CC metric.
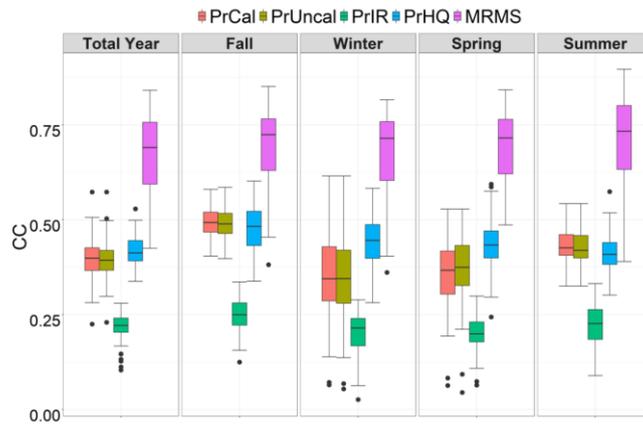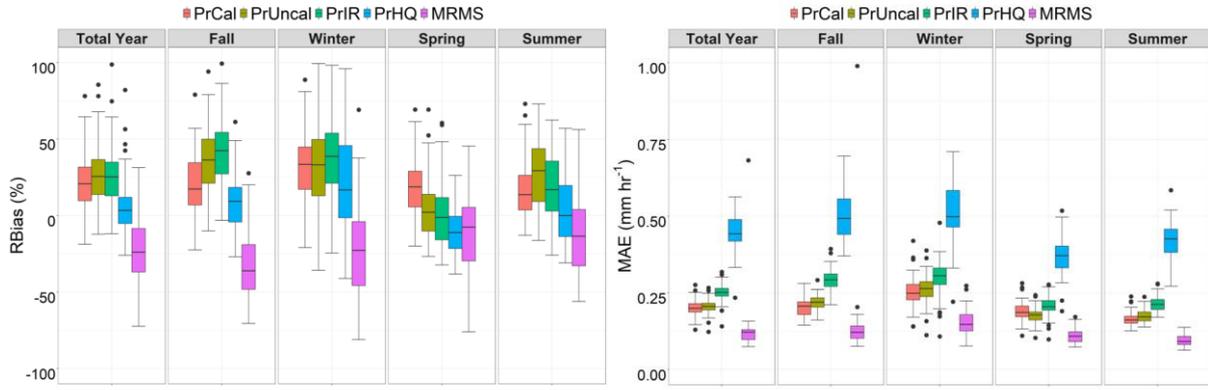
(**Zone-1**)
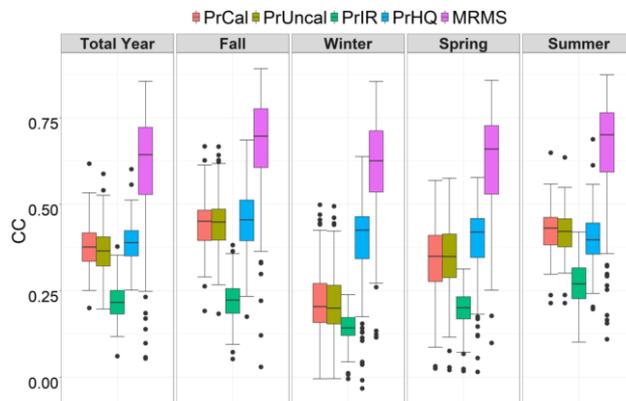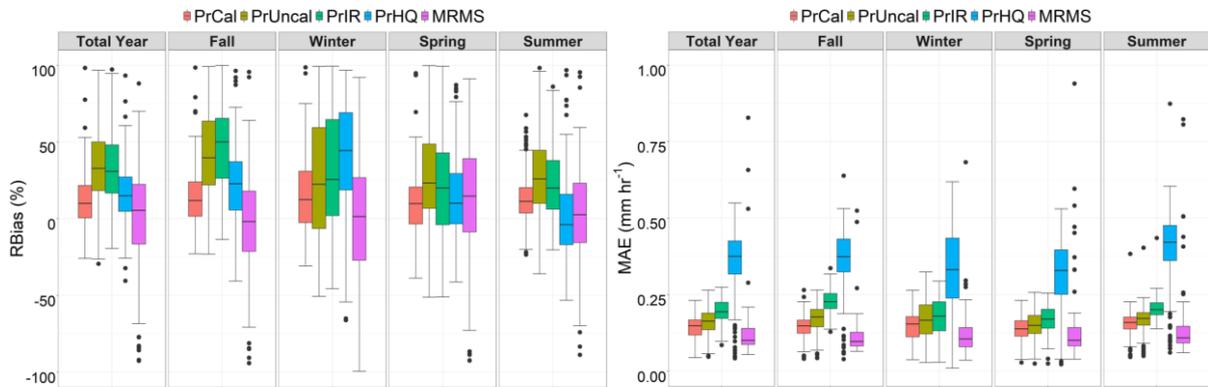
(**Zone-2**)
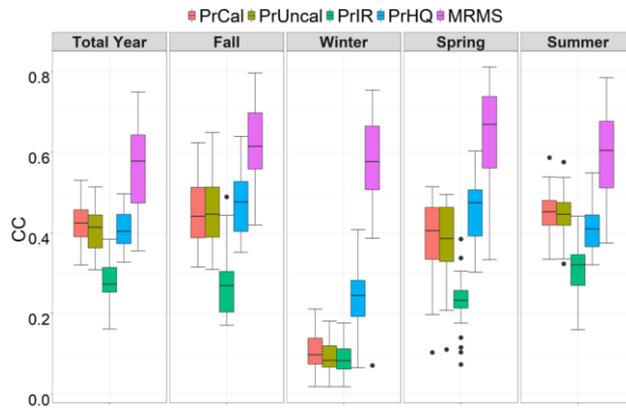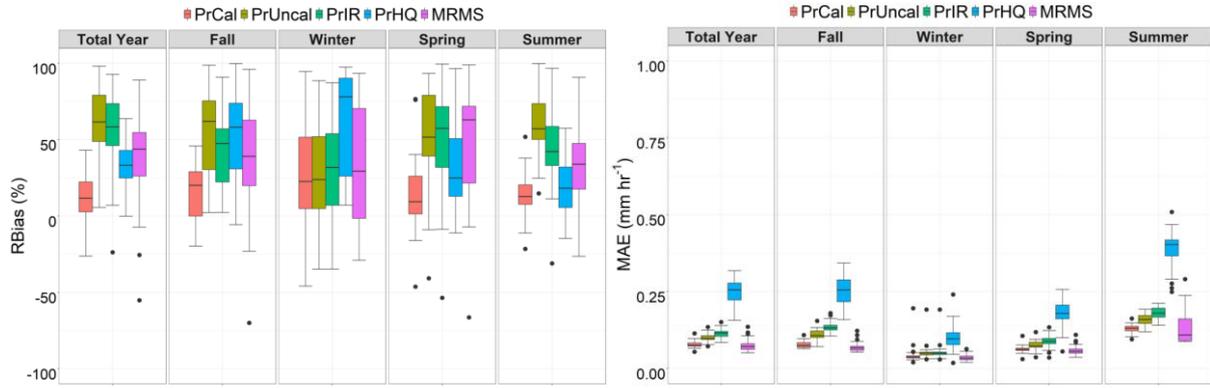
(**Zone-3**)
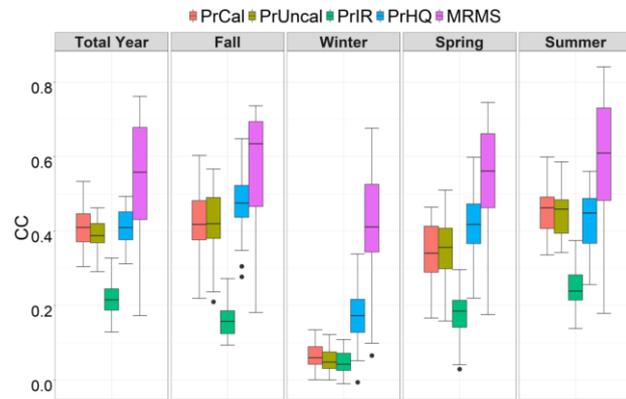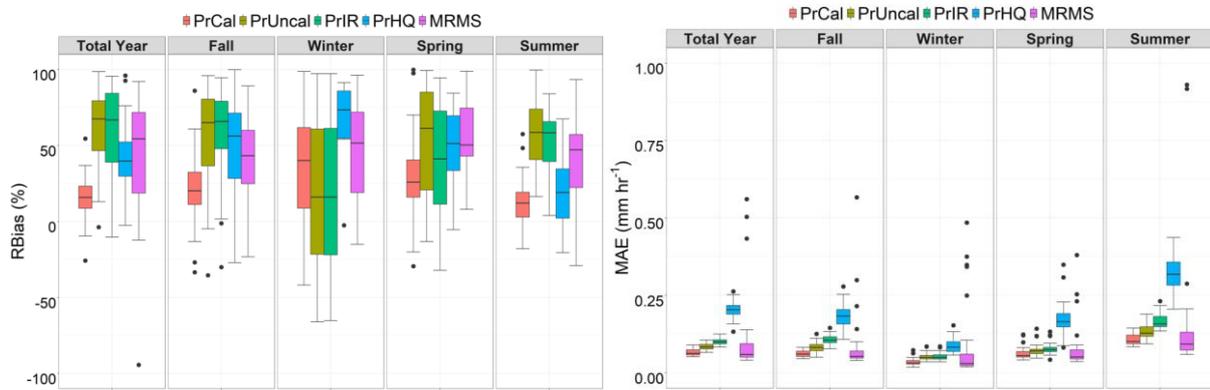
(**Zone-4**)
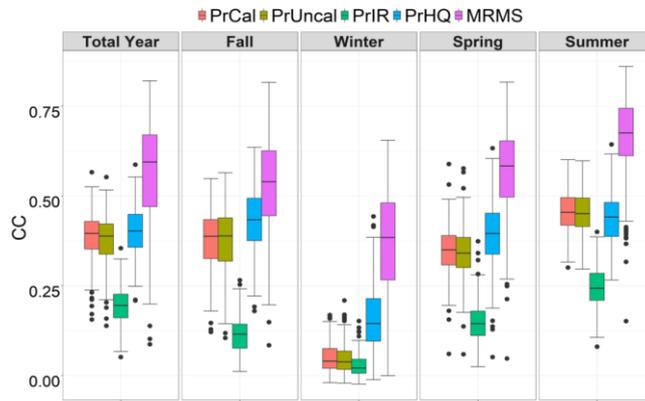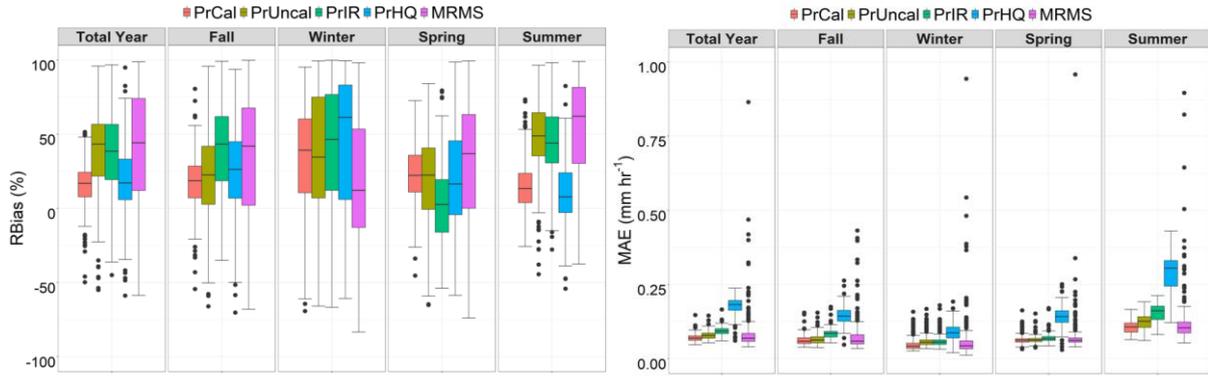
(**Zone-5**)
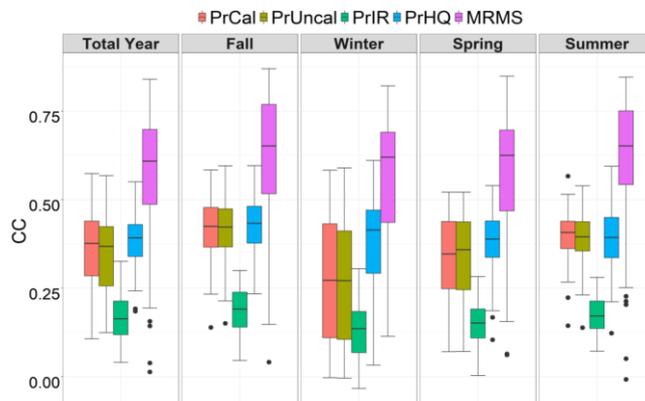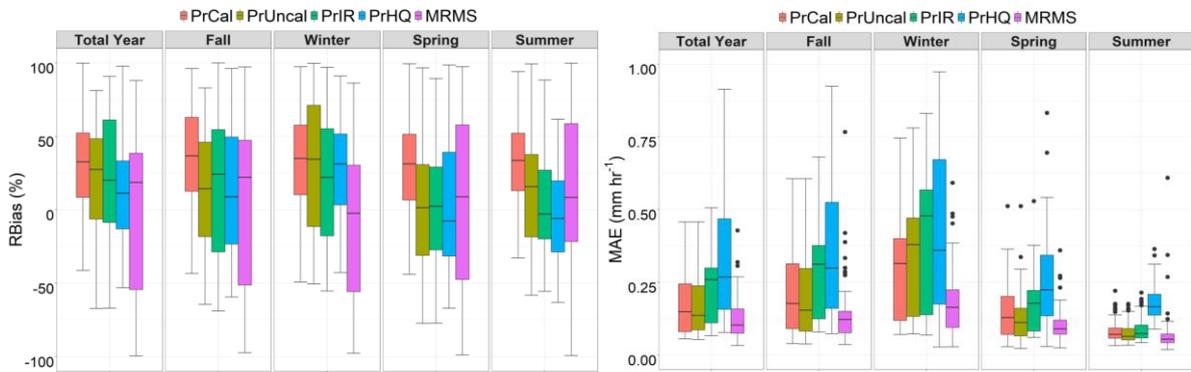
(**Zone-6**)
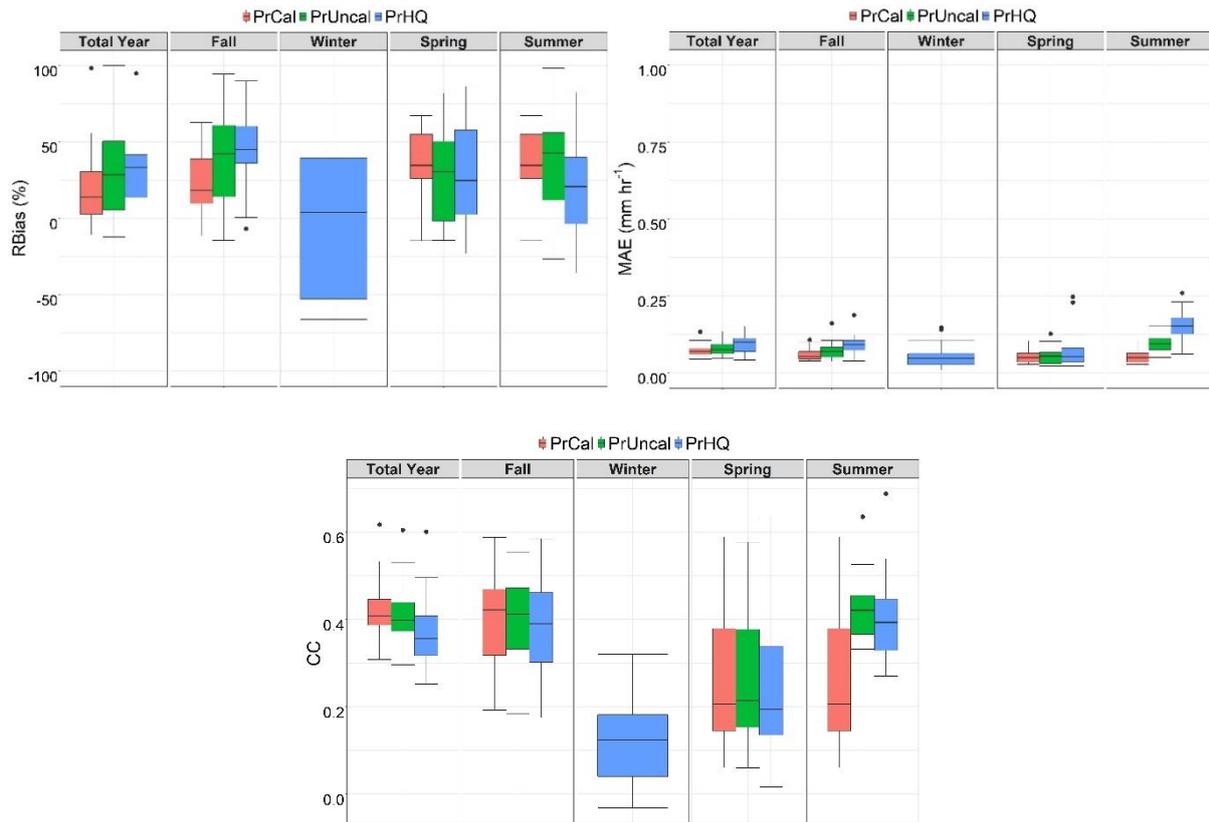
(Zone-7)

Figure 10. Performance evaluation of IMERG V06 (four products) and MRMS at hourly timescale based on Rbias (%), MAE (mm h^{-1}), and CC for different seasons over seven zones

The capability of the IMERG and MRMS products to detect the occurrence of precipitation is further assessed for each zone across the study area for the four seasons. Figure 11 shows boxplots of the three categorical statistics (i.e. POD, FAR, and CSI) corresponding to all products at hourly timescale. Overall, MRMS outperforms other products particularly over eastern and western coasts and during warm periods, with CSI values around 0.5. PrHQ follows MRMS closely as both products are microwave-based and can observe hydrometeor profiles relatively accurately (Huffman et al. 2019b). PrHQ shows the best performance in detecting precipitation events among other IMERG products with the highest values of POD (~ 0.65-0.75), and CSI (~ 0.45-0.55) corresponding to zones-1 and 6, and zones-2 to 5, respectively. It also has the lowest values of FAR among the other products across all zones and all seasons.
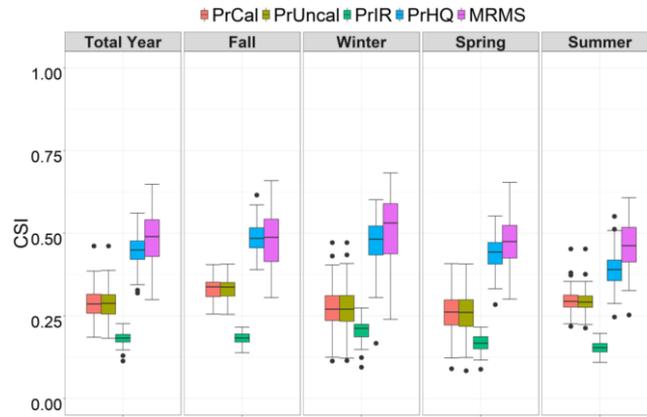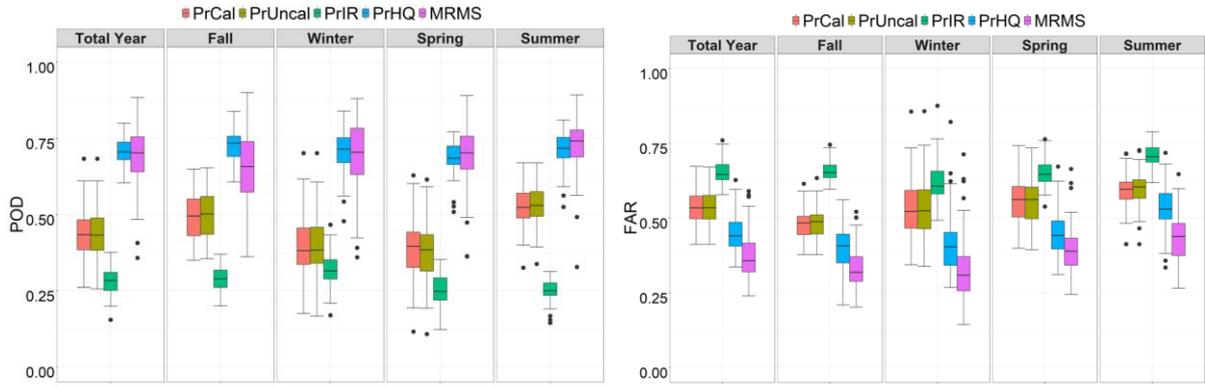
(**Zone-1**)
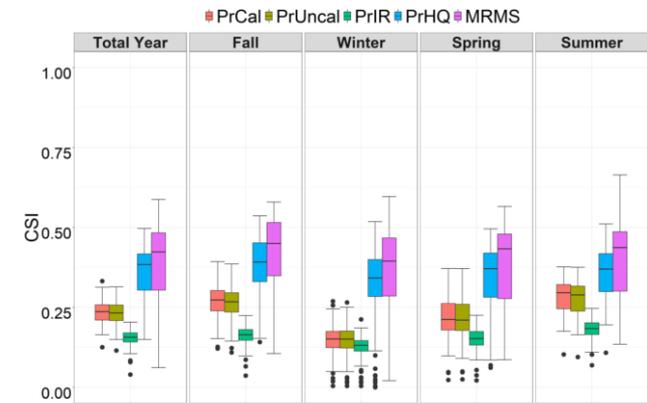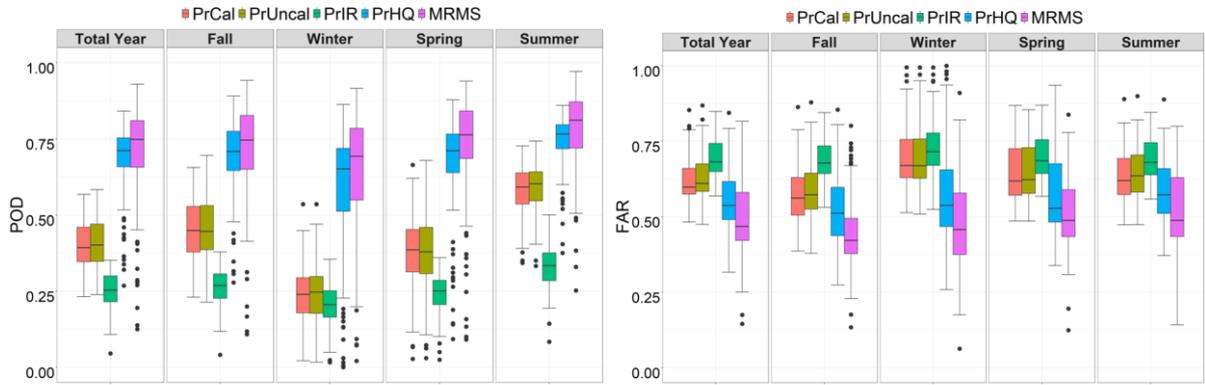
(**Zone-2**)
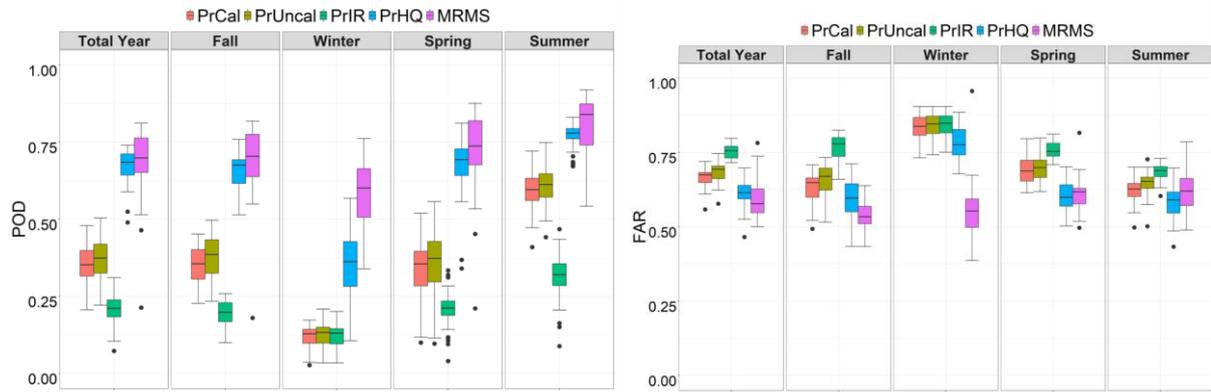
(**Zone-3**)
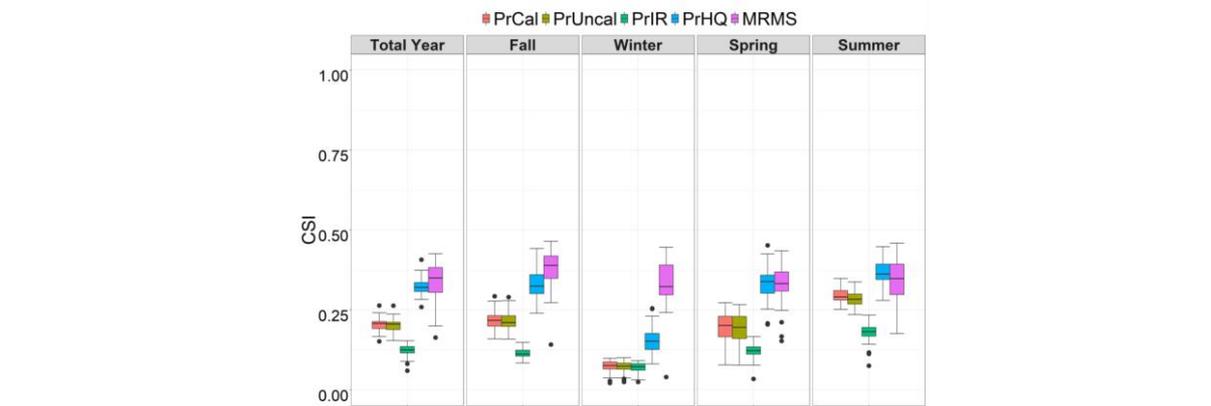
(**Zone-4**)
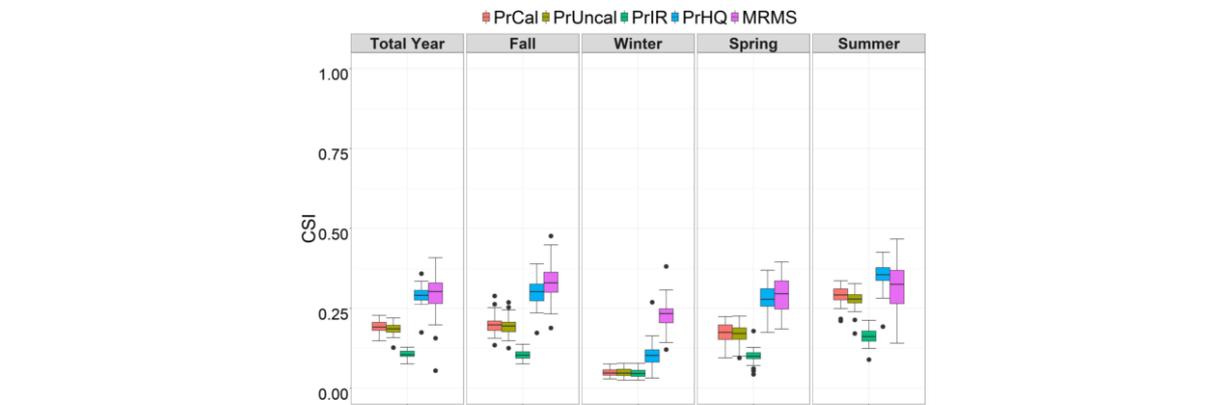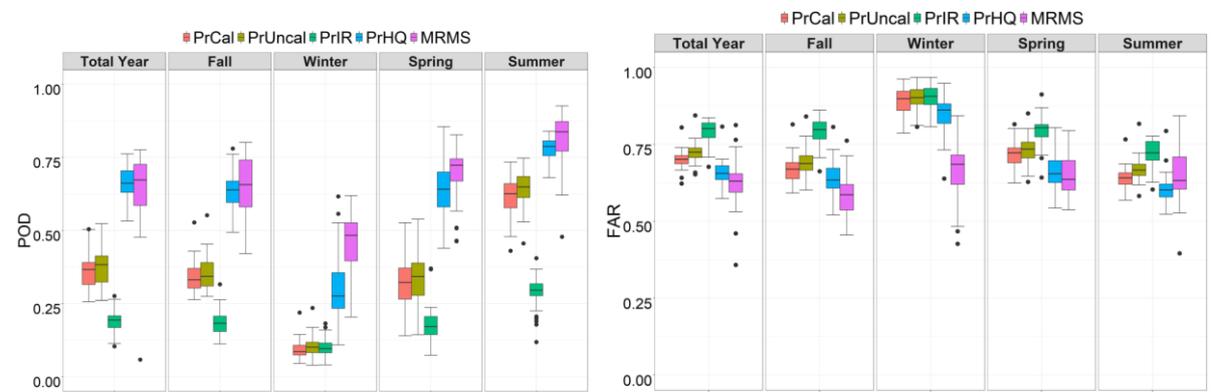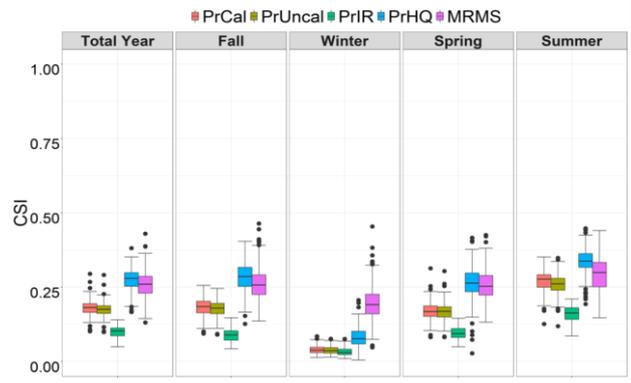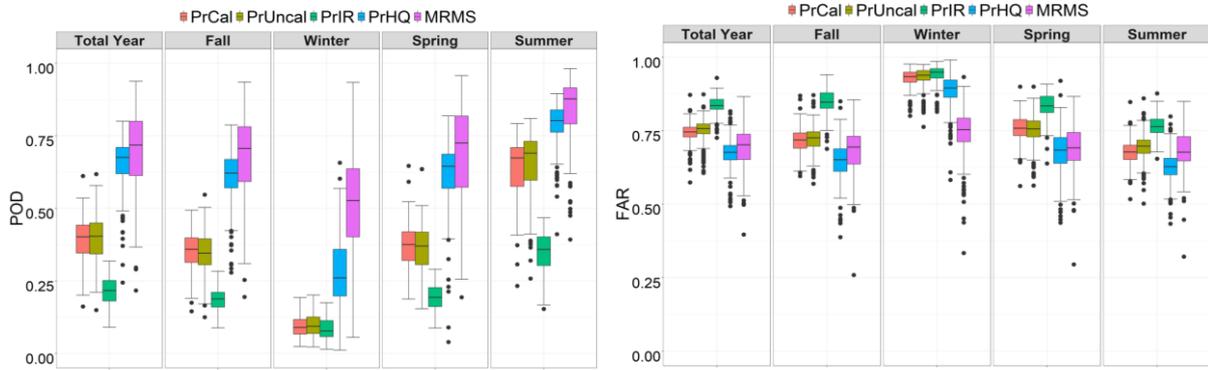
(**Zone-5**)
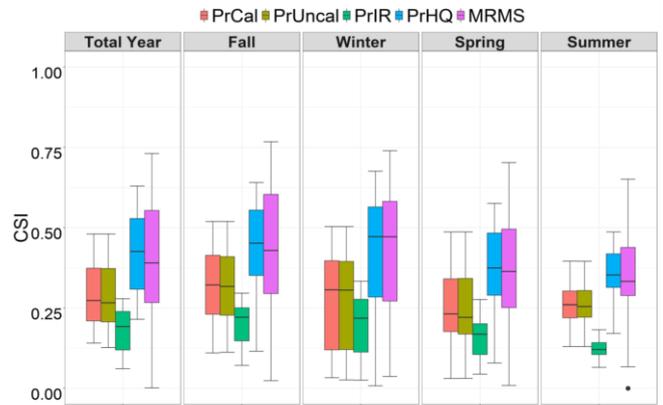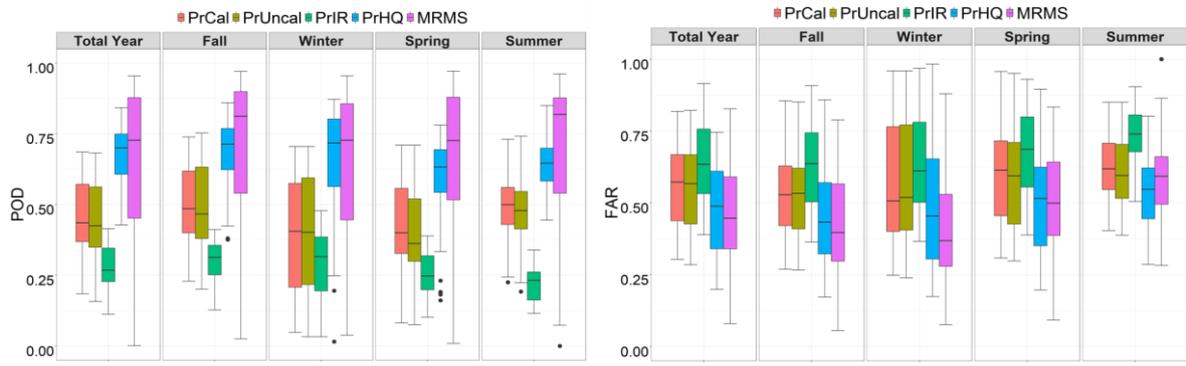
(**Zone-6**)
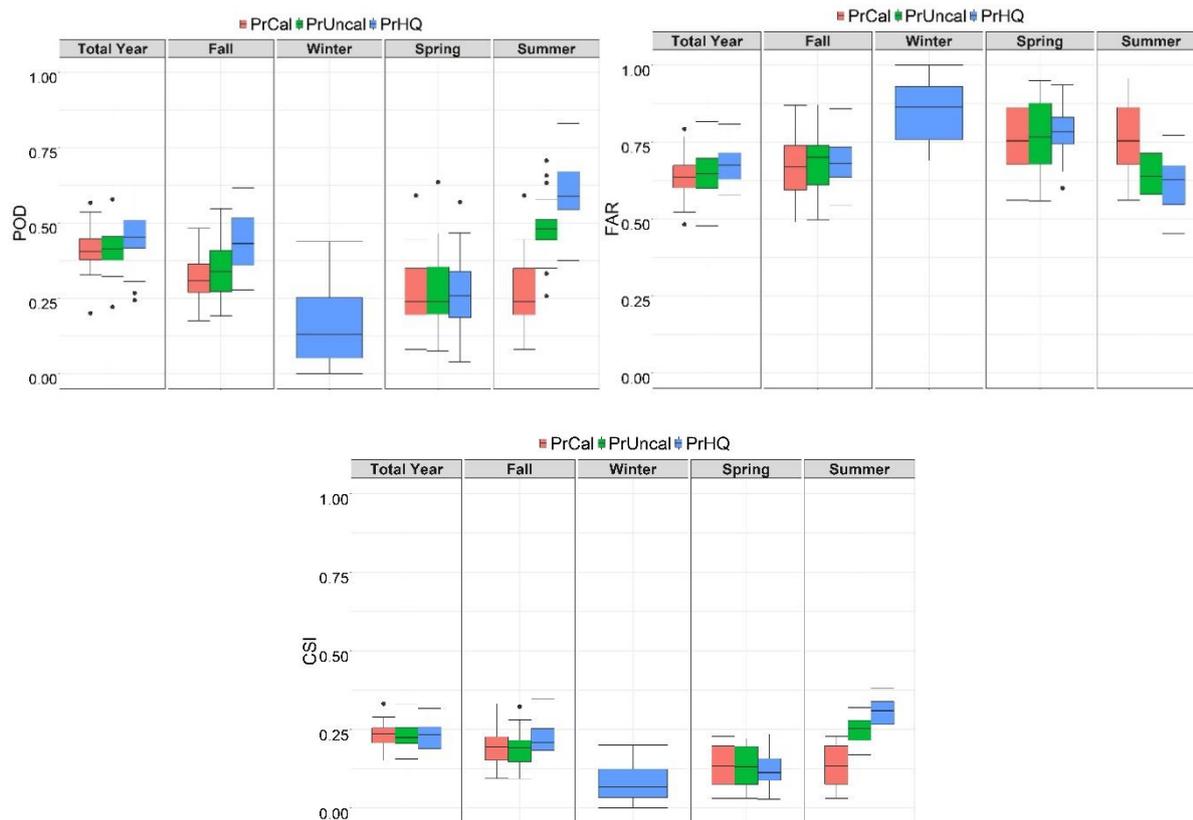
(**Zone-7**)

Figure 11. Evaluation of the IMERG V06 and MRMS products based on POD, FAR, and CSI; the lengths of boxplots represent the range of performance across different sites for seven zones

The hbias, mbias, fbias, and cnbias performance metrics corresponding to the IMERG and MRMS precipitation products are presented in Figure 12. The hit bias, where both satellite and observed data show precipitation values above 0.1mm h^{-1}, ranges between -10% and 10% for the PrCal product. It shows underestimations during summer and overestimations during other months. Analyses show relatively large false and miss biases for all products and during all seasons. The false (and miss) biases for PrCal are 50% (-40%) during all seasons of the study period, which reaches up to 170% (-80%) during winter. Overall, among the IMERG products, PrHQ shows the lowest false and miss biases (except for winter), and PrCal has the best performance based on the hit bias. Additionally, these bias metrics are relatively lower in the fall compared to the other seasons. MRMS shows lower miss and hit biases compared to the IMERG products.

49

The PrIR and MRMS estimates are not available for regions above 60°N and 55°N respectively (including the gauges in zone-7). Therefore these products are not shown in Figure S1 and S2 for zone-7. Note that the PrCal and PrUncal products of Version 06 IMERG precipitation estimates mask out observed passive microwave estimates over snowy/icy surfaces, so outside the latitude band 60°N-S, where IR estimates are not available, precipitation estimates over non-snowy/icy surfaces are recorded as missing. On the other hand, the merged microwave estimate ("HQ") field of Version 06 IMERG precipitation estimates has values across all swaths at the higher latitudes , so PrHQ product is available outside the latitude band 60°N-S (Huffman et al. 2019b). In addition, only the PrHQ shows estimation during winter.
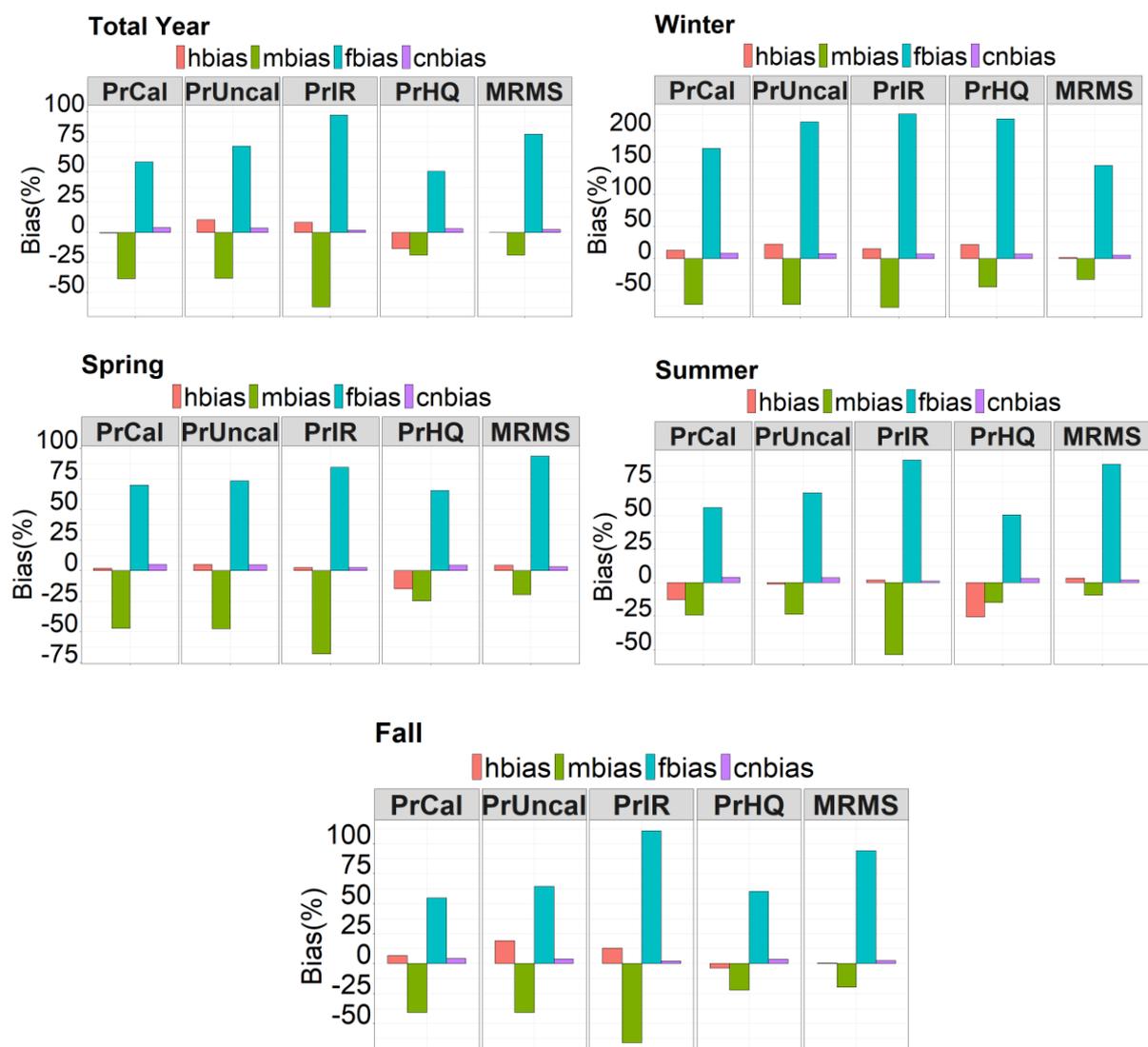
Figure 12. Performance of the hourly IMERG and MRMS products based on hbias, mbias, fbias, and cnbias over 2015-2018

Figure 13 represents the percentage of hit, miss, false, and correct negative events for both IMERG and MRMS. All products show almost similar performance particularly PrCal and PrUncal, except for PrHQ, which has a relatively large false fraction value indicating that it tends to overestimate precipitation. Around 90% (86.7% correct negative and 3.2% hit fractions) of all events are accurately detected by PrCal with only ~10% error (5.5% false and 4.6% miss fractions). MRMS shows lower miss fractions compared to the ones corresponding to the IMERG products and has a larger hit fraction than PrCal, PrUncal, and PrIR. The relatively larger false fraction value (7.1%) suggests that MRMS tends to overestimate precipitation.
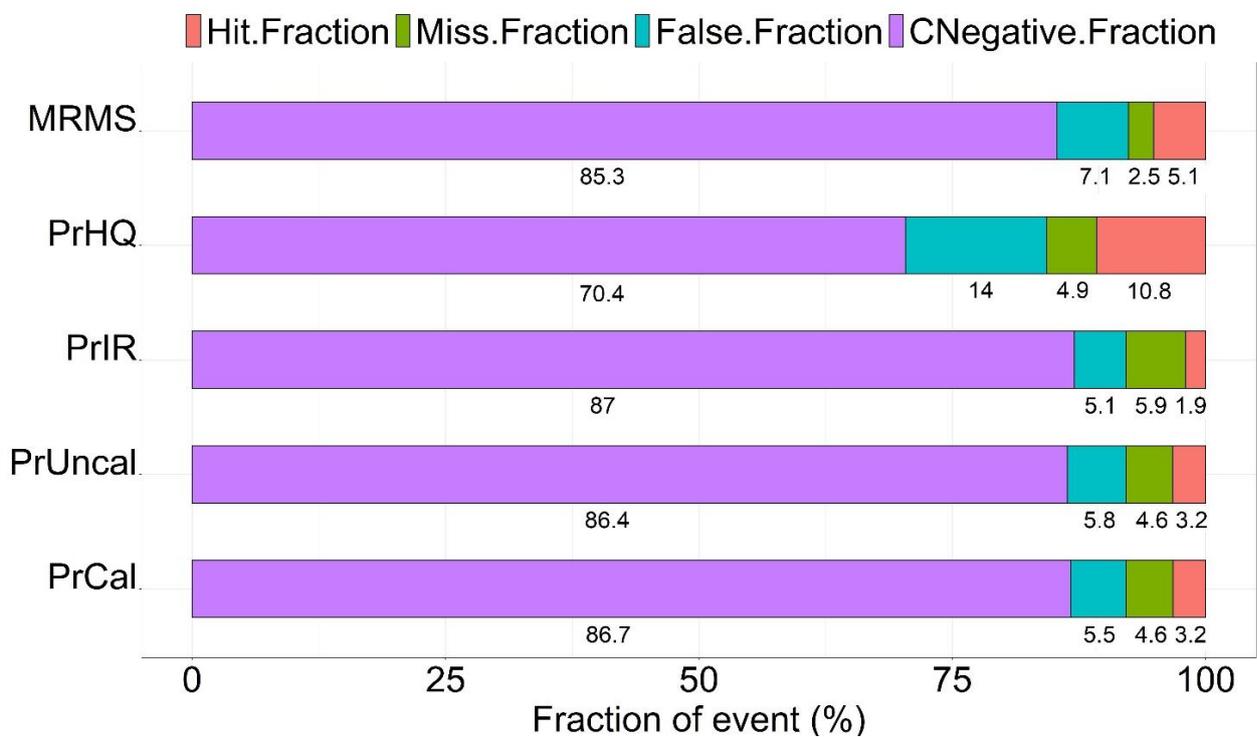
Figure 13. Fractions of hourly hit, miss, false and correct negative events for the IMERG and MRMS products averaged over the study area

Further, we compare the spatially averaged hourly precipitation estimates from IMERG and MRMS with those of the ground-based records across Canada. The corresponding density color scatterplots over the five-year period (2014-2018) for the IMERG products and four-year period (2015-2018) for MRMS are shown in Figure 14. PrCal and PrUncal outperform the other IMERG products but show slight overestimations, while PrHQ has the worst performance and significantly overestimates precipitation. The Q-Q plots of the hourly quantiles averaged over the study sites across the country (Figure 15) indicate that IMERG products, except for

PrHQ which shows significant overestimations at almost all quantiles, tend to slightly overestimate intense precipitation events (i.e. values corresponding to higher quantiles) and underestimate light precipitation (lower quantiles), which is in agreement with the findings of Sunilkumar et al. (2019). MRMS follows the straight line in high quantiles, while it overestimates low and middle quantiles.
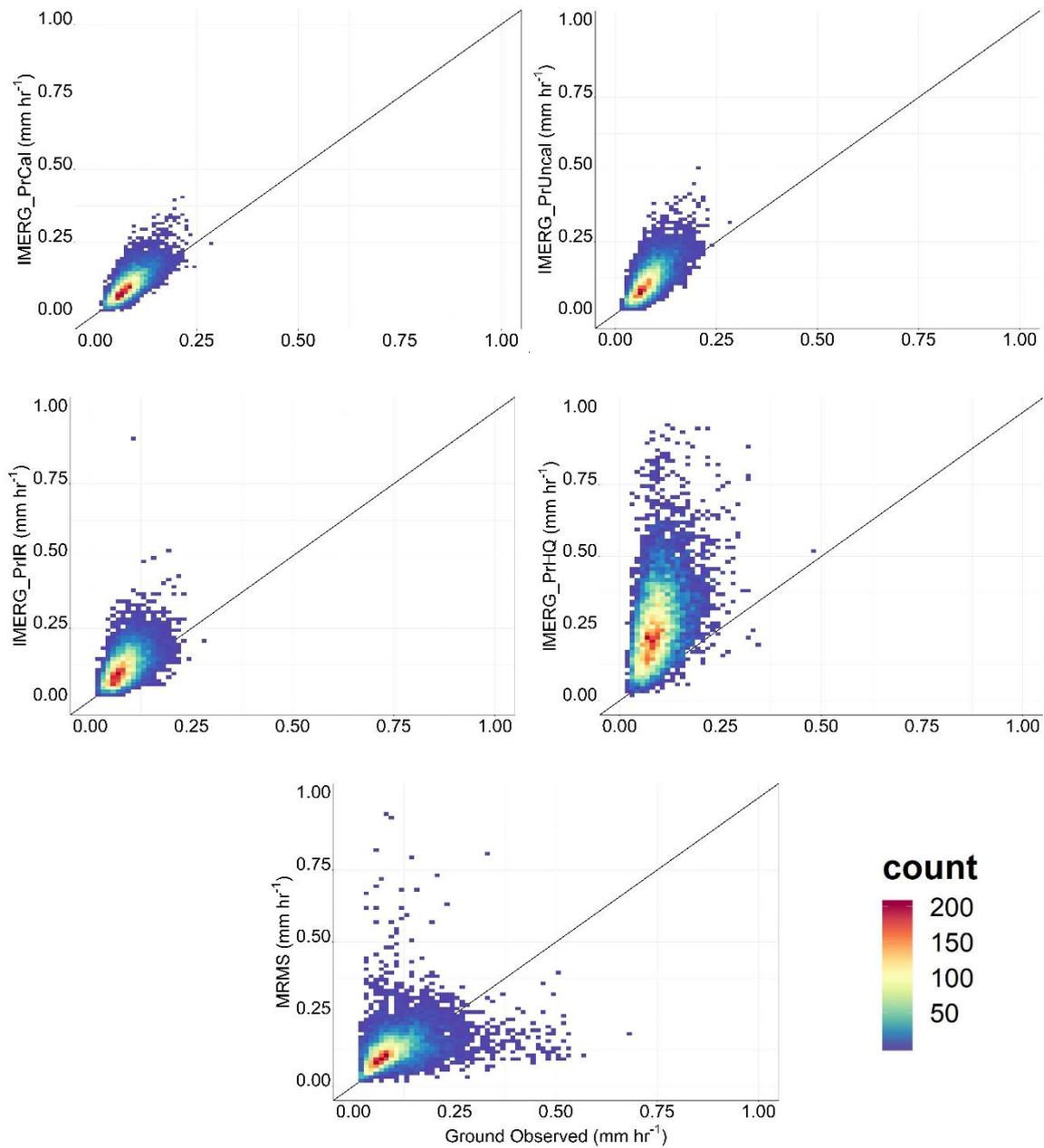
Figure 14. Scatterplots of the spatially averaged true (ground-based) precipitation and IMERG and MRMS precipitation products (mm h^{-1}) at hourly timescale (the blue line is slope one line with an intercept equal to zero)
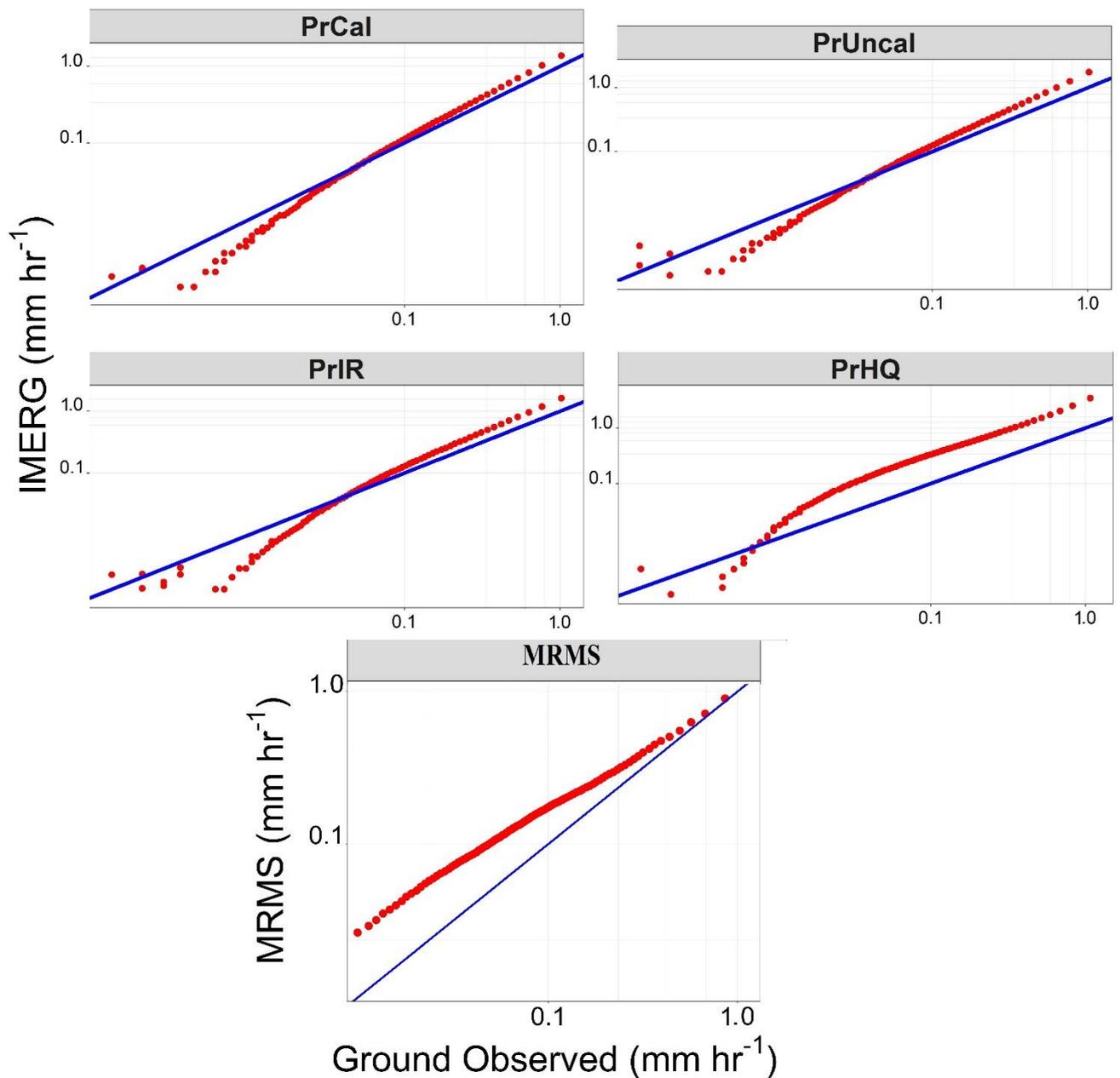
Figure 15. Comparison between the spatially averaged quantiles of the true precipitation and IMERG and MRMS precipitation products (mm h^{-1}) (x and y axes are in log space, (the blue line is slope one line with an intercept equal to zero)

Figures 16 and 17 depict the performance of IMERG at hourly and daily and MRMS ar hourly time scale in capturing the time series pattern against ground observed data. These time series are created based on the spatially averaged value of precipitation over the study area.
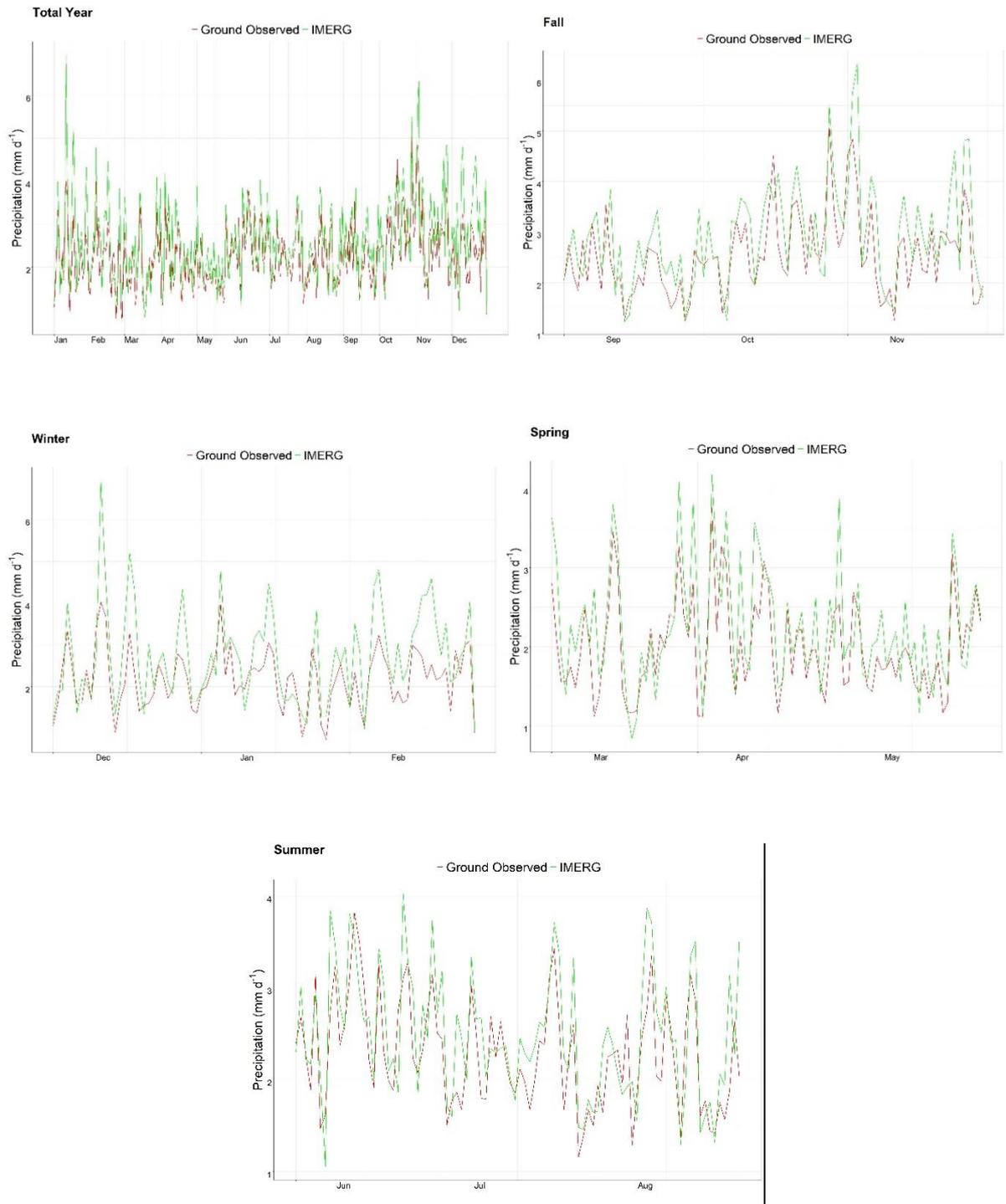
Figure 16. Comparison between spatially averaged IMERG PrCal product and gauge data records at daily timescale
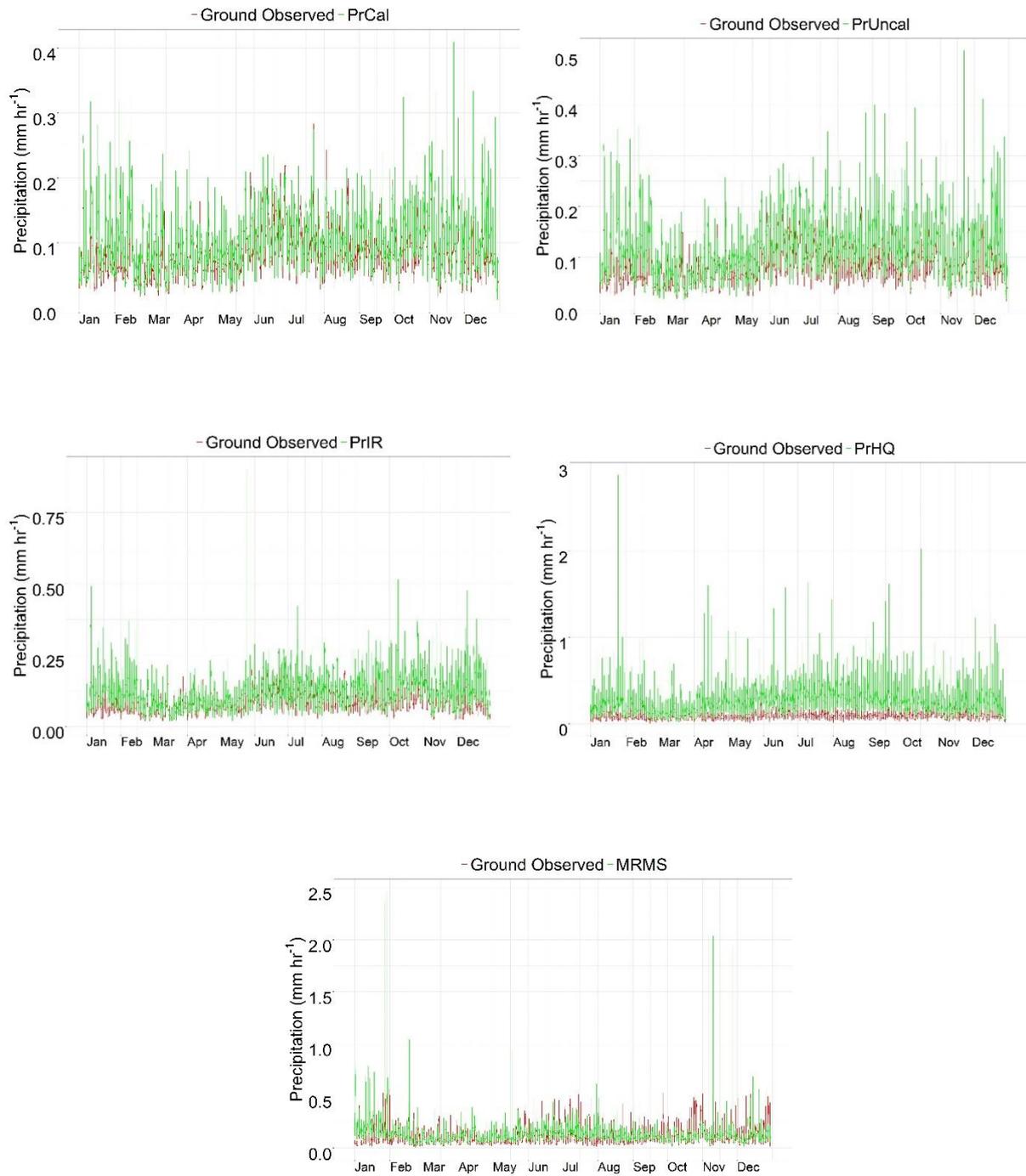
Figure 17. Comparison between spatially averaged IMERG and MRMS products and gauge data records at hourly timescale

3.6. Discussion

We perform a comprehensive analysis of IMERG and MRMS products at hourly and daily time scales over Canada. Overall, the evaluation results show promising performance of these remotely sensed data in representing local precipitation at high spatial and temporal resolution. The PrCal product, which provides a combination of both PMW and IR estimates represent the best regional performance among available IMERG products across Canada with an average Rbias value of 20%, MAE value of 0.15mm h^{-1}, and correlation of 0.45 over 2014-2018. In contrast, the two products of PrIR and PrHQ are relatively less reliable as the former shows low correlation coefficients over the country and the latter indicates high values of RMSE.

The findings of this research regarding IMERG performance agree with the other studies in terms of overall overestimations, better detection capability over plains, and less uncertainty during warm months. Tan et al. (2019) provided the first analysis of IMERG V06 and showed its improved performance in depicting the diurnal cycle of precipitation around the world compared to the previous version (IMERG V05). Their evaluations against the U.S. ground-based observations showed extensive agreements in capturing summertime diurnal peak of precipitation in the central United States by IMERG. Evaluating the IMERG hourly precipitation product against hourly ground-based observations showed slight overestimations over Mainland China in a study conducted by Tang et al. (2017). Also, comparisons between IMERG V03, V04 and V05 Final run products over the globe at $0.1° \times 0.1°$ spatial and daily temporal resolution conducted by Wang et al. (2018) indicated that all IMERG versions tend to overestimate precipitation by about 12%.

As depicted in Figure 8, RMSE values are lower over central Canada versus the coastal regions, which are characterized by heavy precipitation. This can be attributed to the precipitation estimations from IMERG being influenced by topographic conditions, sea and land locations (Xu et al. 2019). Further, our analyses showed that while in several instances PrHQ slightly outperforms PrCal in maintaining the correlation between QPEs and gauge data (Figures 10, and S1), the PrIR product, which estimates precipitation from empirical cloud top temperature and rainfall relationships performs worst. This is because the infrared wave-based sensor does not capture microwave brightness temperatures of hydrometeor profiles and hence cannot detect the precipitation droplets, accurately. In addition, the correlation values significantly

drop in winter because of changes in the precipitation regime (i.e. snowfall) over zones-1 to 5. The Pacific region (zone-6) shows fewer changes in the correlation values between seasons (with larger variations in winter).

The performance of the IMERG products in representing the occurrence of precipitation is higher over the east and west coasts, based on CSI values. Nonetheless, the uncertainties are relatively high in zone-6 indicating larger variations across this area because of its diverse topographic and climatic conditions. Satellite sensors have difficulties in detecting low-level orographic rainfall events that often occur at elevations higher than 3000m elevations (Chen et al. 2019), such as parts of the Rocky Mountains in the west of Canada (zone-6). The PMW-based data (PrHQ) that are derived based on sensors with higher frequency range display more stable performance at high elevations as expected (Chen et al. 2019).

The performance of all products is mostly consistent during fall, spring, and summer, with weaker performance in winter. One source of uncertainty in winter precipitation estimation is the lack of reliable ground-based precipitation observations during cold seasons. As mentioned in the data section, precipitation gauges across most parts of Canada show poor performance in solid precipitation measurement. In addition, previous studies have shown problems associated with satellite snowfall estimates because of PMW sensors used in satellite products. Chen et al. (2019) argued that the PWM retrieval, which is in contact with the precipitation particles, has problems in distinguishing between precipitation and frozen surface. Further, the IR input that utilizes the morphing technique is directly inferred from cloud top temperature and is less affected by the impact of seasonal variation on retrieval results. Huffman et al. 2019 stated that, all merged PMW estimates have low accuracies in regions with frozen or icy surfaces. Thus, PrHQ has relatively low/high uncertainties in summer/winter, while IR input looks more stable and smoother across time. Currently, IMERG classifies rainfall and snowfall using wet-bulb temperature with a uniform temperature threshold, over the globe. It is necessary to acquire a spatially distributed map of temperature thresholds for more precise rainfall and snowfall separation (Tang et al. 2020).

The MRMS data performs relatively well across southern Canada but with uncertainties associated with false and miss estimates. This is partly because precipitation features are

detected aloft and evaporating before reaching the surface. Furthermore, while the MRMS domain covers southern Canada to 55°N latitude, the 30 Canadian radars ingested by MRMS do not cover the entire domain. This could potentially influence some of the statistics used in these analyses. In addition, the results of MRMS in this study indicate its low performance during winter which is in agreement with the findings of Cocks et al. (2016). They assessed the performance of MRMS radar-only QPE for the cold season over the United States and found that it has a distinct negatively biased QPE during the cold season. They argued that during the cool season radar beam overshoot is more common because of shallower precipitation systems and lower cloud bases.

3.7. Summary and Conclusion

This study evaluates the most recent satellite and radar (i.e. IMERG-V06 and MRMS) precipitation estimates using ground-based observations across Canada. A suite of performance metrics is used to assess various characteristics of the RSE products seasonally at daily and hourly timescales over 2014-2018 for IMERG and 2015-2018 for MRMS.

The hourly evaluations suggest that PrCal outperforms other IMERG products in estimating the precipitation amount. Although PrHQ (the microwave-based product) shows large biases in the intensity, it detects the occurrence of precipitation more accurately. The robustness of such sensors to detect precipitation suggests that they can be used to improve PrCal estimates. As expected, IMERG performs better at daily timescale compared to the hourly based on all metrics. For example, the median Rbias and CSI of daily PrCal are 13% and 52% respectively, which reduce to 18% and 25% at hourly scale.

Overall, IMERG better represents the ground-based precipitation amounts over most parts of the interior plains compared to the rest of the country with lower Rbias and RMSE values. Nonetheless, higher POD and lower FAR values indicate that precipitation occurrence is best captured over the west and east coasts. Based on the seasonal assessments, IMERG provides more reliable precipitation estimates during warm months especially in summer according to correlation coefficients and categorical indices. This is in agreement with the findings of Asong et al. 2017.

In addition, although IMERG is capable of representing the temporal and spatial variations of precipitation over most parts of Canada, it tends to overestimate the moderate to heavy precipitation events and shows relatively weak performance during the cold season.

Assessments of MRMS, as the first study of such high-resolution radar-based precipitation estimates across Canada, show the overall satisfactory performance of this product throughout its coverage area in the southern parts of the country. In addition to maintaining the spatial variations of precipitation in accordance with ground observations, MRMS exhibits a higher average CC value (~ 0.6) than PrCal from IMERG (~ 0.4) and better CSI values over all regions. MRMS, however, tends to underestimate precipitation in the eastern and western parts of Canada and overestimates it in the interior plains. Similar to PrHQ, MRMS can detect precipitation occurrence relatively well as they are both microwave-based products. Further, MRMS has better coverage over the southern parts of Canada with higher spatial and temporal resolution than PrHQ, which has several gaps.

The first comprehensive analysis of the most recent QPE products across the entire ground network stations of Canada suggests that both IMERG and MRMS have considerable capabilities in representing precipitation estimates particularly in the interior and over the east and west coasts, respectively. However, there are systematic and random biases and uncertainties associated with both products that should be adjusted before driving hydrological models or performing risk analyses, among others. A combination of both products can result in more accurate estimations especially for short duration events and in areas that have sparse rain gauges.

3.8. References

Anjum, M.N., Ahmad, I., Ding, Y., Shangguan, D., Zaman, M., Ijaz, M.W., Sarwar, K., Han, H., Yang, M., 2019. Assessment of IMERG-V06 Precipitation Product overDifferent Hydro-Climatic Regimes in the TianshanMountains, North-Western China, Remote Sens. 11, 2314; doi:10.3390/rs11192314.

Chen, H., Yong, B., Gourley, J.J., Liu, j., Ren, L., Wang,W., Honge, Y., Zhang, J., 2019. Impact of the Crucial Geographic and Climatic Factors on the Input Source Errors of GPM based Global Satellite Precipitation Estimates. Journal of Hydrology. 575, Pages 1-16. https://doi.org/10.1016/j.jhydrol.2019.05.020.

Cocks, S., Martinaitis, S., Kaney, B., Zhang, J., Howard, K., 2016. MRMS QPE performance during the 2013/14 cool season. J. Hydrometeor., 17, 791–810, doi:10.1175/JHM-D-15-0095.1.

Cocks, S.B., Zhang, J., Howard, K., 2017. MRMS QPE Performance East of the Rockies during the 2014 Warm Season, JOURNAL OF HYDROMETEOROLOGY, American Meteorological Society, DOI: 10.1175/JHM-D-16-0179.1.

Gebregiorgis, A. S., Kirstetter, P.-E., Hong, Y. E., Gourley, J. J., Huffman, G. J., Petersen,W. A., Xue, X., Schwaller,M. R., 2018. To what extent is the day 1 GPM IMERG satellite precipitation estimate improved as compared to TRMM TMPART? Journal of Geophysical Research: Atmospheres, 123, 1694–1707. https:// doi.org/10.1002/2017JD027606.

Giangrande, S., Ryzhkov, A., 2008. Estimation of rainfall based on the results of polarimetric echo classification. J. Appl. Meteor., 47, 2445–2462, doi:10.1175/2008JAMC1753.1.

Grams, H.M., Zhang, J., Elmore, K.L., 2014. Automated Identification of Enhanced Rainfall Rates Using the Near-Storm Environment for Radar Precipitation Estimates. Journal of Hydrometeorology.

Gumindoga, W., Rientjes, T. H. M., Haile, A. T., Makurira, H., Reggiani, P., 2019. Performance evaluation of CMORPH satellite precipitation product in the Zambezi Basin, Int. J. Remote Sens., 40, 1–20, https://doi.org/10.1080/01431161.2019.1602791.

Habib, E., Henschke, A., Adler, R., 2009. Evaluation of TMPA Satellite-Based Research and Real-Time Rainfall Estimates During Six Tropical-Related Heavy Rainfall Events over Louisiana, USA. Atmospheric Research 94: 373–388. doi:10.1016/j.atmosres.2009.06.015.

Hossain F., Huffman, G.J., 2008. Investigating Error Metrics for Satellite Rainfall Data at Hydrologically Relevant Scales, Journal of Hydrometeorology.

Hou, A.Y., Kakar, R.K., Neeck, S., Azarbarzin, A.A., Kummerow, C.D., Kojima, M., Oki, R., Nakamura, K., Lguchi, T., 2014. The global precipitation measurement mission. Bull. Am. Meteorol. Soc. 95 (5), 701–722.

Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E.J., Sorooshian, S., Tan, J., Xie, P., 2019a. Algorithm Theoretical Basis Document (ATBD) Version 06, NASA Global Precipitation Measurement (GPM) Integrated Multi satellitE Retrievals for GPM (IMERG), National Aeronautics and Space Administration (NASA).

Huffman, G.J., Bolvin, D.T., Nelkin, E.J., Tan, J., 2019b. Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation, IMERG Tech Document, National Aeronautics and Space Administration (NASA).

Kubota, T., Shige, S., Hashizume, H., Aonashi, K., Takahashi, N., Seto, S., Hirose, M., Takayabu, Y. N., Nakagawa, K., Iwanami, K., Ushio, T., Kachi, M., Okamoto, K., 2007. Global precipitation map using satelliteborne microwave radiometers by the GSMaP project: Production and valida-tion, IEEE Trans. Geosci. Remote Sens., 45, 2259–2275.

Kubota, T., Ushion, T., Shing, S., Kida, S., Kachi, M., Okamoto, K., 2009. Verification of high-resolution satellite-based rainfall estimates around Japan using a gauge-calibrated ground-radar dataset. J. Meteorol. Soc. Jpn. 87A (3), 203–222.

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

Lin, Y., Mitchell, K. E., 2005. The NCEP stage II/IV hourly precipitation analyses: Development and applications. 19th Conf. on Hydrology, San Diego, CA, Amer. Meteor. Soc., 1.2. [Available online at http://ams.confex.com/ams/pdfpapers/83847.pdf.]

Liu, Z., 2016. Comparison of Integrated Multisatellite Retrievals for GPM (IMERG) and TRMM Multisatellite Precipitation Analysis (TMPA) monthly precipitation products: Initial results, J. Hydrometeorol., 17, 777–790, https://doi.org/10.1175/JHM-D-15- 0068.1.

Martinaitis, S.M., Gourley, J.J., Flamig, Z.L., Argyle, E.M., Clark, R.A., Arthur, A., Smith, B.R., Erlingis, J.M., Perfater, S., Albright, B., 2017. The HMT Multi-Radar Multi-Sensor Hydro Experiment, BAMS, American Meteorological Society.

Martinaitis, S.M., Cocks, S.B., Qi, Y., Kaney, B.T., 2015. Understanding Winter Precipitation Impacts on Automated Gauge Observations within a Real-Time System, Journal of OF Hydrometeorology, DOI: 10.1175/JHM-D-15-0020.1

Mei, Y., Nikolopoulos, E.I., Anagnostou, E.N., Zoccatelli, D., Borga, M., 2016. Error Analysis of Satellite Precipitation-Driven Modeling of Flood Events in Complex Alpine Terrain. Remote Sens.2016,8, 293; doi:10.3390/rs8040293.

Mekis, E., Donaldson, N., Reid, J., Zucconi, A., Hoover, J., Li, Q., Nitu R., Melo S., 2018. An Overview of Surface-Based Precipitation Observations at Environment and Climate Change Canada, Atmosphere-Ocean, 56:2, 71-95, DOI:10.1080/07055900.2018.1433627.

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. doi:10.1016/j. atmosres.2013.08.016.

Moazami, S., Golian, S., Hong, Y., Chen S., Kavianpour, M.R., 2016. Comprehensive evaluation of four high-resolution satellite precipitation products under diverse climate conditions in Iran. Hydrological Sciences Journal, DOI: 10.1080/02626667.2014.987675.

Ochoa- Rodriguez, S., Wang, L.P., Willems, P., Onof, C., 2019. A review of radar- rain gauge data merging methods and their potential for urban hydrological applications. Water Resources Research, 55, 6356–6391. https://doi.org/10.1029/2018WR023332.

Petersen, W.A., Christian, H.J., Rutledge, S.A., 2005. TRMM observations of the global relationship between ice water content and lightning. Geophys. Res. Lett. 32, 2471–2494.

Qi, Y., Martinaitis, S., Zhang, J., Cocks, S., 2016. A Real-Time Automated Quality Control of Hourly Rain Gauge Data Based on Multiple Sensors in MRMS System, J. Hydrometeor, DOI: 10.1175/JHM-D-15-0188.1, American Meteorological Society.

Rasmussen, R., Baker, B., Kochendorfer, J., Meyers, T., Landolt, S., Fischer, A. P.,…Smith, C. (2012). How well are we measuring snow: The NOAA/ FAA/NCAR winter precipitation test bed. Bulletin of the American Meteorological Society, 93(6), 811–829.

Singh H., Pirani F.J., Najafi M.R., 2020. Characterizing the temperature and precipitation covariability over Canada. Theoretical and Applied Climatology,139(3):1543-58.

Stampoulis, D., Anagnostou, E.N., 2012, Evaluation of Global Satellite Rainfall Products over Continental Europe. Journal of Hydrometeorology. DOI: 10.1175/JHM-D-11-086.1

Sunilkumar, K., Akiyo Yatagai, A., Masuda, M. 2019, Preliminary Evaluation of GPM IMERG Rainfall Estimates Over Three Distinct Climate Zones With APHRODITE, Earth and Space Science, 6, 1321–1335. https://doi.org/10.1029/2018EA000503.

Tan, J., Huffman, G. J., Bolvin, D. T., Nelkin, E. J., 2019. Diurnal cycle of IMERG V06 precipitation. Geophysical Research Letters, 46, 13,584–13,592. https://doi.org/10.1029/2019GL085395.

Tang, G., Ma, Y., Long, D., Zhong, L., Hong, Y., 2016. Evaluation of GPM Day-1 IMERG and TMPA Version 7 legacy products over Mainland China at multiple spatiotemporal scales, J. Hydrol., 533, 152–167, https://doi.org/10.1016/j.jhydrol.2015.12.008.

Tang, G., Clark, M.P., Papalexiou, S.M., Ma, Z., Hong Y., 2020. Have satellite precipitation products improved over last two decades? A comprehensive comparison of GPM IMERG with nine satellite and reanalysis datasets, Remote Sensing of Environment, Volume 240, 111697, https://doi.org/10.1016/j.rse.2020.111697.

Tian, Y., Lidard-Peters, C. D., Choudhury, B. J., Garcia, M., 2007. Multitemporal Analysis of TRMM-Based Satellite Precipitation Products for Land Data Assimilation Applications. Journal of Hydrometeorology 8: 1165–1183. doi:10.1175/2007JHM859.1.

Villarini, G., Krajewski, W. F., 2007. Evaluation of the research version TMPA three-hourly 0.25_ _0.25_ rainfall estimates over Oklahoma. Geophys. Res. Lett., 34, L05402, https://doi.org/10.1029/2006GL029147.

Villarini, G., Mandapaka, P.V., Krajewski, W.F., Moore, R.J., 2008. Rainfall and sampling uncertainties: A rain gauge perspective. Journal of Geophysical Research. VOL. 113, D11102. doi:10.1029/2007JD009214.

Wang J., Wolff, D.B., 2012. Evaluation of TRMM Rain Estimates Using Ground Measurements over Central Florida, Journal of Applied Meteorology and Climatology. doi:10.1175/JAMC-D-11-080.1. American Meteorological Society.

Wang, C., Tang, G., Han, Z., Guo, X., Hong, Y., 2018. Global intercomparison and regional evaluation of GPM IMERG Version-03, Version-04 and its latest Version-05 precipitation products: Similarity, difference and improvements, Journal of Hydrology.

Xu, F., Guo, B., Ye, B., Ye, Q., Chen, H., Ju, X., Guo, J., Wang, Z., 2019. Systematical Evaluation of GPM IMERG and TRMM 3B42V7 Precipitation Products in the Huang-Huai-Hai Plain, China, Remote Sens.,11, 697; doi:10.3390/rs11060697

Yong, B., Liu, D., Gourley, J. J., Tian, Y., Huffman, G. J., Ren, L., and Hong, Y., 2015. Global view of real-time Trmm Multisatellite Precipitation Analysis: Implications for its successor global precipitation measurement mission, B. Am. Meteorol. Soc., 96, 283– 296, https://doi.org/10.1175/BAMS-D-14-00017.1.

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., Grams, H., Wang, Y., CoCKs, S., Martinaitis, S., Arthur, A., Cooper, K., Brogden, J., Kitzmiller D., 2016. Multi-Radar Multi-Sensor (MRMS) Quantitative Precipitation Estimation: Initial Operating Capabilities", Bulletin of the American Meteorological Society. DOI:10.1175/BAMS-D-14-00174.1.

Chapter 4

4. Bias correction of satellite IMERG data

4.1. Introduction

As mentioned in the previous chapters, although satellite-based precipitation products provide reasonably well estimates at high spatiotemporal resolution, the measurements indicate systematic biases due to inherent uncertainty related to remotely sensed data. It is noted that two types of errors are associated with SPPs named random and systematic errors (AghaKouchak et al., 2012). Systematic errors that are consistent and predictable come from the sensors and devices that satellite algorithms use in precipitation estimates, however, random errors as unpredictable errors are related to the spatial and temporal sampling. Previous studies have attempted to evaluate and correct the systematic biases by comparing the SPPs with the most reliable ground-based (gauge and radar) precipitation observations. For correcting this type of error, at first, understanding the error characteristics and its variation over different regions are important which was performed in Chapter 3 of this study. Therefore, in this chapter, a method to adjust the bias of IMERG PrCal product as the most applicable and reliable data at daily temporal and $0.1°$ spatial scales is developed over the entire study area including gauged and ungauged locations. This method is based on a variety of processes that are explained in detail in the methodology section of this chapter.

In recent decades the use of satellite precipitation estimates has increased dramatically over different parts of the globe, while studies regarding the uncertainty analysis and bias correction of these useful products have conducted. As the satellite data are grid-based and in contrast, the ground-based true data are provided in point scales, comparison between two datasets and consequently bias correction of satellite areal data have some challenges (Tang et al., 2018). A simple approach is correcting the spatially averaged bias over the interested region (Seo et al., 1999). However, these methods may not be suitable for large-scale basins (> 5000 km2) where rainfall varies in space significantly (Habib et al., 2014).

Some other bias correction techniques consider spatially distributed patterns in bias and attempt to adjust them (Yang et al., 2016; Lu et al., 2019). These methods use different regression-

based techniques such as stepwise regression (STEP) and geographically weighted regression (GWR) to find a relation between satellite precipitation and other covariates, as well as interpolation techniques such as inverse distance weighting, nearest neighbor, and kriging to correct the biases, especially at ungauged pixels. Müller and Thompson (2013) applied such a method in Nepal and concluded that space-variant correction schemes are effective in reducing the bias for CMORPH and TRMM. However, the limitation of this method is that they can not capture the time variation of true data and adjust the biases in timing of the events (Ajaaj et al, 2016). Furthermore, interpolation techniques lead to additional uncertainties, and the quality of interpolated precipitation in grid pixels without gauges is typically lower than that in grid pixels with gauges (Tang et al., 2018). In the following, more details regarding some of the widely-used bias correction methods with their main advantages and disadvantages are provided.

Mean bias removal technique

In this technique the mean bias is estimated at each year by calculating the difference between observed and estimated rainfall as follows:

$$Bias = \frac{\sum_{i=1}^{N}\left(P_{O_i} - P_{S_i}\right)}{N} \quad (4.1)$$

where P_{O_i} is mean monthly observed precipitation at year i, and P_{S_i} is satellite precipitation at year i, N is the number of years. This 'mean bias' is then applied to uncorrected satellite data at each month for a given year to get the corrected precipitation in that year (Davis, 1976; Kharin and Zwiers, 2002). The main advantage of this method is its simplicity, however, it does not capture the climate pattern and does not remove the bias that is associated with higher precipitation rate (Alharbi, 2019).

Multiplicative shift technique

At first the ratio between observed (P_{O_i}) and satellite estimates (P_{S_i}) for each year is obtained by

$$m = \frac{\sum_{i=1}^{N} P_{O_i}}{\sum_{i=1}^{N} P_{S_i}} \qquad (4.2)$$

The value *m* should be calculated using cross-validation (1 year among the total dataset is withdrawn for 'test' and the remaining dataset is used as the 'training' dataset, which has been used for calculation of all statistics). It is then multiplied by the original estimated value for generating bias-corrected data (Ines and Hansen, 2006; Acharya et al., 2013). It can remove the bias from mean monthly rainfall and it can be used for correcting daily rainfall. In this method biases will remain in rainfall intensity and frequency specifically for dry periods. It also fails to correct year to year variations in rainfall intensities (Ajaaj et al., 2016).

Linear regression (R)

In this technique, the general linear regression equation is used at the monthly time scale, which is given by:

$$R_t = a_0 + a_t \times F_t \quad (4.3)$$

where a_0 and a_t are called the constant and coefficient of the linear regression data and R_t is the corrected estimated satellite precipitation (Hay et al., 2000; Lafon et al., 2013). The coefficients of the linear fit can be evaluated using the least square estimates as follows.

$$a_t = Cov\ (Y,F)/Var(F) \quad (4.4)$$

$$a_0 = Y - a_t \times F \qquad (4.5)$$

Where $Cov\ (Y, F)$ is the covariance for the observed (Y) and the estimated (F) precipitation, respectively. $Var(F)$ is the variance of the estimated (F) data.

This method is simple, but it may not capture inter-monthly variability and possibly affect moments of the probability distribution of daily precipitations (Diaz-Nieto and Wilby 2005).

Nonlinear correction method

The linear bias correction methods work by correcting the bias in the mean without correcting the bias in the variance. Therefore, the non-linear bias correction approach that has an exponential form

$$P_{BC} = a.P^b \quad (4.6)$$

shown by Equation 4.6, is used to correct the variously estimated precipitations by SPPs. In Equation 4.6, a and b are scale factors, and P is the original precipitation rate that is estimated by SPP. P_{BC} is the bias-corrected rate. The parameter b can be estimated iteratively by matching the monthly coefficient of variations of the bias-corrected SPPs with the monthly coefficient of variations of the ground observations. The non-linear bias correction works by matching the mean and variance of ground observations (Alharbi, 2019). This technique is simple and just needs monthly observed statistics, however, biases associated with the higher-order moments may not be removed by this method (Ajaaj et al., 2016).

Stochastic modeling

This method was developed by Müller and Thompson in 2013 to spatially aggregate and interpolate the parameters of gauge data that describe the frequency and intensity of the rainfall observed at the satellite grid resolution. The resulting gridded parameters were then used to adjust the probability density function of satellite rainfall observations at each grid cell, accounting for spatial heterogeneity. The procedure for this method is as follows:

1. Extracting rainfall frequency and intensity parameters. For rainfall frequency parameters the authors used a first-order Markov chain model which characterized the probability of a rainy day conditional on the previous day being dry or rainy. For rainfall intensity parameters, they used a gamma distribution with shape and rate parameters to describe the probability distribution of daily rainfall depths on those rainy days. These Stochastic Model Parameters (SMPs) are directly related to a range of relevant metrics that describe rainfall distribution and can thus be used to evaluate the bias adjustment method.

2. Aggregating the SMPs observed at the gauges to the resolution of satellite pixels.

3. Interpolating the aggregated SMPs from the gauged to the ungauged pixels, labeled as SMP_{pix}.

4. Interpolating the SMPs obtained for satellite data at the gauged pixels to the ungauged pixels, labeled as SMP_{sat}.

5. Computing the biases of SMP_{sat} at ungauged pixels by subtracting the result of step 2 (SMP_{pix}) to the result of step 3 (SMP_{sat}).

6. Finally, biases are adjusted by subtracting the modeled bias (SMP_{sat}) from original bias (SMP_{osat}), the local SMPs of sat is calculated as $SMP_{adjusted} = SMP_{osat} - \Delta SMP_{sat}$

$$= SMP_{osat} - (SMP_{sat} - SMP_{pix}).$$

The main advantage of this method is considering the biases of both frequency and intensity of satellite products and correct them simultaneously. Besides, it can correct the biases at the satellite pixels devoid of gauges. Nevertheless, it still suffers from incapability to capture the time variation of precipitation data particularly for small scale rainfall features.

Quantile mapping method

Quantile Mapping (QM) (also referred to as quantile matching, cumulative distribution function matching, quantile-quantile transformation) attempts to find a transformation,

$$P_o = h(P_m) \qquad (4.7)$$

of a modeled variable P_m such that its new distribution equals the distribution of the observed variable P_o. If the distribution of the variable of interest is known, the transformation h is defined as

$$P_o = F_O^{-1}\big(F_m(P_m)\big) \qquad (4.8)$$

where F_m is the CDF of P_m and F_O^{-1} is the inverse CDF (or quantile function) corresponding to P_o (Gudmundsson et al., 2012).

QM can be achieved by using theoretical distributions to solve Eq. (4.8). For this purpose, different parametric distributions such as Gamma, Bernoulli, and Weibull can be used to model the precipitation characteristics (occurrence and intensity). The parameters of the distributions can be estimated by maximum likelihood methods for both observed and modeled data.

A common approach, that is also applied in this study, to solve Eq. (4.8) is the empirical CDF of observed and modeled values instead of assuming parametric distributions. The empirical CDFs are approximated using tables of the empirical percentiles. Values in between the percentiles are approximated using linear interpolation (Gudmundsson et al., 2012).

In climate change analyses to bias correct General Circulation Models (GCMs), QM equates cumulative distribution functions (CDFs) $F_{o,h}$ and $F_{m,h}$ of observed data $x_{o,h}$ denoted by the subscript o, and modeled data $x_{m,h}$, denoted by the subscript m, in a historical period, denoted by the subscript h, respectively. This leads to the following transfer function,

$$x_{m,p}(t) = F_{o,h}^{-1}\big\{F_{m,h}\big[x_{m,h}(t)\big]\big\} \qquad (4.9)$$

where $x_{m,p}(t)$ is a bias-corrected value at time t within some projected (future) period, denoted by the subscript p. F^{-1} denotes the Inverse CDF (ICDF) of observed data. If CDFs and inverse CDFs (i.e., quantile functions) are estimated empirically from the data, the algorithm can be illustrated with the aid of a quantile–quantile plot, which is the scatterplot between empirical

quantiles of observed and modeled data. In this case, QM amounts are found by interpolating between points in the CDF plot of the observed data. The transfer function is constructed using information from the observed gauges. QM, like all statistical post-processing algorithms, relies strongly on an assumption that the climate model biases to be corrected are stationary (i.e., that characteristics in the historical period will persist into the future) (Cannon et al., 2015; Najafi et al., 2016 and 2017).

As the most widely recognized and recent technique in bias correction of SPPs, QM has been used in several studies (Zhang and Tang, 2015; Yang et al., 2016; Gumindoga et al., 2019; Alharbi, 2019). In the field of SPPs the Eq. 4.9 will be changed slightly. In this case

$$\widehat{P_s} = F_O^{-1}\big(F_s(P_s)\big) \qquad (4.10)$$

where F_s is the CDF of P_s (satellite precipitation), F_O^{-1} is the inverse CDF of P_o (observed precipitation), and $\widehat{P_s}$ is bias-corrected satellite precipitation. There are several statistical transformations related to the QM method for modeling the CDFs. The distribution-derived transformation uses theoretical distribution to solve Eq. 4.10. Parametric transformations are used directly to model the quantile-quantile relationship

$$P_o = h(P_s) \qquad (4.11)$$

Instead of assuming parametric distributions, nonparametric transformations use empirical CDFs to solve Eq. 4.10 or nonparametric regressions such as cubic smoothing splines (e.g. Hastie et al., 2001）to solve Eq. 4.11. The smoothing spline is only fitted to the fraction of the CDF corresponding to observed wet days and modeled values below this are set to zero. The nonparametric QM approach is highly valued for bias adjustment, because it does not rely on any predetermined function and as such provides more flexibility. The main advantage of QM method is that it can effectively capture the evolution of mean and variability of estimated data (SPPs) while matching all statistical moments (Cannon et al., 2015). However, it may not capture the occurrences of the daily precipitation (Ajaaj et al., 2016). Beside, QM is unable to

correct estimated values outside the codomain of the observed period (Passow and Donner, 2020).

Regression Quantile Mapping (RQM) method

RQM is a bias correction approach based on (linear) regression models which allow to design of transfer functions over ungauged sites and adjust the biases of temporally and spatially (Passow and Donner, 2019). In this study, the RQM method is used for bias correcting of SPP, which is explained in detail in the methodology section.

4.2. Data

In this chapter, the bias correction technique is applied on the daily IMERG PrCal product aggregated from half hourly original data. Also, daily in situ rain gauge data aggregated from hourly are used as reference true data in bias correction of satellite estimates. These two datasets have already been explained in detail in previous chapter. In this section the NARR covariate data sources required for implementing the regression-based bias correction method are described.

4.2.1. NCEP North American Regional Reanalysis: NARR

Reanalysis a systematic approach to produce data sets for climate monitoring and research. Reanalyses are created via an unchanging ("frozen") data assimilation scheme and model(s) which ingest all available observations every 6-12 hours over the period being analyzed. This unchanging framework provides a dynamically consistent estimate of the climate state at each time step. The one component of this framework which does vary are the sources of the raw input data. This is unavoidable due to the ever-changing observational network which includes, but is not limited to, radiosonde, satellite, buoy, aircraft, and ship reports (Dee et al., 2014). The NARR project is an extension of the NCEP Global Reanalysis which is run over the North American Region. It covers 1979 to near present and is provided in gridded output at ~32 km (0.3°) resolution at three-hourly, daily, and monthly on a Northern Hemisphere Lambert Conformal Conic grid for all variables. The NARR model takes in, or assimilates, a great

amount of observational data to produce a long-term picture of weather over North America. The data that are assimilated in order to initialize the model to real-world conditions are temperatures, winds, and moisture from radiosondes as well as pressure data from surface observations. Also included in this dataset are dropsondes, pibals, aircraft temperatures and winds, satellite radiance (a measure of heat) from polar (orbiting Earth) satellites, and cloud drift winds from geostationary (fixed at one location viewing Earth) satellites.

In this study we used two different types of NARR variables, first, daily time-varying variables in order to apply as covariates in regression quantile mapping which help to predict bias-corrected satellite estimates over both gauged and ungauged pixels. Second, the longterm monthly mean variables that are utilized to calculate the climatic distance between satellite pixels used in the interpolation process of bias-corrected data. The first group of covariates includes 1) daily accumulated total precipitation, the daily mean of 2) air temperature, 3) convective potential energy, 4) non-convective cloud cover, 5) dew point temperature at 2m, 6) specific humidity at 2m, 7) pressure at mean sea level, 8) pressure vertical velocity. This group of variables is selected based on some criteria examined (explained in the methodology chapter) in order to make sure that there is a relationship between them as independent variables and gauge/satellite precipitation data as the dependent variable. The second group of variables includes 1) invariant parameters such as elevation (the elevation in this study is provided from Canadian Digital Elevation Model (CDEM) which has the 20 meter base spatial resolution and then is resampled to 10 km (consistent with IMERG spatial resolution) using nearest neighbor resampling technique in GIS software), and the long-term monthly mean of 2) total precipitation, 3) convective potential energy, 4) non-convective cloud cover, 5) dew point temperature at 2m, 6) specific humidity at 2m, 7) Shortwave Radiation Flux, 8) Geopotential Height, 9) Vegetation Index.

Some of the abovementioned variables are defined as follows:

Convective potential energy:

In meteorology, convective available potential energy (commonly abbreviated as CAPE), is the integrated amount of work that the upward (positive) buoyancy force would apply on a given mass of air (called an air parcel) if it rose vertically through the entire atmosphere. Positive CAPE will cause the air parcel to rise, while negative CAPE will cause the air parcel

to sink. Nonzero CAPE is an indicator of atmospheric instability in any given atmospheric sounding, a necessary condition for the development of cumulus and cumulonimbus clouds with attendant severe weather hazards (Barry and Chorley, 1998).

Shortwave Radiation Flux:

Shortwave flux is a result of specular and diffuse reflection of incident shortwave radiation by the underlying surface. This shortwave radiation, like solar radiation, can have a profound impact on certain biophysical processes of vegetation, such as canopy photosynthesis and land surface energy budgets, by being absorbed into the soil and canopies.[4] As it is the main energy source of most weather phenomena, the solar shortwave radiation is used extensively in numerical weather prediction (Kantha and Clayson, 2000).

Geopotential Height:

Geopotential height approximates the actual height of a pressure surface above mean sea-level. Therefore, a geopotential height observation represents the height of the pressure surface on which the observation was taken. Since cold air is denser than warm air, it causes pressure surfaces to be lower in colder air masses, while less dense, warmer air allows the pressure surfaces to be higher. Thus, heights are lower in cold air masses, and higher in warm air masses. (Hofmann-Wellenhof and Moritz, 2005).

Vegetation Index:

A vegetation index is a single number that quantifies vegetation biomass and/or plants vigor for each pixel in a remote sensing image. The index is computed using several spectral bands that are sensitive to plant biomass and vigor (Huete et al., 2002). Vegetation affects rainfall through the process of transpiration. When plants convert carbon dioxide and sunlight into carbohydrates via photosynthesis, they lose water through their leaves.

4.3. Methodology

As in this study, bias correction of daily SPP is investigated, the original half-hourly PrCal product of IMERG and corresponding hourly ground observed data are both aggregated to daily time series during the 5-year study period from 2014 to 2018.

Considering previously mentioned limitations of several bias correction methods, in this study a more reliable model that corrects biases of satellite precipitation data spatially and temporally at gauged as well as ungauged sites is developed. The fundamental of the proposed framework is based on RQM that is explained as follows.

4.3.1. Regression Quantile Mapping

At first, the linear quantile regression is described.

Linear quantile regression

Given a desired probability τ and n random continuous variables Y_i ($i=1, \ldots, n$) with CDF $F_{Y_i}(y) = P(Y_i \leq y)$, the τ-th quantile $Q_{Y_i}(\tau)$ of Y_i is defined as

$$Q_{Y_i}(\tau) = F_{Y_i}^{-1}(\tau) = inf\{y\epsilon\mathbb{R}: \tau \leq F_{Y_i}(y)\} \quad (4.12)$$

where the infimum (inf) is the greatest element in $\{y\epsilon\mathbb{R}\}$ that is less than or equal to all elements of $\{y\epsilon\mathbb{R}: \tau \leq F_{Y_i}(y)\}$ for $\tau\epsilon[0,1]$ and $F_{Y_i}^{-1}$ is the quantile function, i.e., the inverse CDF, of Y_i. In linear quantile regression (QR), Q_{Y_i} is assumed to depend linearly on a vector of predictor variables $X_i = (X_{i1}, \ldots, X_{ip})^T$ such that

$$Q_{Y_i}(\tau|X_i) = X_i^T . \beta_\tau \quad (4.13)$$

where $\beta_\tau = (\beta_{\tau,1}, \ldots, \beta_{\tau,p})^T$ is a vector of unknown regression parameters estimated by solving the minimization problem

$$\hat{\beta}_\tau = arg \min_{\beta \epsilon R^p} \sum_{i=1}^n \rho_\tau(Y_i - X_i^T \beta_\tau) \qquad (4.14)$$

where $\rho_\tau(.)$ is the check function given by

$$\rho_\tau(e) = \tau e I_{[0,\infty)}(e) - (1 - \tau)e I_{(-\infty,0)}(e)$$

$$= e\left(\tau - I_{(-\infty,0)}(e)\right) = \left[\frac{1}{2} + \left(\tau - \frac{1}{2}\right) sgn(e)\right]|e| \qquad (4.15)$$

with $I_A(.)$ as the indicator function of some interval A and sgn (.) is the sign function.

In most practical applications of QR, one is interested in the full conditional distribution of the Y_i. Hence, QR is not only used to estimate the conditional sample quantile q_τ for a single τ, but rather to estimate a full set of $q_{\tau_1}, \ldots, q_{\tau_k}$ for various $\tau_1, \ldots, \tau_k$ values. However, since Eq. (4.15) is solved separately for each desired τ, the estimated quantile functions may commonly intersect with each other. In such a case, a quantile of a higher order switches places with a quantile of a lower order. This undesired phenomenon is called quantile crossing (QC) and is a well-known problem of QR. Several approaches are possible to prevent QC; here, a smoothing spline is used to smooth the conditional distribution which is initially estimated independently for each τ (See section 4.3.2).

Regression quantile mapping

The procedure of bias correction of satellite data is based on the true value obtained from ground stations observed records. Thus, at first satellite pixels where there is at least one rain gauge located inside them are determined and their biases are corrected using the method developed here. Then, biases of other ungauged pixels are adjusted.

RQM is an extension of QM which uses conditional inverse CDFs instead of usual inverse CDFs. In this case, in addition to removing the systematic biases, time or state-dependent characteristics reflected by time or other climate variables as covariates (e.g., long term trends or the annual cycle) and the response of the climate variable to changes in associated climate predictors (here denoted as X) are also adjusted. Thus, the equation is provided as:

$$\hat{y}_s = F^{-1}_{y_o|X}\left(F_{y_s|X}(y_s)\right) \qquad (4.16)$$

where y_s is the quantile of original satellite data, $F_{y_s|X}$ is the CDF of satellite original data conditional on covariate X, $F^{-1}_{y_o|X}$ is the inverse CDF (or quantile function) of the observations y_o conditional on the covariates X, and $\hat{y}_s$ is the bias-corrected satellite data.

In this study the linear QR is used as we assume a simple linear relationship between the predictors (covariates) and precipitation (dependent variable). It is noted that RQM algorithm requires the predictive modeling of CDFs and therefore the QR approach and regression model should be selected with respect to the intrinsic properties of the climate variable and predictors considered (e.g., for modeling seasonality, a harmonic regression function concerning time varying covariate might be an appropriate choice) (Passow and Donner, 2020). As in this study, the aim is bias correcting of daily time series of SPP, so various covariates at daily scale as independent variables have been selected and applied for developing the regression model (See section 4.2).

There are different ways in which the CDFs can be approximated, including both parametric and nonparametric methods. A common approach is to estimate empirical CDFs as an initial approximation. Since these CDFs are step functions, piecewise linear interpolation is used to estimate probabilities and quantiles for values in between.

For the approximation of the conditional CDFs and inverse CDFs we combine QR with a nonparametric smoothing spline (see next section) estimator. Let $\tau_1,\ldots,\tau_k$ be a set of probabilities, where $\tau\epsilon[0,1]$ and $\tau_1 < \ldots < \tau_k$. QR is applied to each time series (y_o and y_s) to estimate the regression parameters

$$\hat{\beta}_{\tau z} = \left(\hat{\beta}_{\tau 1, z}, \dots, \hat{\beta}_{\tau k, z} \right)^{T} \quad (4.17)$$

for the observation ($z=o$) and the satellite data ($z=s$) for each τ_i separately.

With the linear QR parameters known, we can use the corresponding QR model (Eq. 4.11) to predict quantile curves q_τ for each time point t in the study period of interest (Tp) with respect to the predictors X_p of that period and time-dependent (linear) effect of the predictors (i.e., $\hat{\beta}_{\tau o}$ and $\hat{\beta}_{\tau s}$) as

$$\hat{q}_{\tau, z}(t) = X_p . \hat{\beta}_{\tau, z} \,\forall\, t \in T_P \quad (4.18)$$

In Eq. (4.18), two different sets of quantile curves ($\hat{q}_{\tau, o}$ and $\hat{q}_{\tau, s}$) are obtained, which are estimated based on the same set of predictors (X_p) and can be evaluated over the same period (Tp). The difference between them only being the regression coefficients $\hat{\beta}_{\tau, z}$. [The actual quantile matching (Eq. 4.16) performed in a later step of the algorithm will make use of these $\hat{q}_{\tau, z}$.] Therefore, RQM is the only bias correction method in which all necessary CDFs and inverse CDFs are defined for Tp.

At this point, the $\hat{q}_{\tau, z}$ still resemble step functions of $F_{y_o|X}^{-1}$ and are therefore not suitable for distribution mapping. To obtain a continuous approximation of $F_{y_o|X}^{-1}$ we use the smoothing spline estimator to provide estimates for the unknown quantile values.

4.3.2. Smoothing Spline

In mathematics, a spline is a special function defined piecewise by polynomials. The term "spline" is used to refer to a wide class of functions that are used in applications requiring data interpolation and/or smoothing. The data may be either one-dimensional or multi-dimensional.

The B-spline basis is a commonly used spline basis that is based on a special parametrization of a cubic spline. The B-spline is based on the knot sequence (De Boor, 1978). Cubic splines

are created by using a cubic polynomial in an interval between two successive knots. A spline of order n is a piecewise polynomial function of degree n-1 in a variable x. The values of x where the pieces of the polynomial meet are known as knots denoted $t_0, t_1, t_2, \ldots, t_n$ and sorted into non-decreasing order. When the knots are distinct, the first n-2 derivatives of the polynomial pieces are continuous across each knot. When r knots are coincident, then only the first n-r-1 derivatives of the spline are continuous across that knot.

In this study "create.bspline.basis" function in R is used to make a connection between known quantiles and estimate unknown ones by computing the appropriate linear combination. Functional data objects are constructed by specifying a set of basic functions and a set of coefficients defining a linear combination of these basis functions. The B-spline basis is used for non-periodic functions. B-spline basis functions are polynomial segments jointed end-to-end at argument values called knots, breaks, or join points. The segments have specifiable smoothness across these breaks. B-splne basis functions have the advantages of very fast computation and great flexibility (Ramsay et al. 2009).

In this study RQM is developed by fitting the linear model between precipitation and some time-varying predictors as independent variables at each quantile. For selecting predictors, the main part is checking the multicollinearity and the strength of the correlation between multiple independent variables. Multicollinearity occurs when independent variables in a regression model are correlated. Fortunately, there is a simple test to assess multicollinearity in the regression model. The variance inflation factor (VIF) identifies the correlation between independent variables and the strength of that correlation. Statistical software calculates a VIF for each independent variable. VIFs start at 1 and have no upper limit. A value of 1 indicates that there is no correlation between this independent variable and any others. VIFs between 1 and 5 suggest that there is a moderate correlation, but it is not severe enough to warrant corrective measures. VIFs greater than 5 represent critical levels of multicollinearity where the coefficients are poorly estimated, and the p-values are questionable. Therefore, among around 15 climatic variables, just 6 of them which had more correlation with satellite and observed precipitations individually as well as a reasonable VIF value (the VIF value for each of those six variables was about 1.5), have been selected for fitting the RQM model. It is also worthwhile to mention that, the original gridded covariates have been provided in ~32 km, so,

before using, they are resampled to ~10 km in order to be matched with satellite pixels' resolution.

4.3.3. Clustering

One objective of this chapter is to correct the bias over ungauged pixels. For this purpose, we use the parameters extracted from RQM developed at gauged pixels, then interpolate them over ungauged pixels. In order to have more reliable interpolation over the entire study area (Canada wide) with diverse climatic conditions, we clustered the area based on the climatic distance between satellite pixels. Then, the information of gauged pixels at each cluster is used to be interpolated over ungauged pixels within that cluster. In the following the clustering methods utilized in this study are explained.

- **K-means Clustering method**

A cluster refers to a collection of data points grouped because of certain similarities.

At first, a target number k, which refers to the number of centroids needed in the dataset is defined. A centroid is the imaginary or real location representing the center of the cluster. Every data point is allocated to each of the clusters by reducing the in-cluster sum of squares.

K-means clustering is an extensively used technique for data cluster analysis. It is easy to understand and delivers training results quickly. The K-means algorithm identifies k number of centroids, and then allocates every data point to the nearest cluster while keeping the centroids as small as possible. The *'means'* in the K-means refers to averaging of the data; that is, finding the centroid. To process the learning data, the K-means algorithm in data mining starts with the first group of randomly selected centroids, which are used as the beginning points for every cluster, and then performs iterative (repetitive) calculations to optimize the positions of the centroids. It halts creating and optimizing clusters when either:

The centroids have stabilized — there is no change in their values because the clustering has been successful.

The defined number of iterations has been achieved.

However, slight variations in the data could lead to high variance. Furthermore, clusters are assumed to be spherical and evenly sized, something which may reduce the accuracy of the K-means clustering (Amorim and Hennig, 2015)

- **Hierarchical Clustering method**

Hierarchical clustering is an alternative approach to k-means clustering for identifying groups in the dataset. It does not require to pre-specify the number of clusters to be generated as is required by the k-means approach. Furthermore, hierarchical clustering has an added advantage over K-means clustering in that it results in an attractive tree-based representation of the observations, called a dendrogram.

Hierarchical clustering can be divided into two main types: agglomerative and divisive.

Agglomerative clustering: It's also known as AGNES (Agglomerative Nesting). It works in a bottom-up manner. That is, each object is initially considered as a single-element cluster (leaf). At each step of the algorithm, the two clusters that are the most similar are combined into a new bigger cluster (nodes). This procedure is iterated until all points are member of just one single big cluster (root) (see figure below). The result is a tree which can be plotted as a dendrogram.

Divisive hierarchical clustering: It is also known as DIANA (Divise Analysis) and it works in a top-down manner. The algorithm is an inverse order of AGNES. It begins with the root, in which all objects are included in a single cluster. At each step of iteration, the most heterogeneous cluster is divided into two. The process is iterated until all objects are in their own cluster (Figure 22).

Note that agglomerative clustering is good at identifying small clusters. Divisive hierarchical clustering is good at identifying large clusters. For both K-means and Hierarchical, we measure the (dis)similarity of observations using Euclidean distance (i.e. climatic distance) measure. In R, the Euclidean distance is used by default to measure the dissimilarity between each pair of

observations. However, to measure the dissimilarity between two clusters of observations, we used the Ward method in this study. Ward's minimum variance method minimizes the total within-cluster variance. At each step, the pair of clusters with minimum between-cluster distance are merged (Ward, 1963).

It is noted that to perform cluster analysis, generally, the data should be prepared as follows:

Rows are observations (individuals) and columns are variables. Any missing value in the data must be removed or estimated. The data must be standardized (i.e., scaled) to make variables comparable. Standardization consists of transforming the variables such that they have mean zero and standard deviation one.

In this study, a number of nine covariates for each month at a given pixel are utilized to calculate the climatic distance between that pixel and other pixels. The climatic distance is then used between covariates for clustering the study area. It should be noted that all processes of bias correction are conducted at each month of the study period separately, so the long term monthly covariates of NARR data are utilized which give a unique value at each month for a given pixel. These values have been provided in the monthly mean from 1979 to 2019. Having the all information at each resampled 0.1-degree pixel makes it possible to implement clustering techniques for classifying all satellite pixels in several groups.

Both K-means and Hierarchical methods were used to select the best number of clusters. However, differences were negligible between two methods and based on "silhouette" criterion for obtaining the optimal number of clusters, nine clusters were selected. Silhouette refers to a method of interpretation and validation of consistency within clusters of data. In this study, the silhouette is calculated with the Euclidean distance which is a metric for measuring the "ordinary" straight-line distance between two points. The silhouette ranges from -1 to $+1$, where a high value indicates that the object is well matched to its own cluster and poorly matched to neighboring clusters. If most objects have a high value, then the clustering configuration is appropriate (for more detail and equations see Rousseeuw, 1987). Figure 18 shows the average value of 0.52 for all clusters which means a reasonable structure has been found (Rousseeuw, 1987).
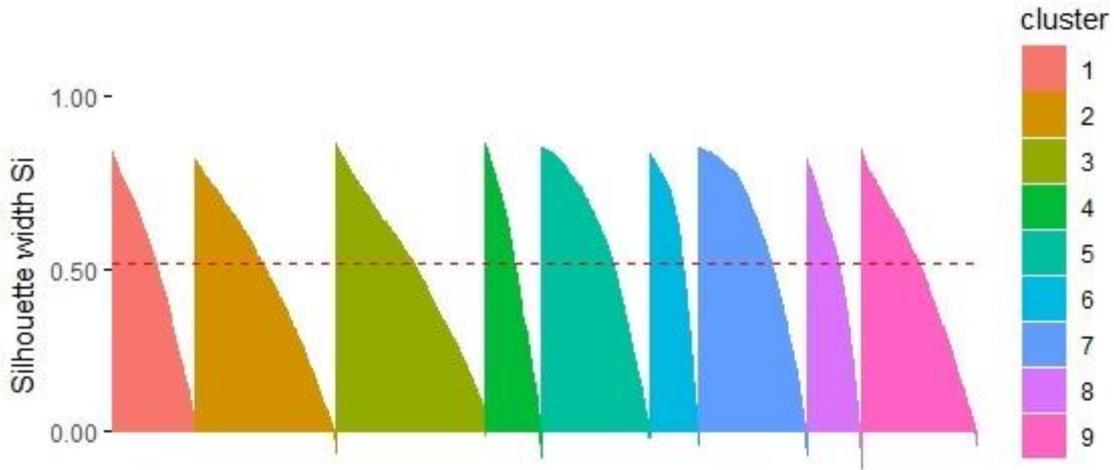
Figure 18. The average silhouette width of all nine clusters over the study area

4.3.4. CDF Segmentation

The idea of using partitioned CDF in this study came from Grillakis et al. 2013 who presented an improved quantile mapping based bias correction method named multi-segment statistical bias correction (MSBC) quantile mapping for general circulation model (GCM) simulated daily precipitation. The method used different instances of gamma function that were fitted on multiple discrete segments on the precipitation cumulative distribution function (CDF), instead of the common quantile-quantile approach that uses one theoretical distribution to fit the entire CDF. This improved the ability of the method to better transfer the observed precipitation statistics to the raw GCM data (Grillakis et al. 2013)

In the proposed RQM method of this study, we split the precipitation CDF into five equal, discrete, sequential segments from 0 to 0.2, 0.21 to 0.4, 0.41 to 0.6, 0.61 to 0.8, and 0.81 to 1. Then RQM is performed at each segment separately to correct the CDF of satellite data in the range of that segment. For example, at first, the data with CDF between 0 and 0.2 is corrected, then the data with CDF between 0.2 and 0.4, etc. Also, the procedure is applied for each calendar month separately.

4.3.5. Interpolating/Extrapolating by IDW

In this research, we use the Inverse Distance Weighted (IDW) interpolation method to obtain the parameters of the developed RQM model at gauged pixels over ungauged ones. IDW is a deterministic spatial interpolation approach to estimate an unknown value at a location using some known values with corresponding weighted values. The basic formula can be seen in equation 4.17. x* is unknown value at a location to be determined, w is the weight, and x is known point value. The weight is inverse distance of a point to each known point value that is used in the calculation. Simply the weight can be calculated using equation 4.20.

$$x^* = \frac{w_1 x_1 + w_2 x_2 + \ldots + w_n x_n}{w_1 + w_2 + \ldots + w_n} \qquad (4.19)$$

$$w_i = \frac{1}{d_{ix^*}^p} \qquad (4.20)$$

where P variable stands for Power. There is no particular rule in defining the P value, but from the equation, we can see that the higher P value will give lower weight. An experiment is suggested in defining the optimum P value. It could be done by taking a small portion of sample point as testing/validation dataset. Then start with a small P value, do the IDW interpolation and calculate the RMSE between the interpolation result and the actual sampling value. Iterate by increasing the P value step by step and calculate the RMSE. The lowest RMSE is the optimum P value which is given the smallest error between the interpolation and actual value. Here P value of 2 with minimum RMSE was selected. Also, in this study, d is the distance between ungauged pixel and gauged one and x is the covariates of ungauged pixel.

4.3.6. Validation using Bootstrap Technique

In this research to validate the performance of bias correction method implemented over the ungauged pixels, the bootstrap technique is used. Because at ungauged pixels there is no gauge to compare the bias-corrected satellite precipitation with the observed data directly, we assume some gauged pixels at each cluster as ungauged, then validate the bias correction model at these pixels. For taking samples of gauged pixels (considering as ungauged) and to make sure that

the samples are randomly selected over each cluster bootstrap can be used an appropriate technique. Bootstrap is a resampling technique used to estimate statistics on a population by sampling a dataset with replacement.

This procedure is summarized as follows:

- Draw a number of samples with replacement with the chosen size
- Calculate the statistic on the samples
- Calculate the mean of the calculated sample statistics

In this study, we validate the proposed bias correction method over the 10% of gauged pixels at each cluster by using the bootstrap sampling technique. For this purpose, we consider each cluster separately and take 10% of gauged pixels assumed as ungagged in a given cluster. Then, the bias-corrected parameters extracted from that 90% gauged pixels are interpolated over the taken 10% assumed ungauged pixels. Afterward, by applying the covariates into interpolated parameters, the bias-corrected satellite data is obtained for 10% pixels. By comparing the bias-corrected data with observed gauge records, the statistics of RBias, RMSE, and CC are calculated for those 10% pixels. By iterating the sampling of 10% pixels for 1000 times by bootstrapping (sampling a dataset with replacement), each pixel may select several times. Therefore, the mean metrics values of each pixel are calculated and considered as statistic indices of validation the bias correction method.

4.3.7. Methodology Overview

To summarize the methodology of bias correction used in this study the outlines of all previously explained steps are provided as follows:

- Aggregating half-hourly satellite data and hourly ground gauges to daily
- Assigning satellite pixels containing gauges
- Extracting gridded daily time series of covariates and resampling them at each satellite pixel resolution

- Extracting and resampling long term monthly covariates at each satellite pixel (during the long term period from 1979 to 2019 for each month there is just one value)

- Extracting and resampling invariant covariates (e.g. elevation) at each satellite pixel

- Clustering all satellite pixels using K-means and Hierarchical methods based on the climatic distance calculated between pixels by utilizing invariant covariates

- Partitions the CDF of observed data of gauged pixels into discrete segments

- Applying RQM method at gauged pixels to each CDF segment separately

- Performing smoothing spline to smooth the conditional distribution initially estimated independently for each quantile

- Extracting the parameters of RQM at all bias-corrected gauged pixels

- Interpolating/extrapolating the parameters over ungauged pixels for each cluster separately using the IDW method (the weights are calculated based on the distance between gauged pixels and ungauged ones inside each cluster)

- Applying the interpolated bias-corrected parameters to the time-varying covariates at each ungauged pixel in order to obtain bias-corrected satellite estimates at those pixels

- Validation of the results by comparing the obtained bias-corrected satellite estimates at some random selected gauged pixels put aside in previous steps (assumed as ungauged pixels) with observed data at those pixels using a bootstrap sampling technique

4.4. Results and Discussion

The obtained results of the bias correction technique are presented in two sections, first, for gauged pixels where there is at least one rain gauge inside each pixel. Second, for ungauged pixels by interpolating parameters obtained from implemented RQM in gauged pixels. These results are extracted for each cluster separately. Table 4 lists the information for each cluster. As shown in this table the number of pixels at each cluster and a corresponding number of gauged pixels varies and they are not equal. The reason is that we generated nine clusters based on nine climatic and topographic covariates (previously explained) dissimilarity in the K-means clustering method (Figure 19). So, each cluster contains a different number of gauged pixels that can affect on the interpolation of bias-corrected parameters. Further explanations are

provided under the specific results of this section. Figures 20 to 28 provide the spatial patterns of covariates over the study area. As displayed, the classified study area (Figure 19) is a combination of those nine figures (Figures 20-28). The correlation coefficient between nine covariates is provided in Table 5. As depicted the highest correlation is between "Specific Humidity" and "Dew Point Temperature" with a value of 0.98 and the lowest correlation is between "Non-Convective Cloud Cover" and "Dew Point Temperature" with a value of -0.03.
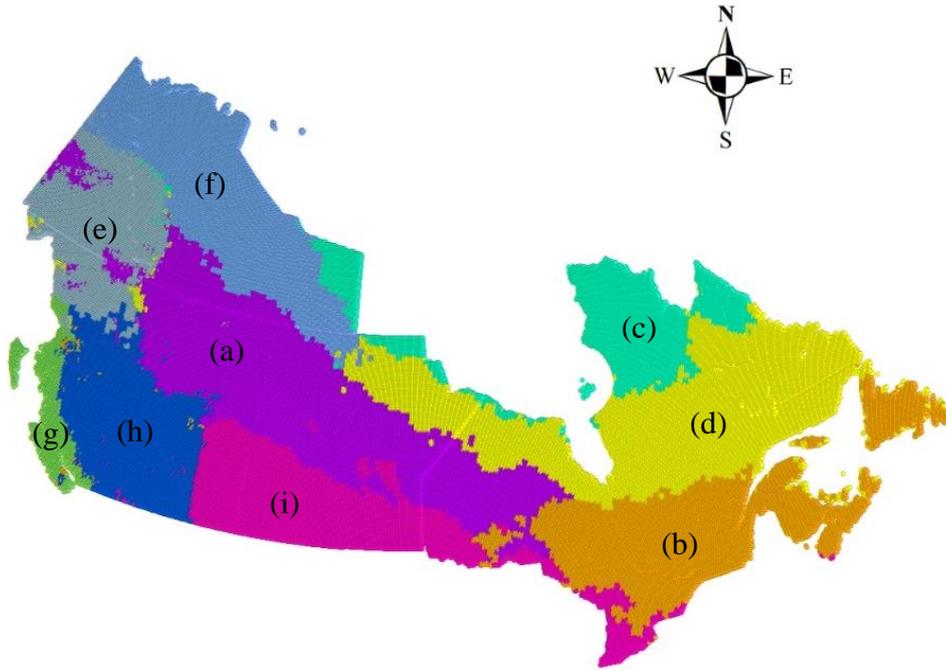
Figure 19. Clusters pattern of the study area by K-means ((a) represents cluster no.1 and (i) represents cluster no.9)

Table 4. Covariates averaged at each cluster

Cluster No.	Number of Pixels	Number of Gauged Pixels	Elevation (m)	Precipitation (mm/month)	Specific Humidity (g/m^3)	Dew Point Temperature (°C)	Convective Potential Energy (J/Kg)	Non-Convective Cloud Cover (%)	Shortwave Radiation Flux (W/m^2)	Geopotential Height (m)	Vegetation Index (%)
1	15363	42	255	30.6	2.84	-5.4	5.34	58.3	62	9674	29.6
2	14448	102	248	32	3.2	-3.67	7.81	57.7	69	10033	42.7
3	13128	29	294	31	2.58	-7.15	4.4	63.3	60	9595	30
4	13226	86	375	42.5	3.0	-4.8	4.3	64.7	53	9819	38.8
5	11507	32	1525	25.5	2.5	-8.1	6.5	56.8	50	9236	24.5
6	13534	26	402	21.3	2.15	-9.6	2.15	56.9	45	9120	17.8
7	5871	55	803	86	3.3	-3.84	14.36	60.6	59	9857	42.4
8	11901	111	1603	43	3.3	-3.67	14.05	55.4	65	9881	36.22
9	11131	103	518	22.1	3.23	-3.8	11.55	54.8	67	9874	26.5

Table 5. Correlation Coefficient between covariates

Covariates	Elevation	Precipitation	Specific Humidity	Dew Point Temperature	Convective Potential Energy	Non-Convective Cloud Cover	Shortwave Radiation Flux	Geopotential Height	Vegetation Index
Elevation	1	-0.21	-0.45	-0.55	-0.29	-0.04	0.62	0.1	-0.12
Precipitation	-0.21	1	0.68	0.63	0.6	0.24	-0.33	0.31	0.15
Specific Humidity	-0.45	0.68	1	0.98	0.87	-0.1	-0.4	0.55	0.45
Dew point Temperature	-0.55	0.63	0.98	1	0.8	-0.03	-0.54	0.44	0.44
Convective Potential Energy	-0.29	0.6	0.87	0.8	1	-0.22	-0.21	0.52	0.33
Non-Convective Cloud Cover	-0.04	0.24	-0.1	-0.03	-0.22	1	-0.47	-0.43	-0.15
Shortwave Radiation Flux	0.62	-0.34	-0.4	-0.54	-0.22	-0.47	1	0.43	0.02
Geopotential Height	0.1	0.31	0.55	0.44	0.52	-0.43	0.43	1	0.66
Vegetation Index	-0.12	0.15	0.45	0.44	0.33	-0.15	0.02	0.66	1
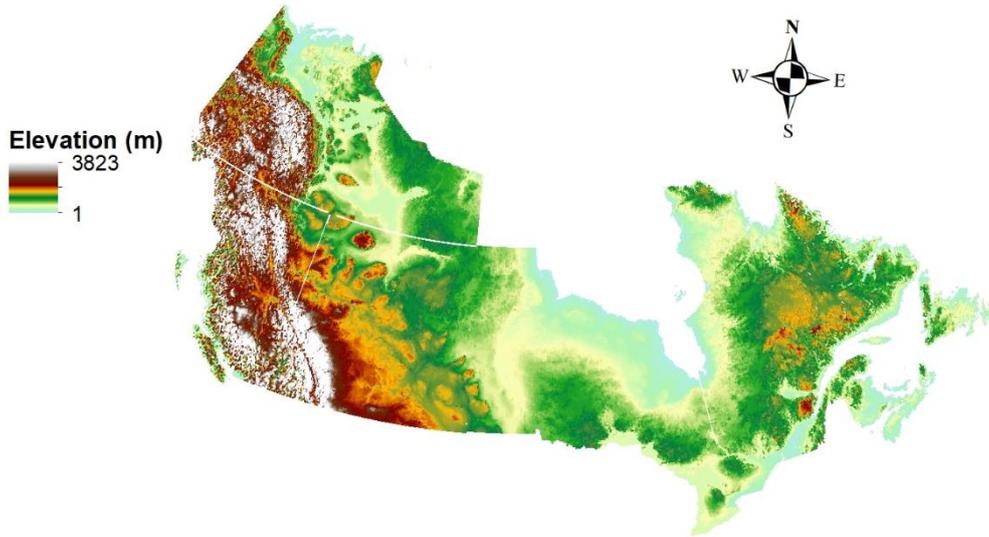
Figure 20. Elevation map of the study area
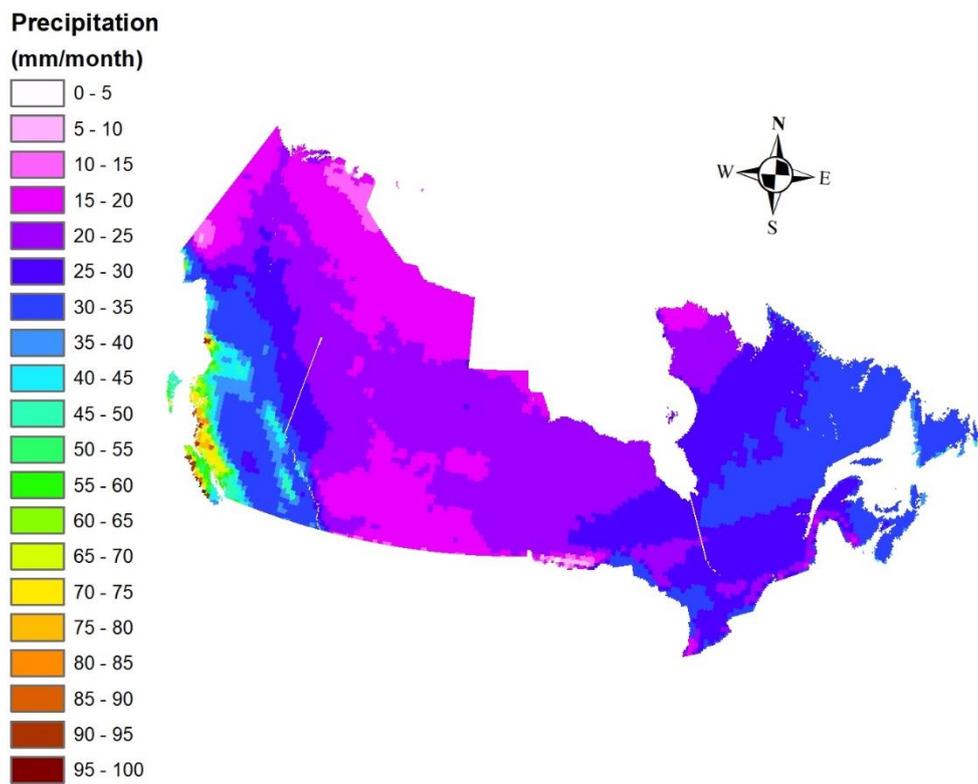
Figure 21. Monthly mean precipitation from 1979 to 2019
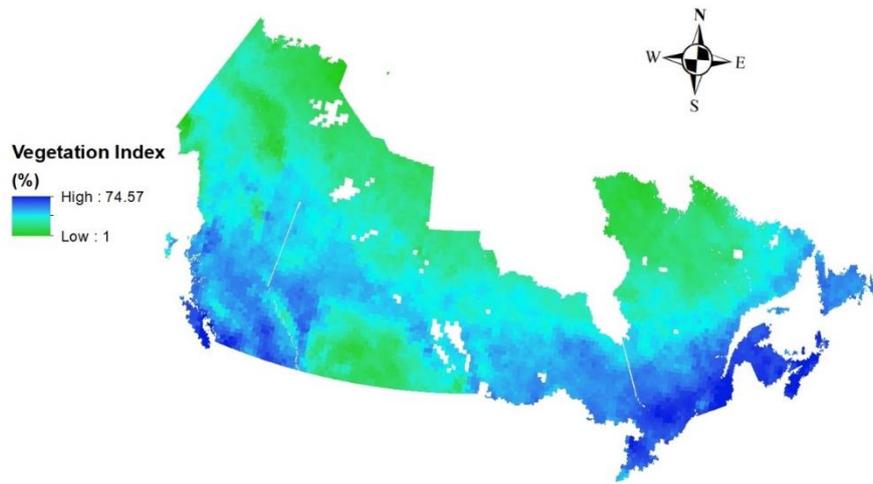
Figure 22. Monthly mean vegetation index from 1979 to 2019
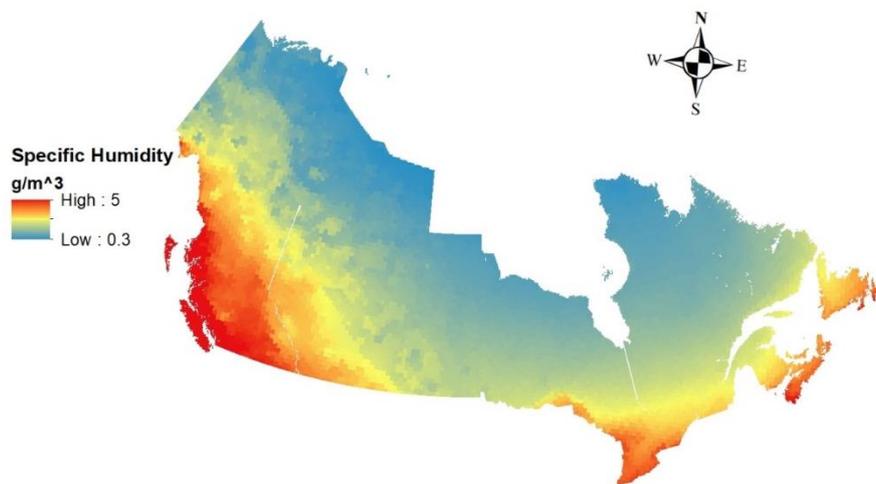
Figure 23. Monthly mean specific humidity from 1979 to 2019

Figure 24. Monthly mean dew point from 1979 to 2019
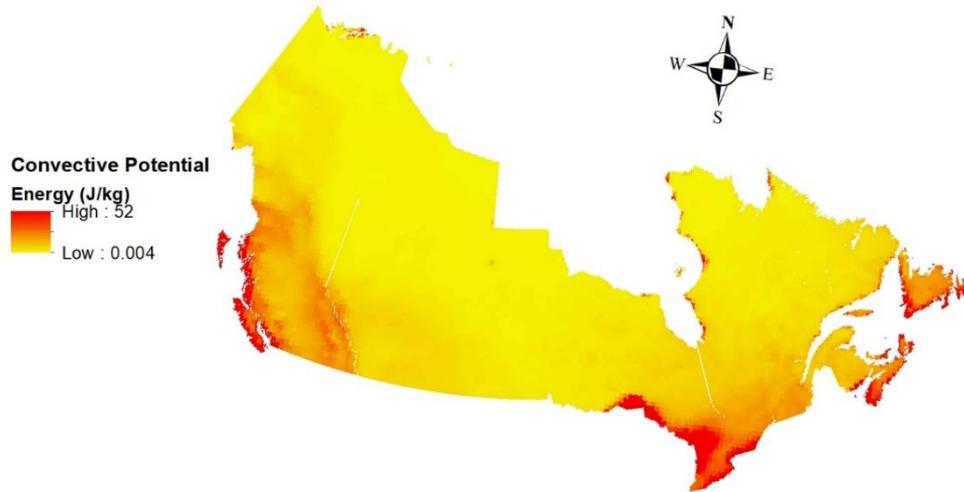
Figure 25. Monthly mean convective potential energy from 1979 to 2019
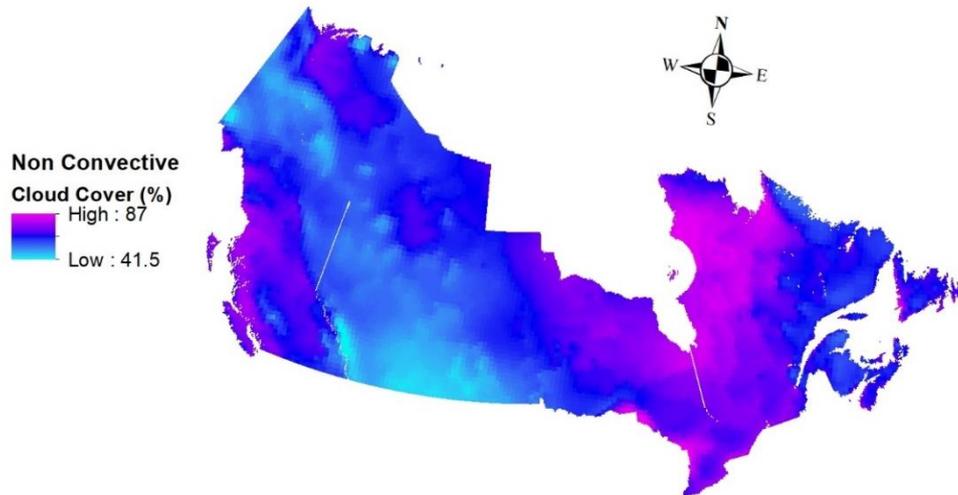
Figure 26. Monthly mean non-convective cloud cover from 1979 to 2019

Figure 27. Monthly mean downward shortwave radiation flux at the surface from 1979 to 2019
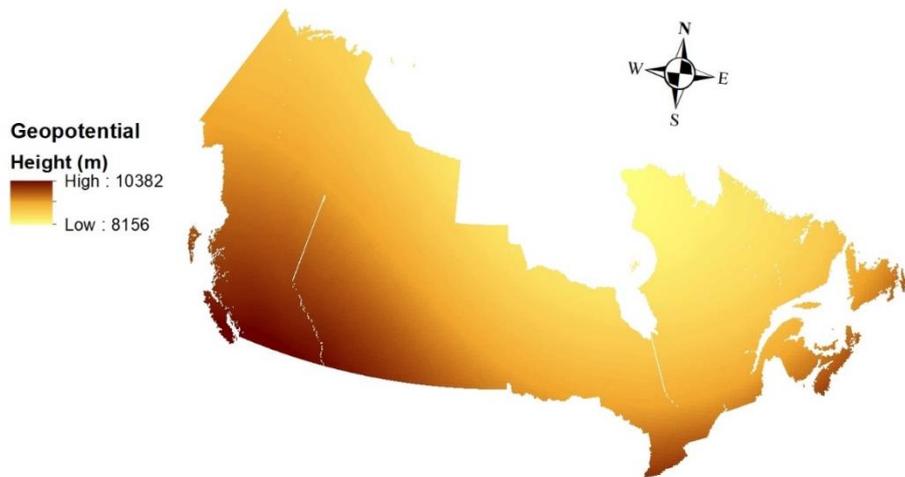
Figure 28. Geopotential height

4.4.1. Results of Bias Correction at Gauged Pixels

The first result in this part is the comparison of time series between ground observed, original (hereafter Ori), and bias-corrected (hereafter BC) IMERG precipitation at daily scale averaged over each cluster of the study area separately. It is worth noting that, this part presents the evaluation results of the bias-adjusted IMERG PrCal product over the gauged pixels by using the RQM model proposed in this study. Then, in the next part, the evaluation results of BC data over some randomly samples ungauged pixels obtained by the interpolation technique of parameters extracted from the RQM developed model at gauged pixels in the previous step are depicted.

4.4.1.1. Time Series Evaluation

The time series here are presented at both mean daily and mean monthly scales over five years from 2014 to 2018 at gauged pixels. In Figure 29 while the original IMERG indicates overestimates/underestimates, the bias-corrected data shows better performance over all months. However, at some clusters (e.g. no.5 and 6 (Figures 29(e) and (f)), BC tends to underestimate the ground observation especially during cold months (November of April). To have a clear comparison, we aggregated the daily precipitation to monthly plotted in Figure 30.
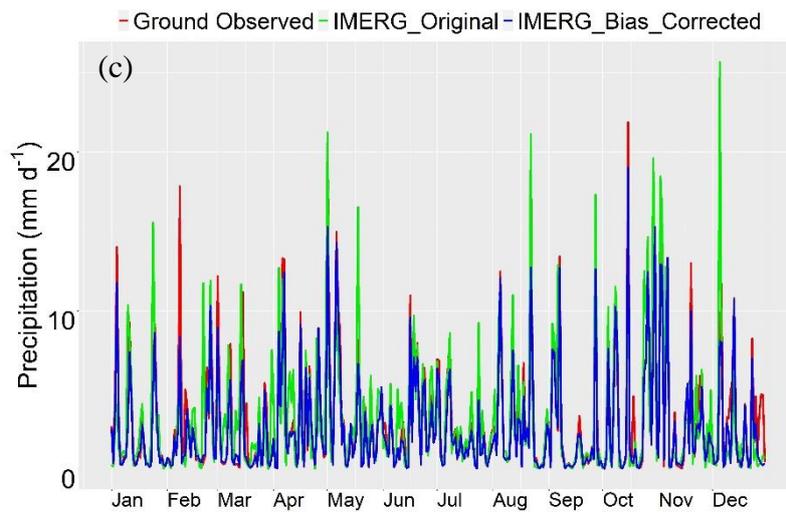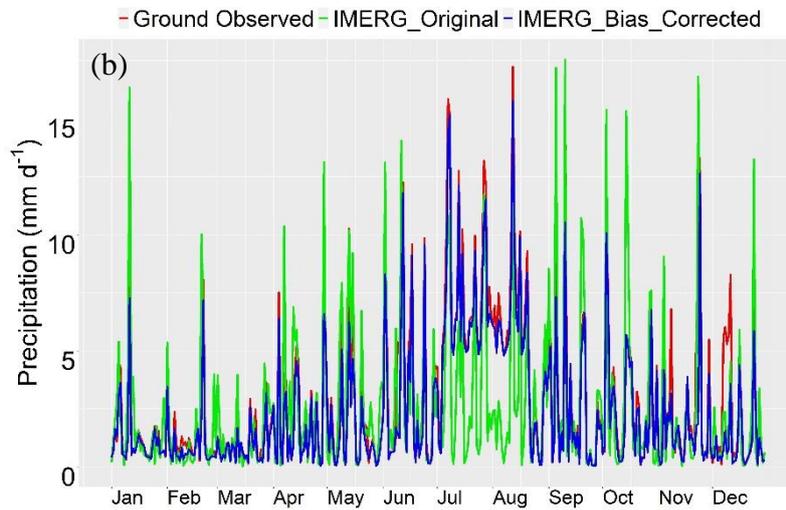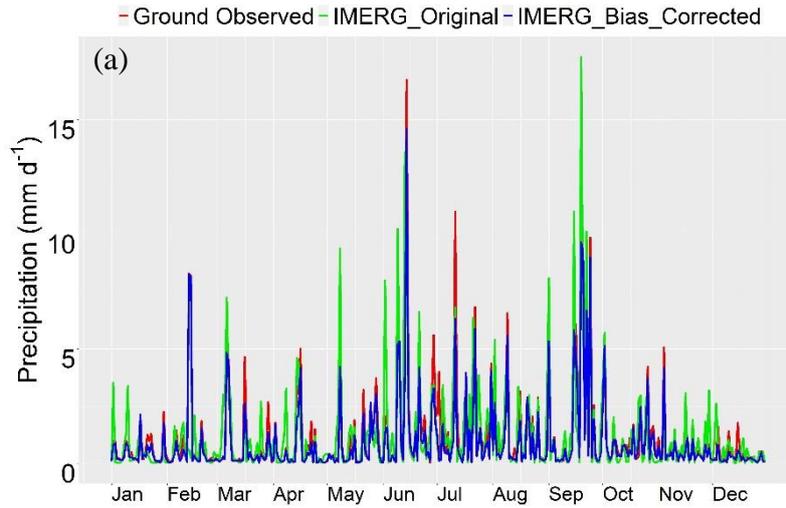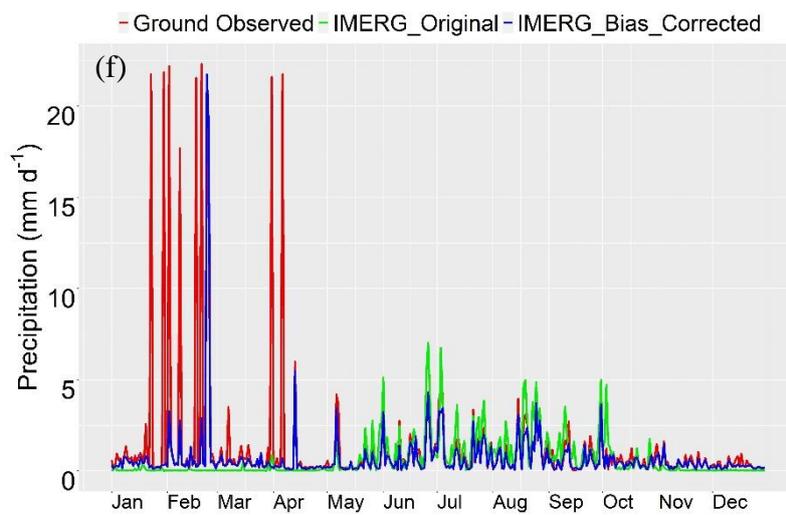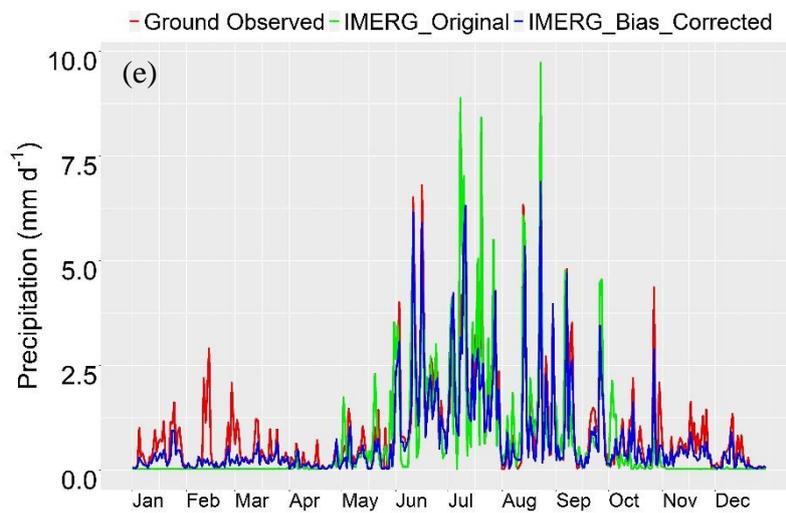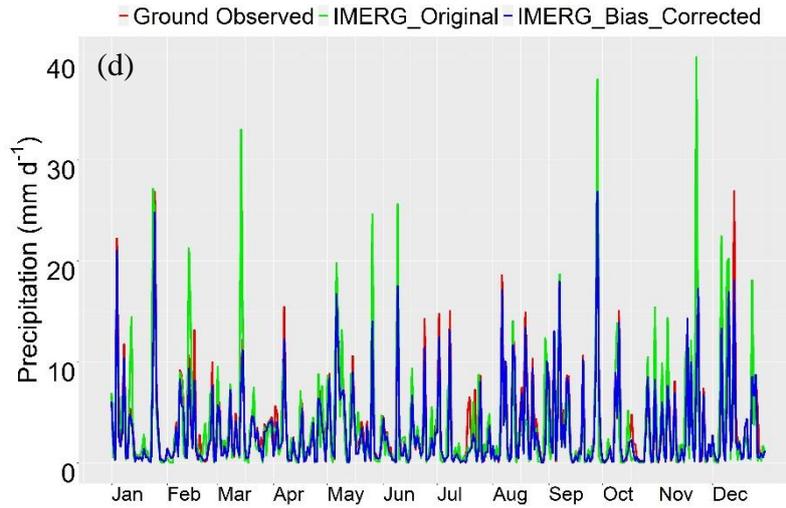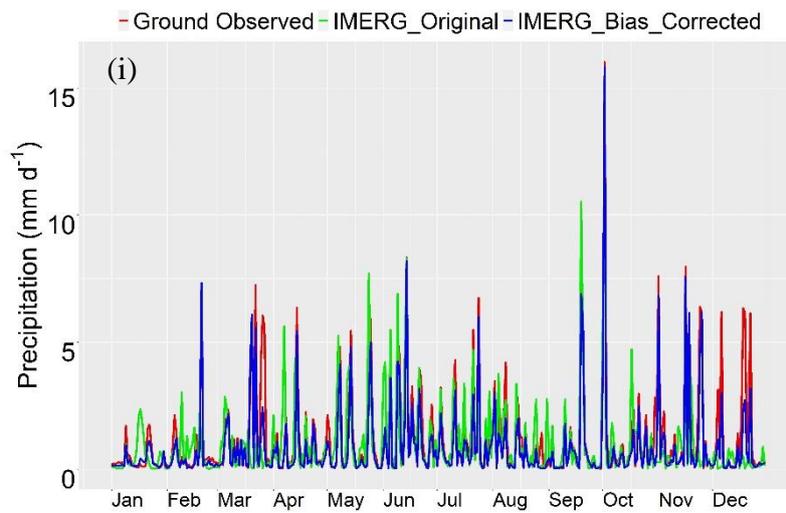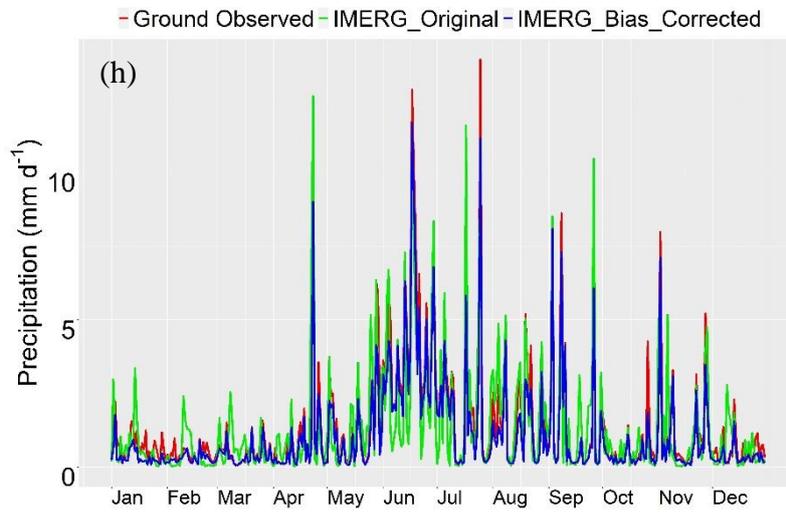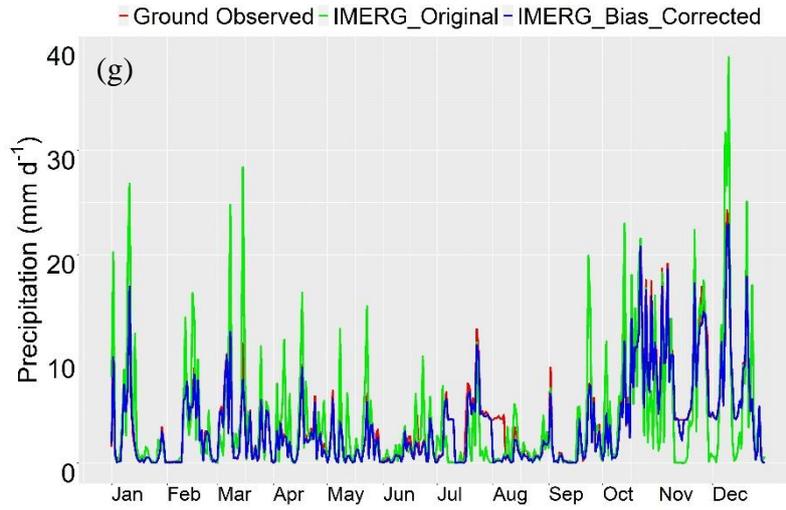
Figure 29. Spatial average of daily precipitation series of gauge observation, IMERG-Ori, and IMERG-BC at gauged pixels over the nine evaluation clusters ((a) represents cluster no.1 and (i) represents cluster no.9)

As seen in Figure 30, at all clusters BC data can reasonably capture the monthly variation of gauge-observed precipitation. Although the improvement of BC to Ori data is obvious, it still displays underestimation. In the following sections by using more statistical metrics the differences between original and bias-corrected data are explained and discussed.
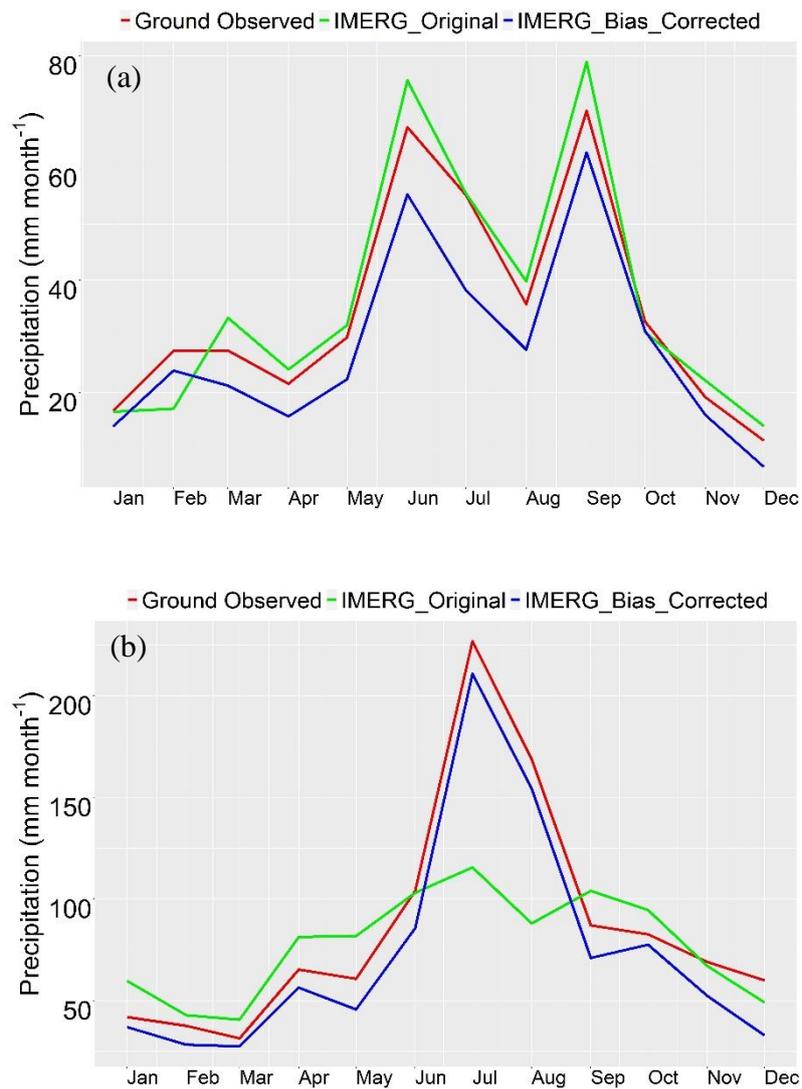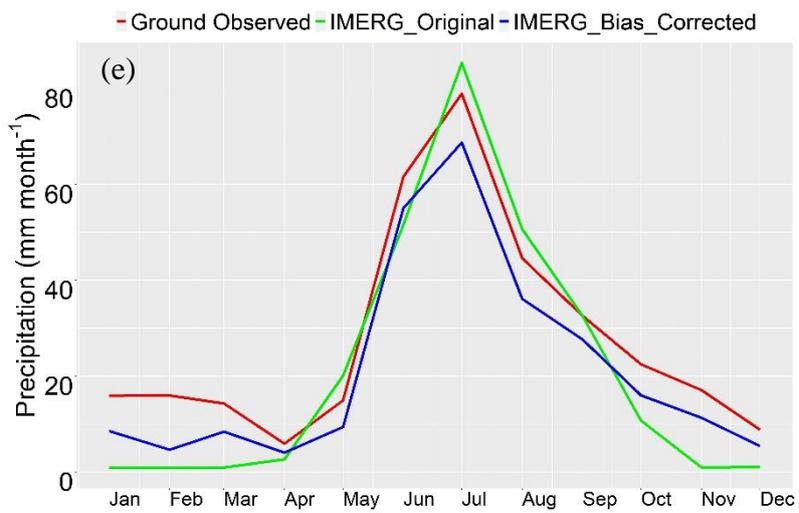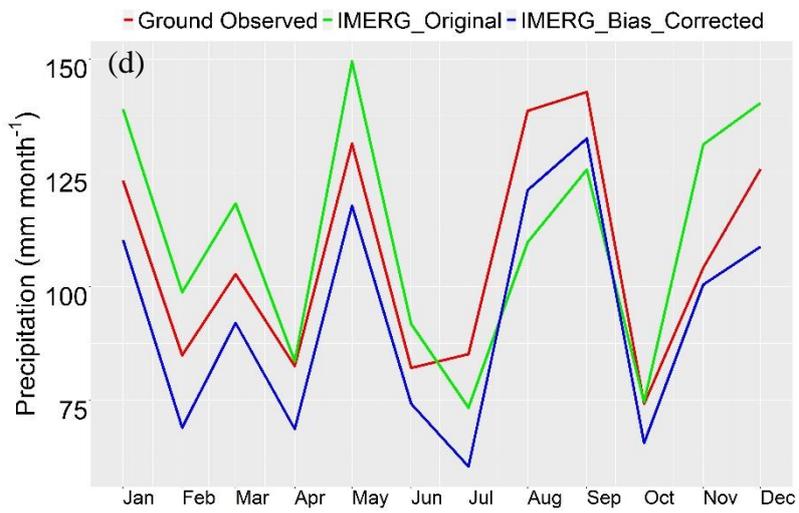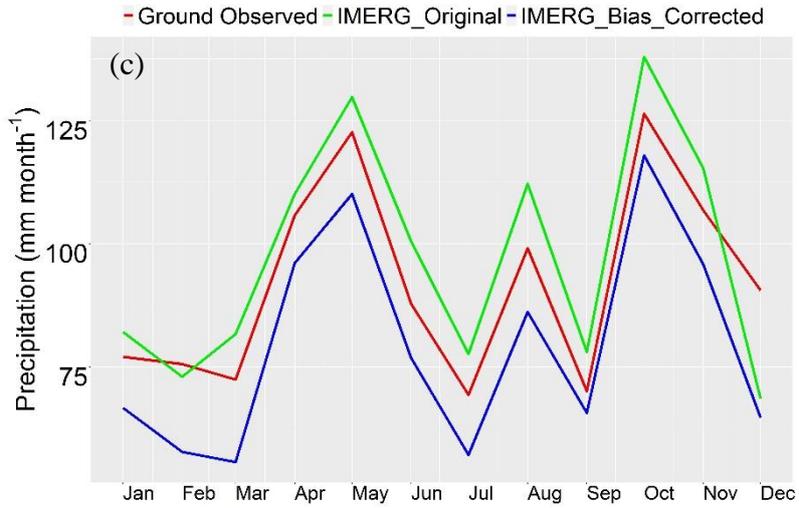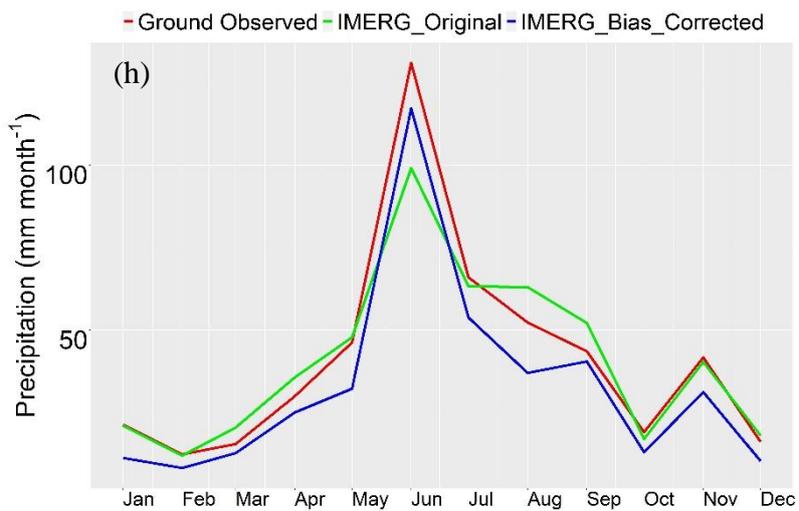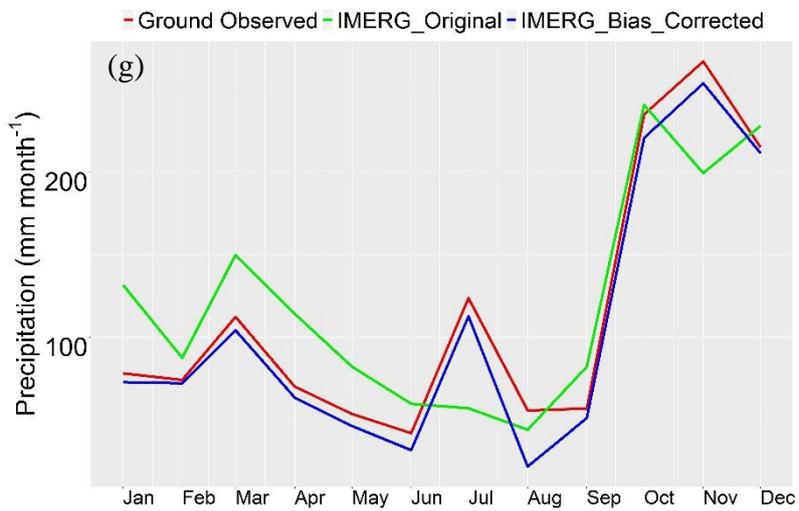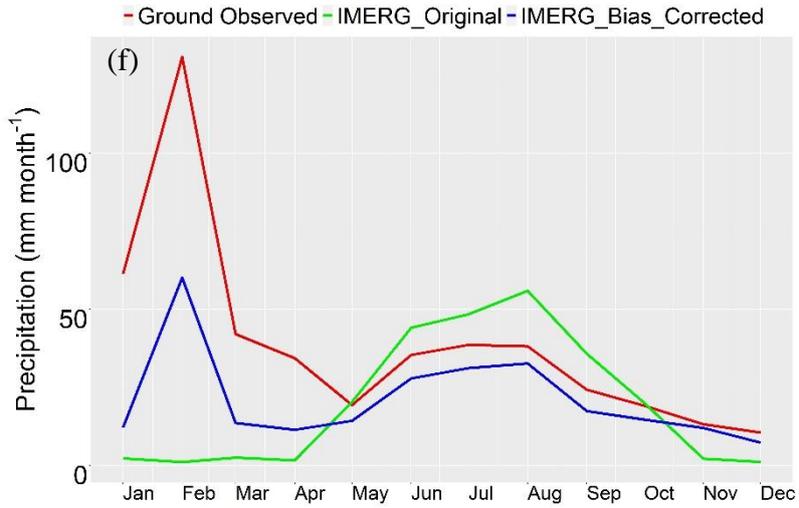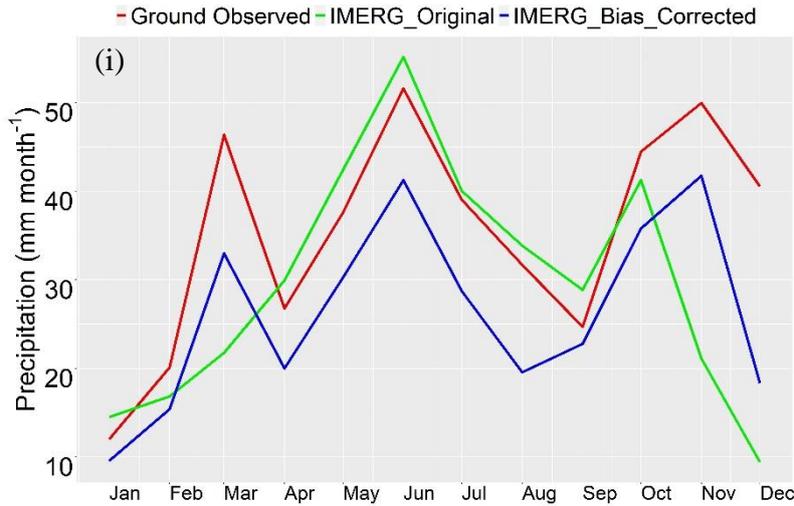
Figure 30. Spatial average of monthly precipitation series of gauge observation, IMERG-Ori, and IMERG-BC at gauged pixels over the nine evaluation clusters ((a) represents cluster no.1 and (i) represents cluster no.9)

- **Statistical Evaluation**

For analyzing the results in more detail, Figures 31-33 represent the evaluation metrics of CC, RBias, and RMSE respectively and for each cluster at each month separately. These metrics are calculated based on the equations provided in Table 2 of Chapter 3. According to the obtained metrics, the performance of the bias correction method is significant for cluster no.1 with improving the CC value from 0.49 to 0.85 and RMSE from 2.9 to 1.5 mm/d (average value of all months). This is a more promising finding as in the cold months where the performance of IMERG is weak with lower CC and higher biases, the BC data can be used instead of Ori biased data. An appropriate BC results over cluster no.2 (east southern part of the country) are produced by the proposed bias correction method. In this area after bias correction, the CC value is improved from 0.55 to 0.88 and RMSE from 8.2 to 4.6 mm/d. For other clusters, the same metrics, as well as RBias, are listed in Table 6. In this table, all metrics have been provided based on the daily mean value of all gauged pixels located within the given cluster for each month separately. We stated in the previous chapter that raw IMERG product has more problems in winter with the lowest precision, as also can see here over the all clusters, the lower performance of IMERG happens during the cold months. Therefore, it provides more room for bias correcting that leads to the largest improvements. For example, the average value of CC over the clusters during the cold months from November to March is improved overall by 267% (from 0.3 to 0.8), however, during warm months from April to October the improvement ranges by 150% (from 0.6 to 0.9). As in the proposed method of bias correction

in this study, the different levels of precipitation based on CDF segmentation are considered and corrected separately, the rate of precipitation doesn't have a considerable effect on the correction. In terms of RBias, as listed in Table 6, over clusters no.2, 5, 6, and 8, the mean value of all months indicates the better performance of Ori to BC data, however, considering Figure 32, the depicted boxplots clarify that the reason is due to the variation of RBias values from negative to positive ranges which generally leads to the less mean value.
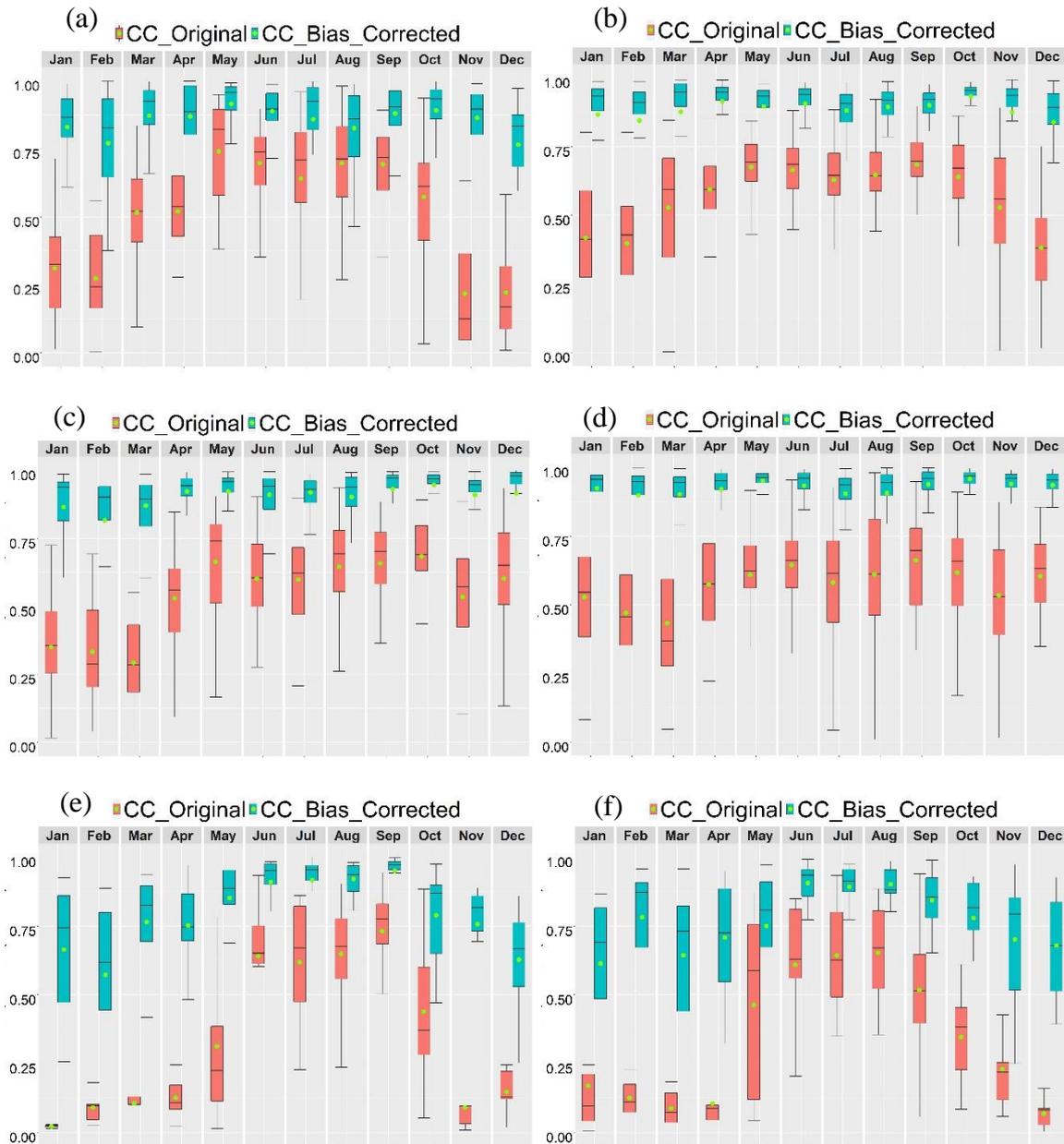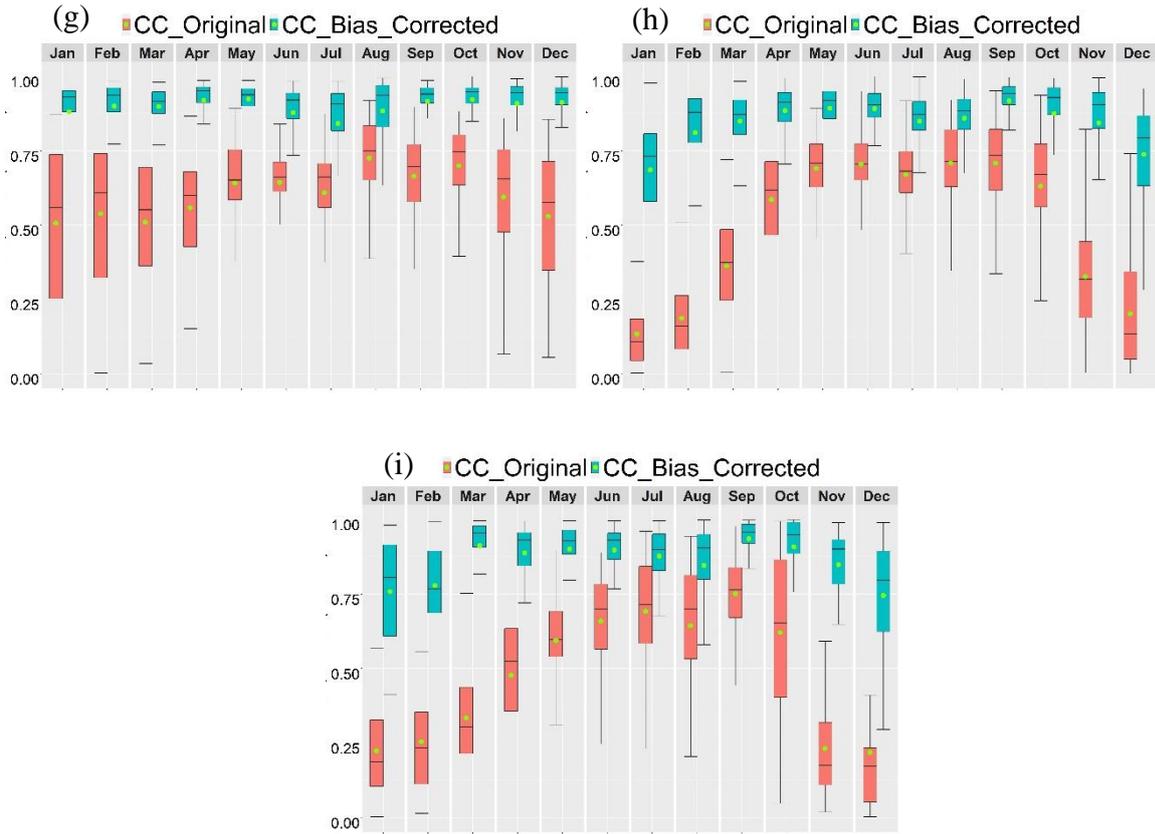
Figure 31. CC (correlation) between spatially averaged daily of IMERG-Ori and IMERG-BC against gauge observation at gauged pixels over the nine evaluation clusters ((a) represents cluster no.1 and (i) represents cluster no.9) (green dots in boxes indicate the median value)
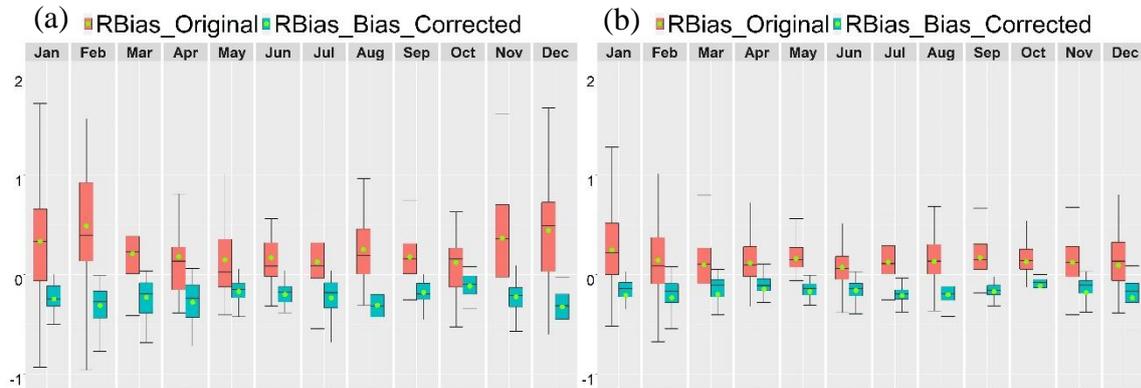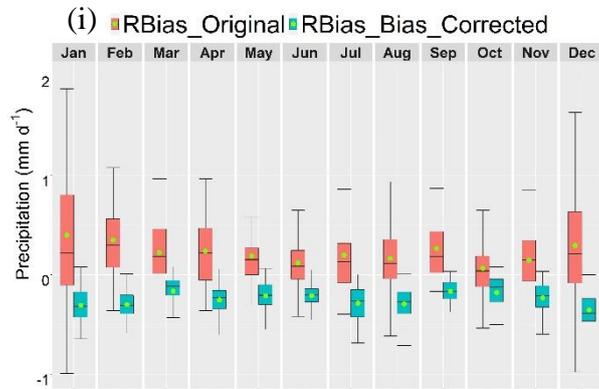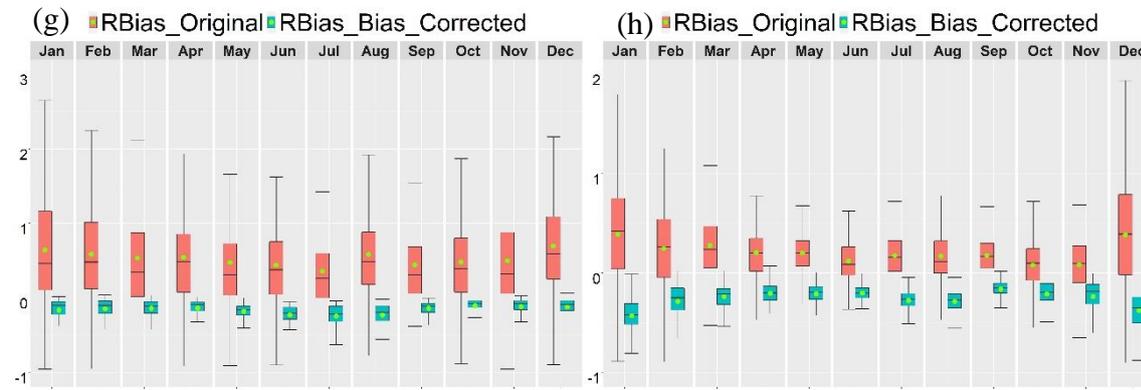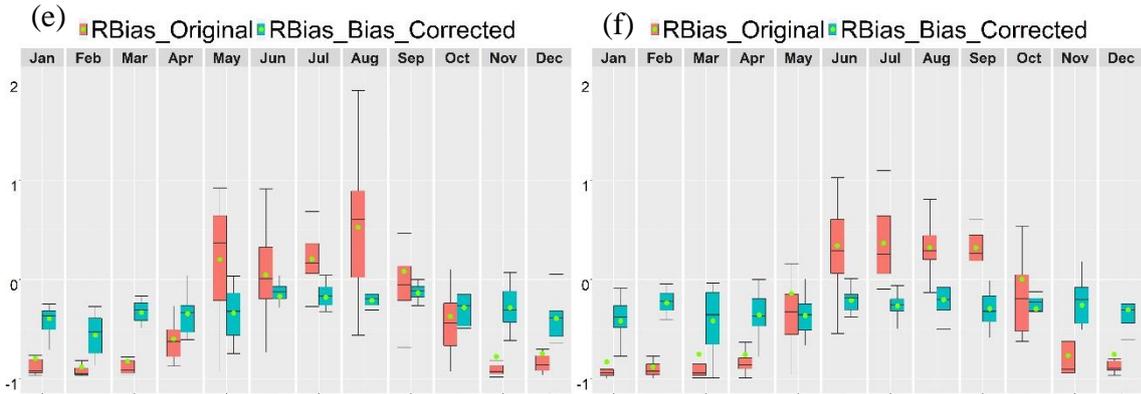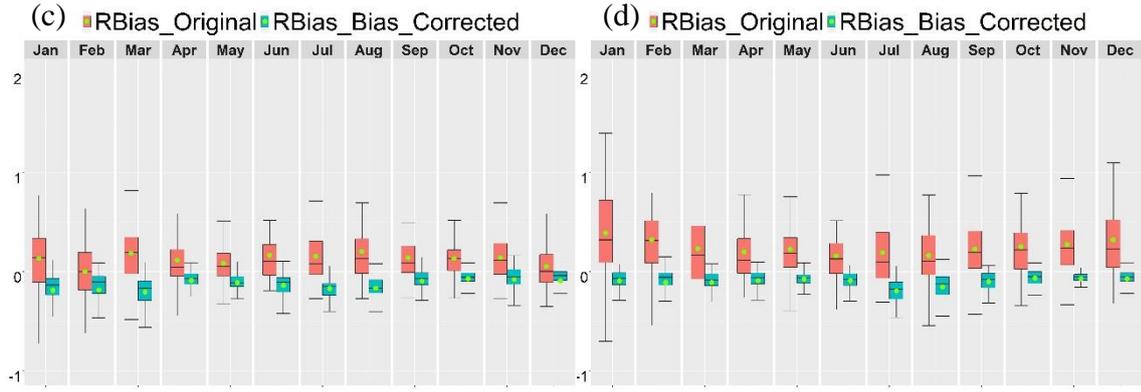
Figure 32. Spatially averaged daily RBias of IMERG-Ori and IMERG-BC against gauge observation at gauged pixels over the nine evaluation clusters ((a) represents cluster no.1 and (i) represents cluster no.9) (green dots in boxes indicate the median value)
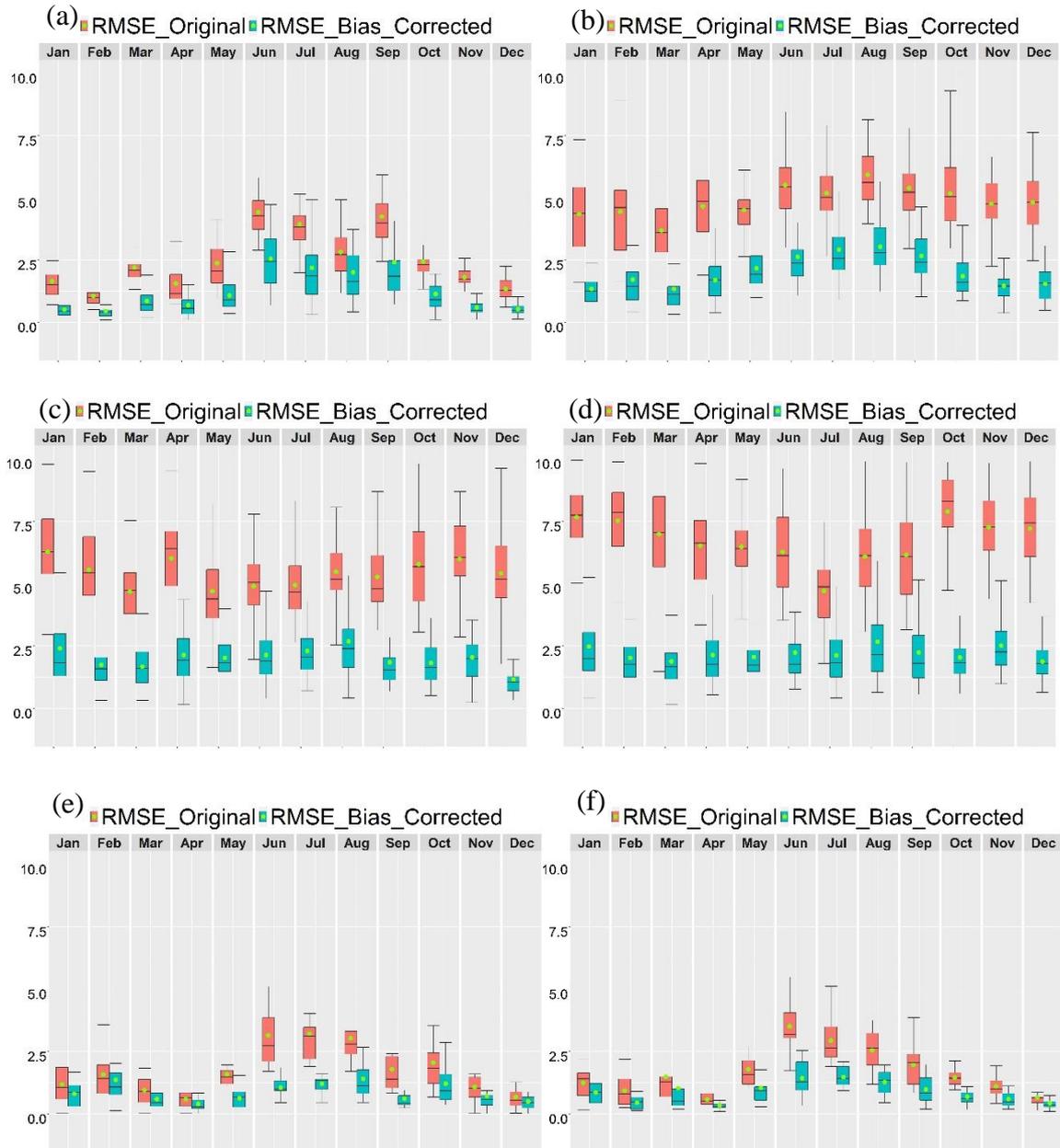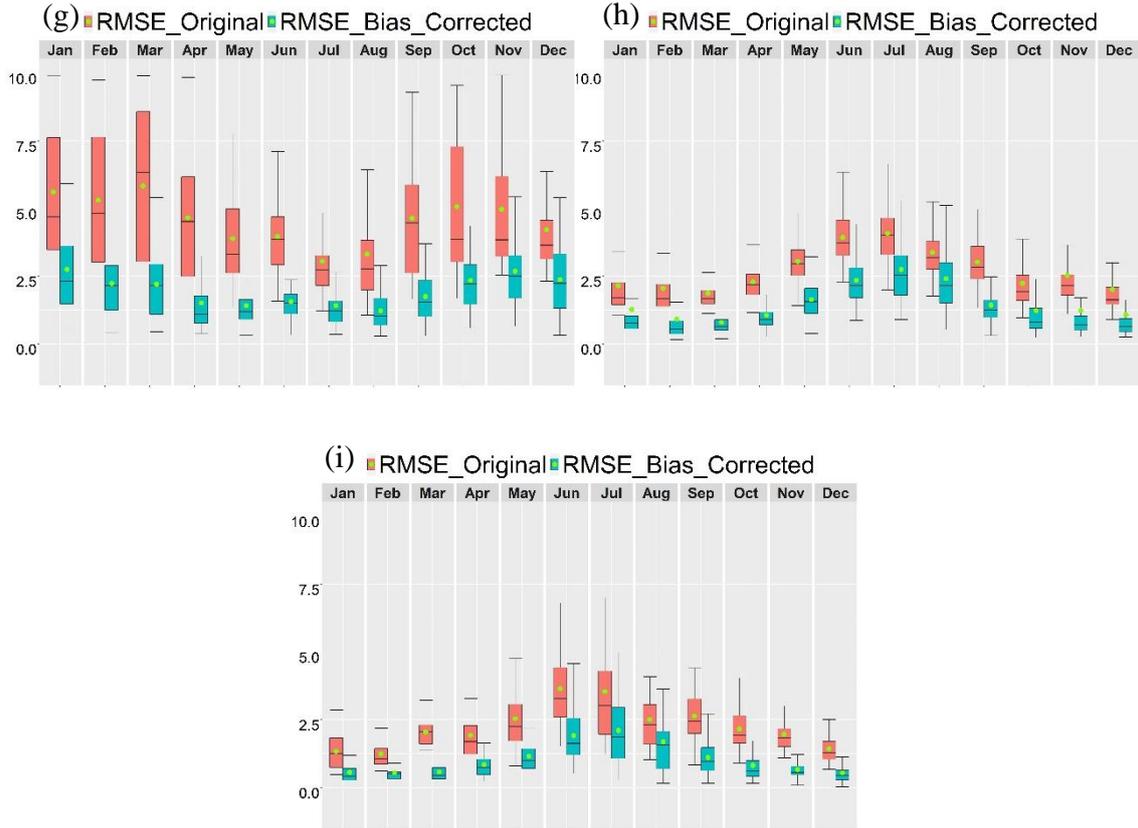
Figure 33. Spatially averaged daily RMSE of IMERG-Ori and IMERG-BC against gauge observation at gauged pixels over the nine evaluation clusters ((a) represents cluster no.1 and (i) represents cluster no.9) (green dots in boxes indicate the median value)

Table 6. Mean statistics value at gauged pixels over each cluster

Time	Statistics	Cluster 1		Cluster 2		Cluster 3		Cluster 4		Cluster 5		Cluster 6		Cluster 7		Cluster 8		Cluster 9	
		Ori	BC	Ori	BC	Ori	BC	Ori	BC	Ori	BC	Ori	BC	Ori	BC	Ori	BC	Ori	BC
	CC	0.28	0.83	0.4	0.87	0.34	0.87	0.53	0.92	0.07	0.66	0.03	0.61	0.48	0.88	0.1	0.68	0.15	0.76
January	RMSE	3.4	1.84	8.36	4.6	9.2	4.6	10.4	3.23	1.18	0.8	5	4.6	12.08	4.09	2.9	1.85	3.73	1.83
	RBias	0.49	-0.24	0.37	-0.2	0.73	-0.18	0.39	-0.1	-0.8	-0.4	-0.83	-0.42	0.77	-0.16	0.43	-0.43	0.7	-0.3
	CC	0.25	0.77	0.36	0.84	0.28	0.82	0.9	0.45	0.03	0.57	0.07	0.78	0.52	0.9	0.13	0.81	0.21	0.78
February	RMSE	2.55	0.45	10.4	5.55	9.4	5.3	9.96	3.17	1.57	1.36	7.29	5.46	10.56	3.17	2.7	1.53	3.96	1.49
	RBias	0.72	-0.31	0.14	-0.23	0.1	-0.18	0.35	-0.1	-0.88	-0.56	-0.88	-0.24	0.75	-0.14	0.25	-0.28	0.35	-0.29
	CC	0.5	0.87	0.5	0.88	0.28	0.87	0.42	0.9	0.003	0.76	0.02	0.64	0.5	0.89	0.35	0.85	0.31	0.91
March	RMSE	2.2	0.85	9.02	6	5.25	2.25	8.4	2.77	0.95	0.59	4.35	3.9	8.5	2.7	2.24	1.08	4.5	2.02
	RBias	0.35	-0.23	0.1	-0.2	0.22	-0.2	0.23	-0.1	-0.82	-0.33	-0.75	-0.42	0.58	-0.14	0.29	-0.23	0.37	-0.16
	CC	0.5	0.87	0.57	0.92	0.52	0.92	0.56	0.92	0.015	0.75	0.02	0.7	0.55	0.92	0.58	0.88	0.48	0.89
April	RMSE	1.55	0.68	7.34	4.31	6.15	2.13	8.1	3.3	0.59	0.39	3.4	3.02	7	2.59	2.53	1.3	1.92	0.84
	RBias	0.18	-0.27	0.11	-0.14	0.12	-0.09	0.2	-0.1	-0.6	0.34	-0.75	-0.35	0.54	-0.13	0.2	-0.2	0.24	-0.25
	CC	0.74	0.91	0.66	0.9	0.65	0.92	0.6	0.95	0.31	0.85	0.35	0.75	0.64	0.92	0.69	0.89	0.59	0.9
May	RMSE	2.37	1.08	6	3.62	5.1	2.18	7.18	2.05	1.59	0.62	1.78	1.04	5.4	1.8	3.27	1.87	2.54	1.15
	RBias	0.2	-0.17	0.16	-0.17	0.09	-0.11	0.22	-0.07	1.98	-0.33	0.04	-0.36	0.47	-0.18	0.2	-0.2	0.19	-0.2
	CC	0.7	0.89	0.66	0.9	0.6	0.91	0.64	0.93	0.53	0.91	0.61	0.9	0.63	0.88	0.7	0.89	0.66	0.9
Jun	RMSE	4.7	2.55	7.03	4	5.15	2.15	6.7	2.5	3.15	1.05	3.52	1.42	5.83	2.65	4.6	2.35	3.65	1.92
	RBias	0.17	-0.19	0.07	-0.16	0.17	-0.13	0.16	-0.1	2.5	-0.17	0.46	-0.21	0.44	-0.22	0.12	-0.2	0.12	0.21
	CC	0.64	0.86	0.63	0.88	0.6	0.92	0.58	0.9	0.62	0.92	0.64	0.9	0.61	0.84	0.67	0.85	0.69	0.88
July	RMSE	5.3	3.55	8.1	4	4.94	2.3	6.1	3.04	3.21	1.21	2.9	1.47	6.4	3.04	4.47	2.8	3.54	2.1
	RBias	0.13	-0.23	0.125	-0.21	0.16	-0.17	0.19	-0.19	0.45	-0.18	0.37	-0.26	0.4	-0.24	0.17	-0.27	0.2	-0.29
	CC	0.7	0.83	0.64	0.9	0.65	0.9	0.6	0.91	0.65	0.92	0.65	0.9	0.71	0.88	0.7	0.86	0.64	0.85
August	RMSE	3.15	2	8.3	4.2	6.23	3.05	9	4	3.03	1.4	2.54	1.26	5.17	2.89	3.56	2.6	2.5	1.69
	RBias	0.26	-0.3	0.13	-0.2	0.26	-0.17	0.16	-0.15	0.53	-0.21	0.32	-0.2	0.65	-0.23	0.17	-0.28	0.17	-0.29
	CC	0.7	0.88	0.68	0.9	0.66	0.93	0.66	0.94	0.73	0.95	0.52	0.84	0.66	0.92	0.7	0.92	0.75	0.94
September	RMSE	4.4	2.41	6.17	3.4	5.9	2.58	9.5	2.5	1.8	0.61	1.96	0.98	7.04	2.83	3.37	1.73	2.63	1.1
	RBias	0.18	-0.18	0.17	-0.17	0.17	-0.09	0.23	-0.1	0.08	-0.14	0.46	-0.3	0.48	-0.13	0.17	-0.16	0.26	-0.16
	CC	0.57	0.89	0.63	0.93	0.68	0.95	0.61	0.96	0.23	0.79	0.29	0.78	0.7	0.92	0.63	0.87	0.61	0.91
October	RMSE	2.43	1.13	8.6	4.03	6.46	2.32	9.9	2.4	2.05	1.21	1.45	0.68	10.67	3.15	2.63	1.58	4.07	1.94
	RBias	0.23	-0.1	0.13	-0.2	0.17	-0.07	0.38	-0.06	-0.37	-0.28	0.003	-0.3	0.48	-0.1	0.08	-0.2	0.1	-0.18
	CC	0.17	0.86	0.52	0.87	0.52	0.91	0.53	0.94	0.007	0.75	0.16	0.7	0.58	0.91	0.32	0.84	0.17	0.85
November	RMSE	1.8	0.59	9.9	5.85	6.83	2.63	10.3	2.5	1.08	0.71	4.9	4.4	13.41	3.62	3.05	1.65	5.6	1.64
	RBias	0.52	-0.22	0.13	-0.11	0.19	-0.08	0.27	-0.07	-0.78	-0.28	-0.77	-0.26	0.5	-0.1	0.1	-0.23	0.34	-0.23
	CC	0.19	0.76	0.36	0.84	0.59	0.91	0.6	0.94	-0.02	0.63	0.015	0.68	0.51	0.91	0.12	0.74	0.07	0.74
December	RMSE	1.34	0.53	9.6	5.7	8.7	3.9	9.5	3.15	0.67	0.49	0.64	0.42	13.31	2.87	2.41	1.37	4.58	2.04
	RBias	0.44	-0.32	0.27	-0.23	0.085	-0.09	0.32	-0.07	-0.75	-0.4	-0.75	-0.3	0.93	-0.12	0.4	-0.38	0.76	-0.35

- **Spatial Pattern Evaluation**

The more visual comparison between BC and Ori-IMERG is demonstrated in Figures 34-37, where each figure compares the relevant metric at gauged pixels for original and bias-corrected IMERG during each season separately. The spatial variation distributed over the study area indicates significant improvements, especially in interior plains where Ori-IMERG is characterized by lower performance. Figure 34 shows the RMSE value over the study area. As seen here and also explained before, the bias correction method has potentially improved the original IMERG especially over the west and east coasts where there was more room for enhancement (the difference between Ori-IMERG and observed data is higher). Considering seasonally evaluation, one can observe that the pattern of bias-corrected performance follows the original data, while it displays better results. For example, both have higher/lower RMSE values during summer/winter.

In terms of CC depicted in Figure 35, the improvement is notable as over all sites it indicates values approximately more than 0.7. The main advantage of the BC-IMERG is its promising value during winter where Ori-IMERG expresses a very low correlation with the observed data.

In addition to the RMSE and CC, two other categorical metrics named POD and FAR (see Table 3 of Chapter 3) are provided in Figures 36 and 37, respectively. POD related to BC data with values more than 0.8 over the most pixels is in good agreement with the gauge-observed during all seasons, nevertheless, Ori data as discussed further in Chapter 3 performs unreliable results in detecting the precipitation events, particularly within the cold seasons. Regarding FAR, again, we can see reasonable values (less than 0.2) associated with BC data for all seasons, while for Ori data it is more than 0.4 that indicates large errors in detecting precipitation events falsely.

Overall, the bias correction technique used in this study not only able to improve the estimates of precipitation intensity but can also detect the precipitation occurrence over the study area remarkably.
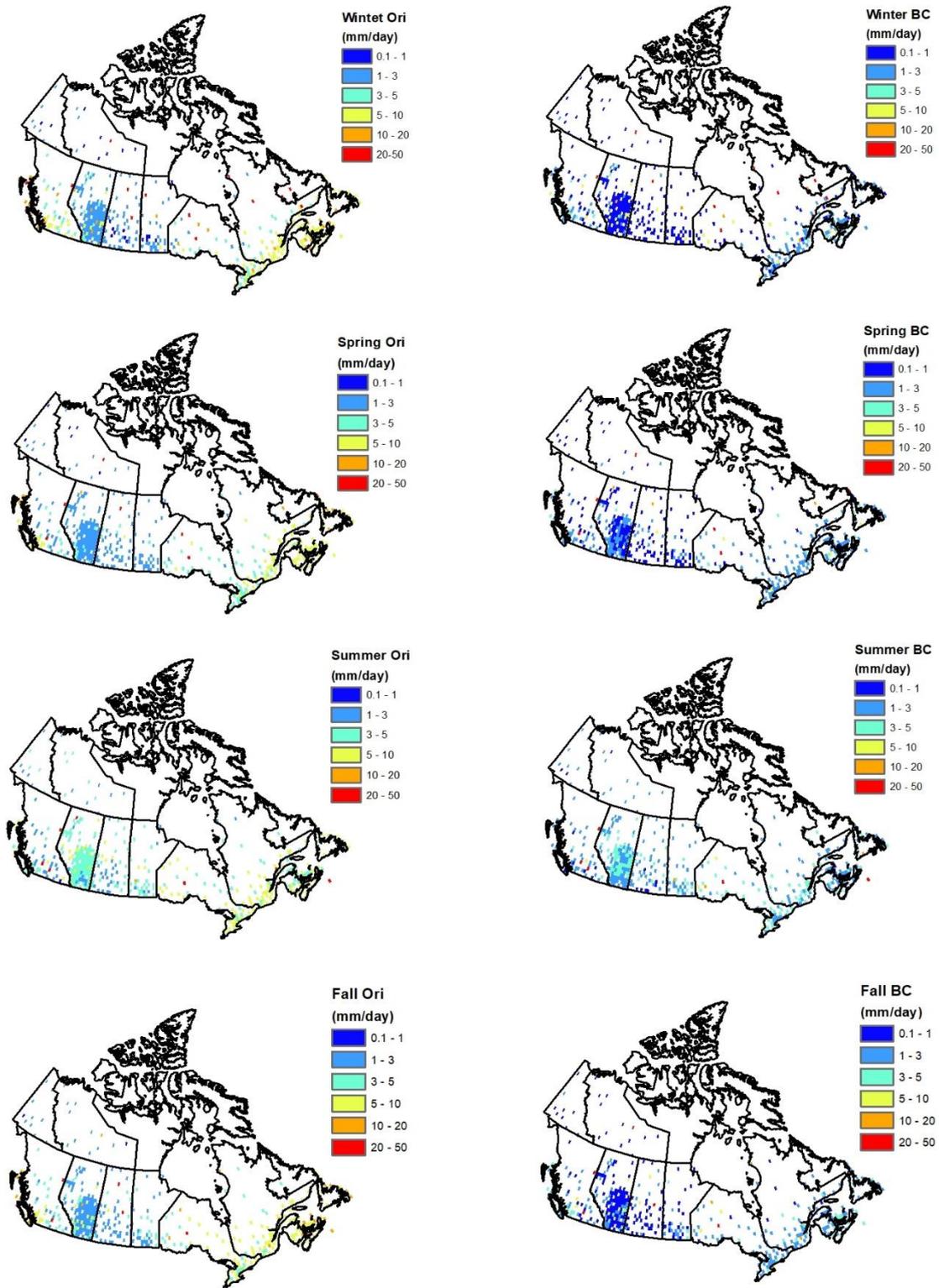
Figure 34. The spatial pattern of RMSE for IMERG-Ori (Ori, left) and IMERG-BC (BC, right) against gauge observation at gauged pixels over the study area based on different seasons
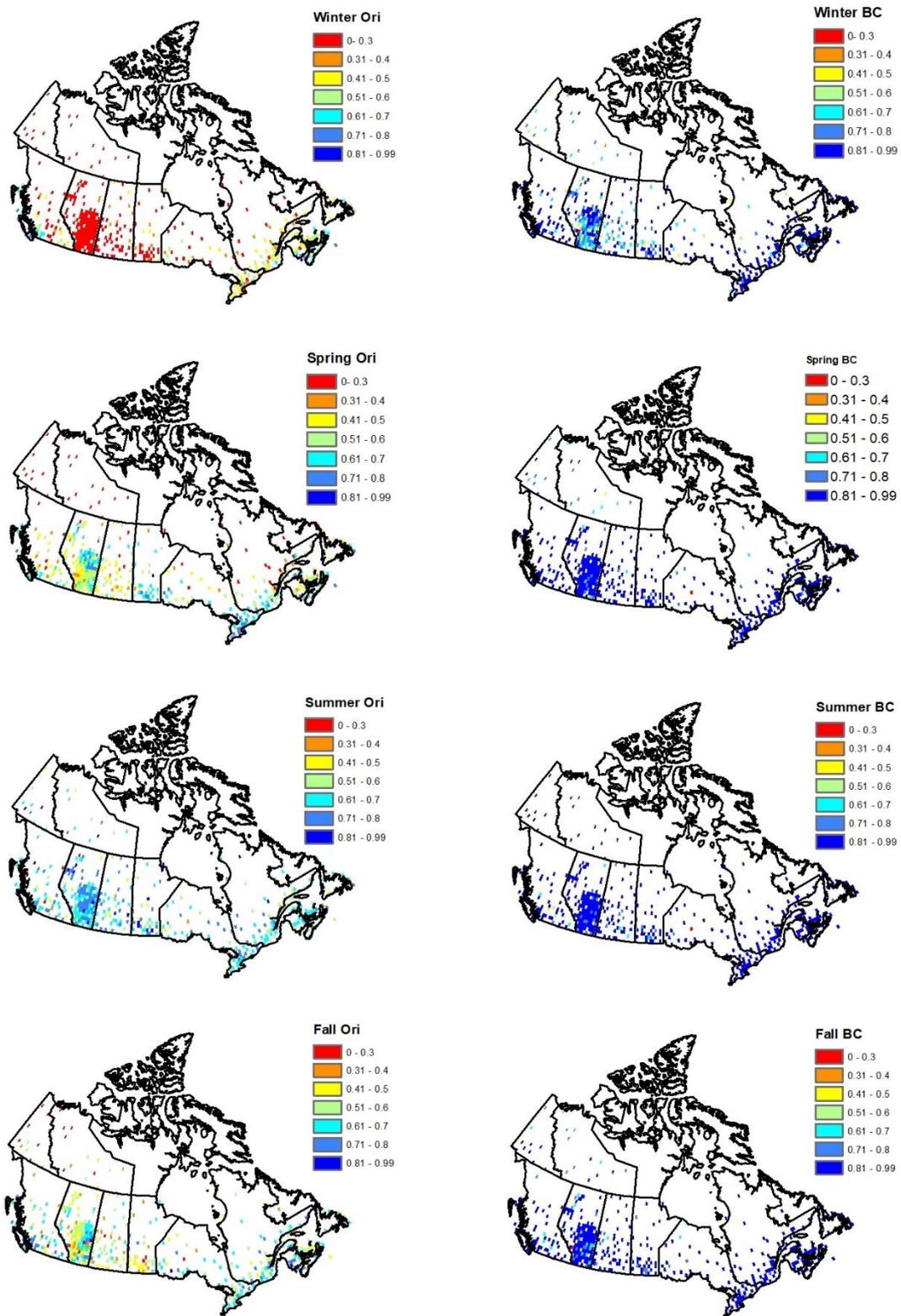
Figure 35. The spatial pattern of CC for IMERG-Ori (Ori, left) and IMERG-BC (BC, right) against gauge observation at gauged pixels over the study area based on different seasons
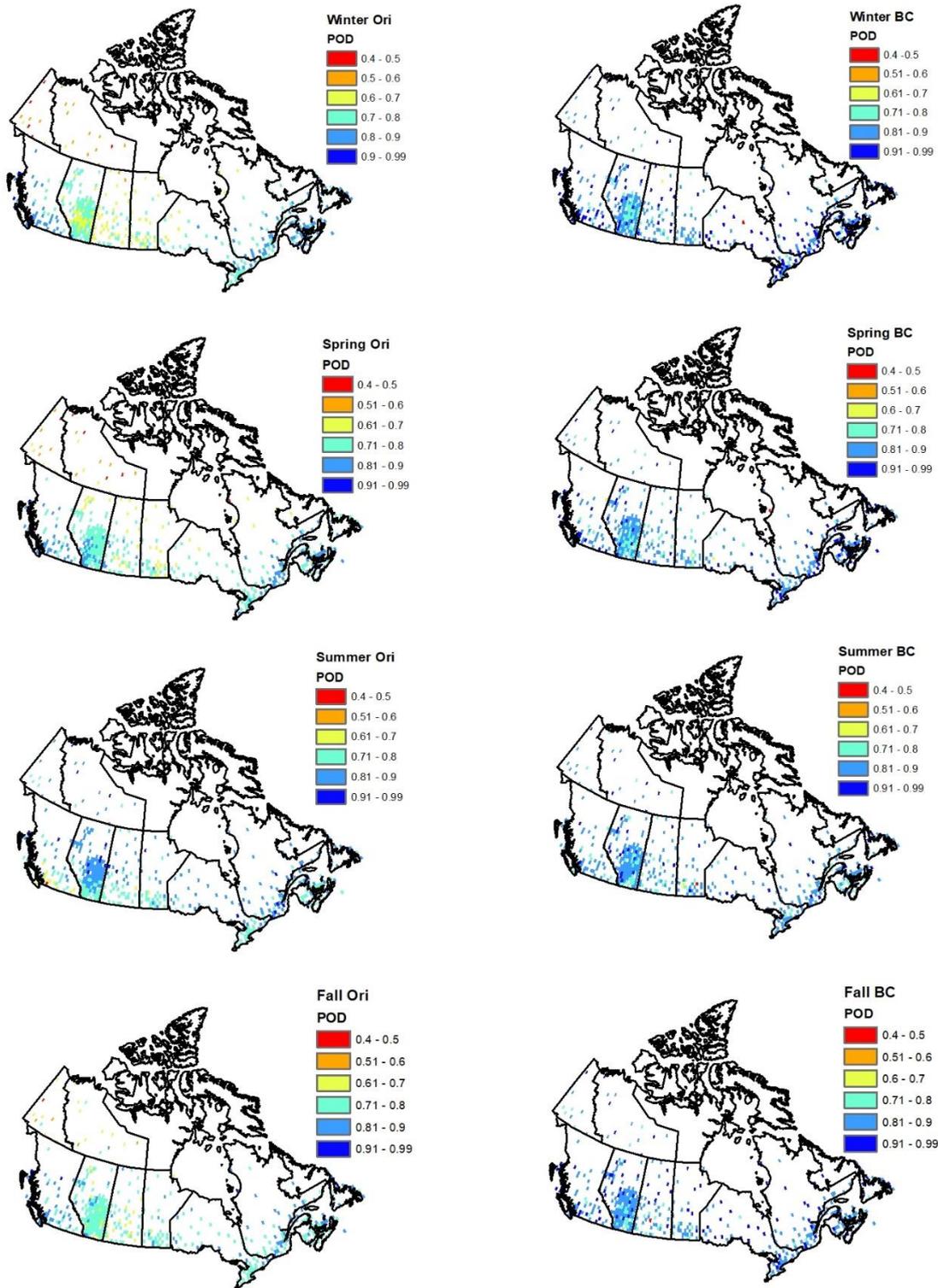
Figure 36. The spatial pattern of POD for IMERG-Ori (Ori, left) and IMERG-BC (BC, right) against gauge observation at gauged pixels over the study area based on different seasons
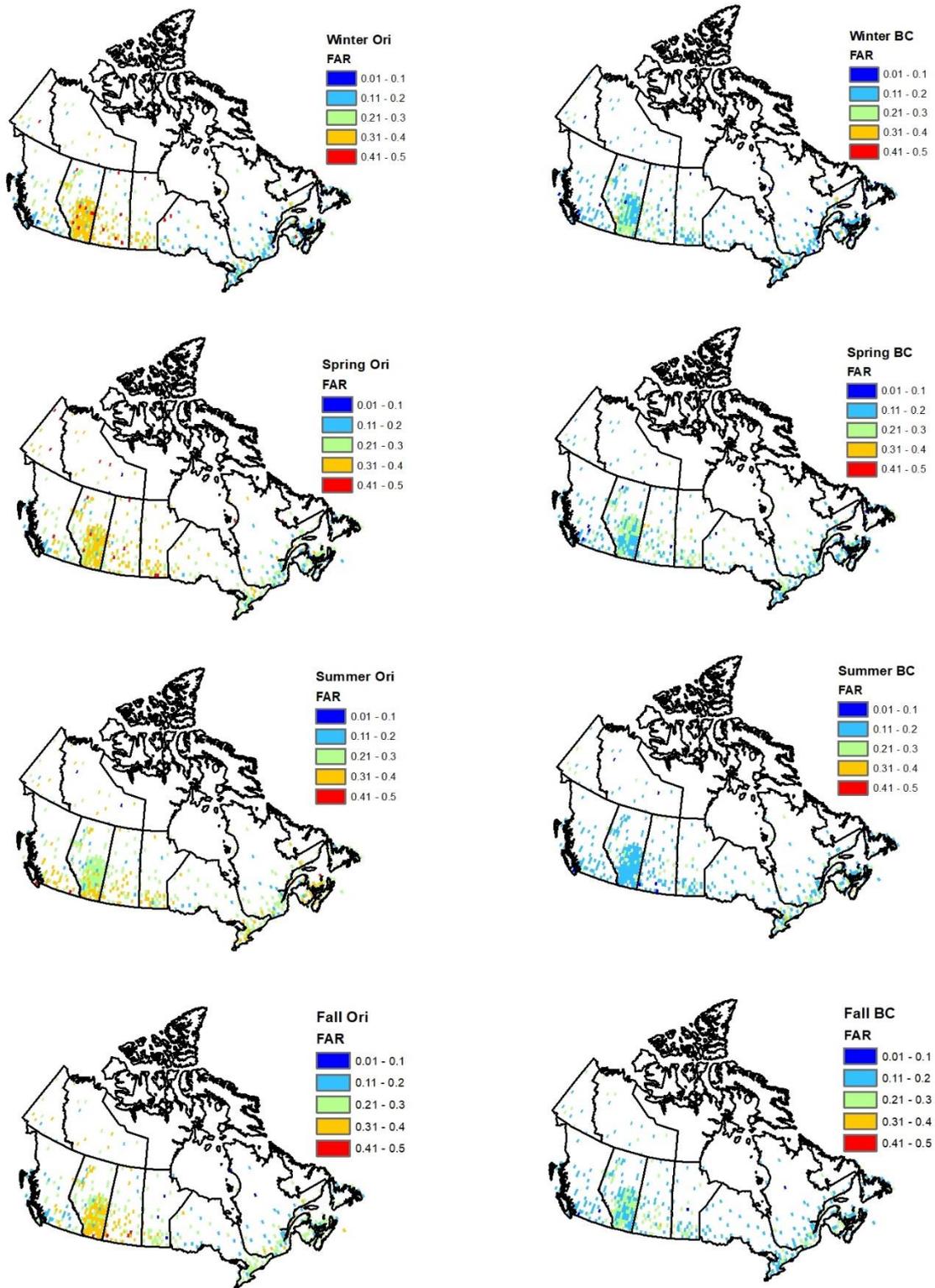
Figure 37. The spatial pattern of FAR for IMERG-Ori (Ori, left) and IMERG-BC (BC, right) against gauge observation at gauged pixels over the study area based on different seasons

4.4.2. Results of Bias Correction at Ungauged Pixels

In this part, the performance of the bias-corrected IMERG product over the ungauged pixels by using the obtained parameters from the developed regression model over those gauged pixels is represented. The approach for comparing the bias-corrected with the raw data is based on the cross-validation by sampling some gauged pixels (10%) at each cluster randomly using the bootstrapping technique for 1000 times and considering the selected pixels as ungauged. Then, the bias correction model is developed over that other 90% of gauged pixels and parameters are calculated and interpolated over 10% assumed ungauged pixels. During 1000 times of repeating this step, we can make sure that all gauged pixels participate in the validation process and the obtained results can cover the entire cluster zone spatially. The following results are the average value of that 10% assuming ungauged pixels (hereafter named validation pixels).

- **Time Series Evaluation**

Figure 38 displays the mean monthly time series of precipitation obtained over the validation pixels at each cluster separately. Although the plots show a reasonable capturing of observed precipitation trend by BC and better performance than Ori data over all clusters, it still is not in good agreement with the gauge-observed during some months. For having a more detailed evaluation, some statistical metrics are provided in Figures 39 to 42.
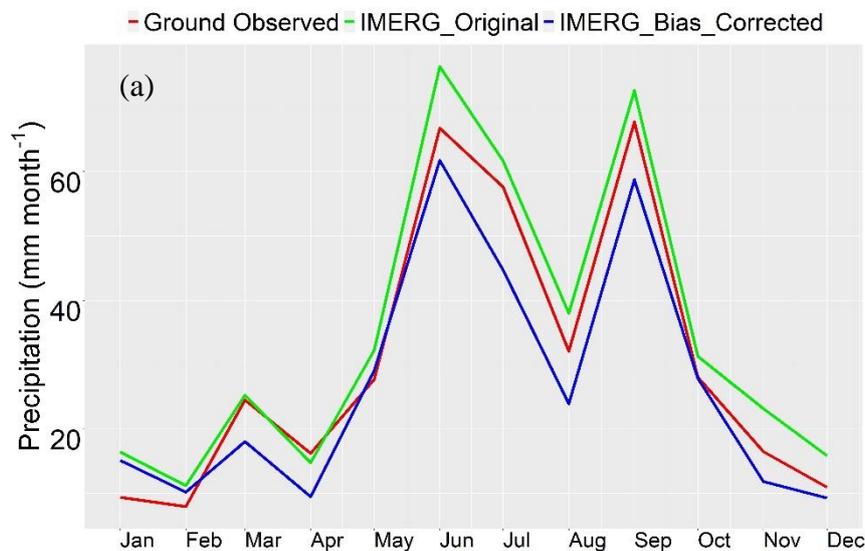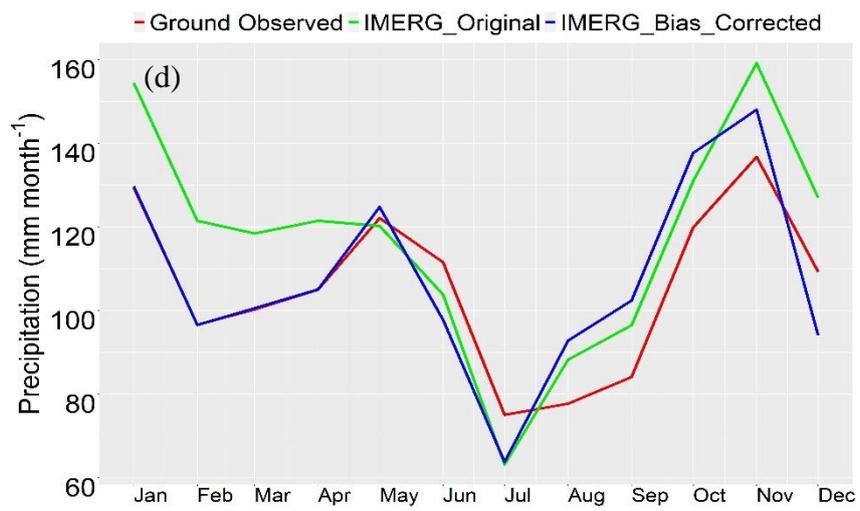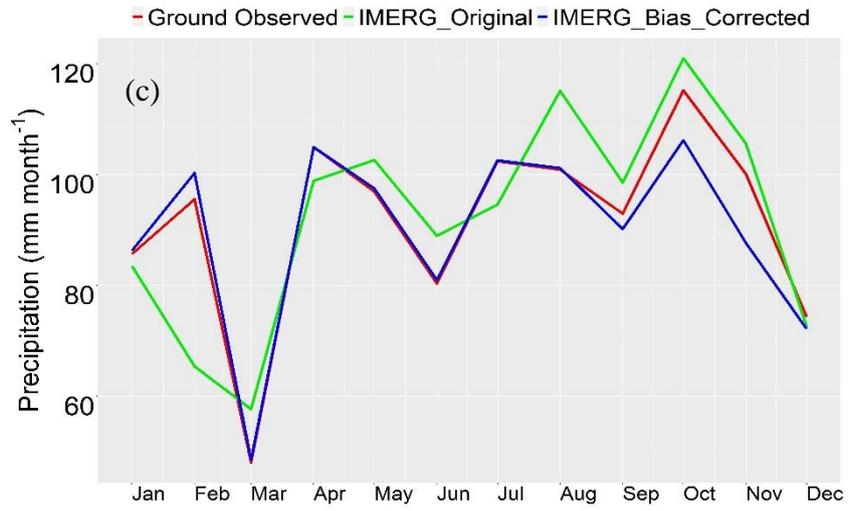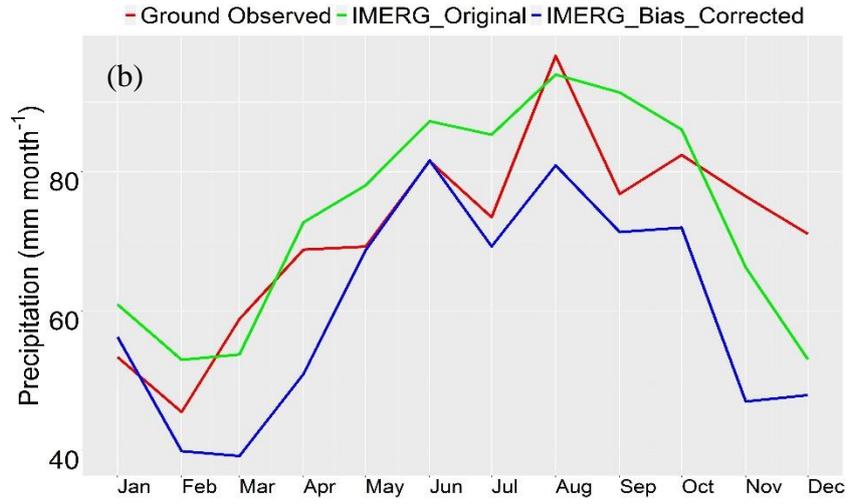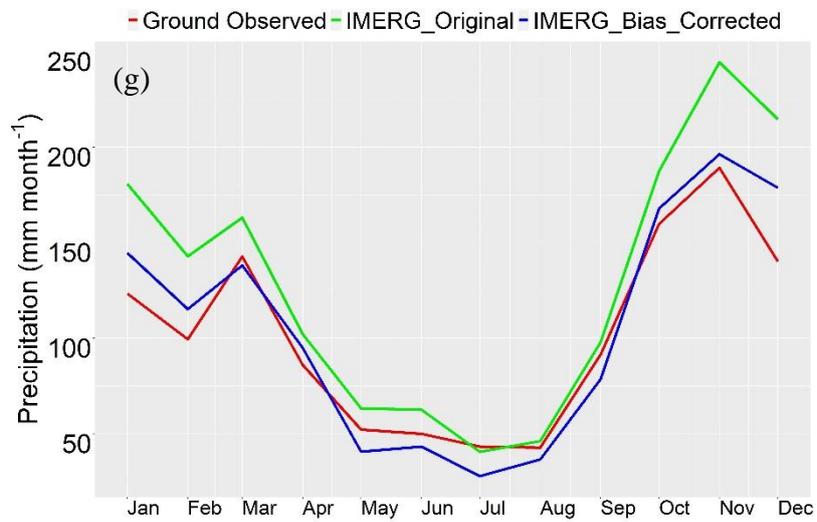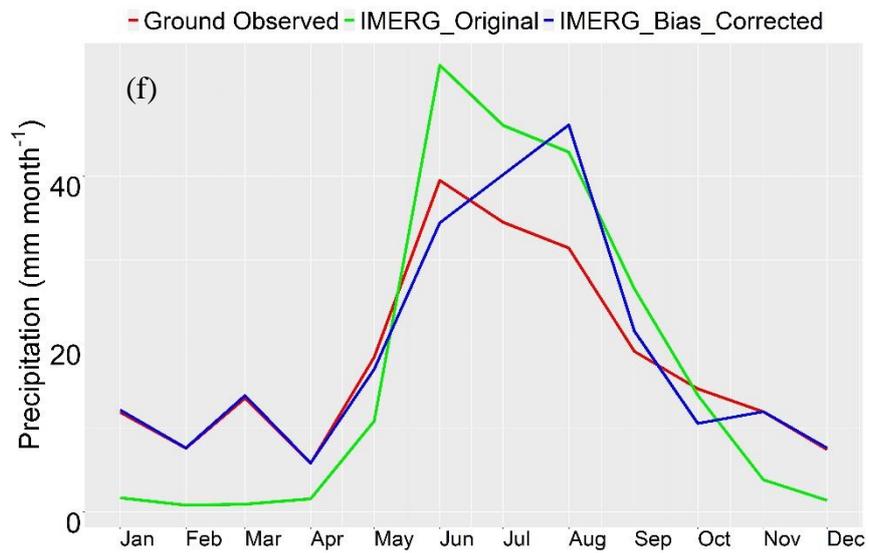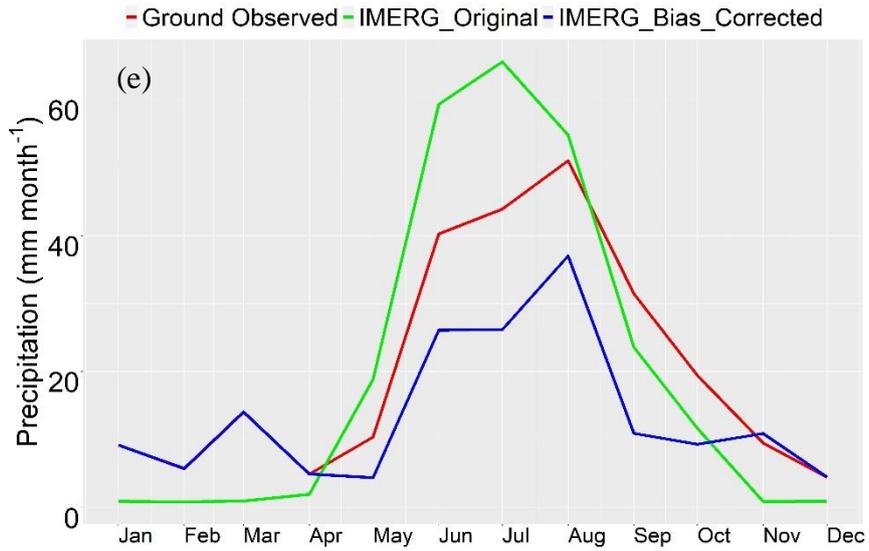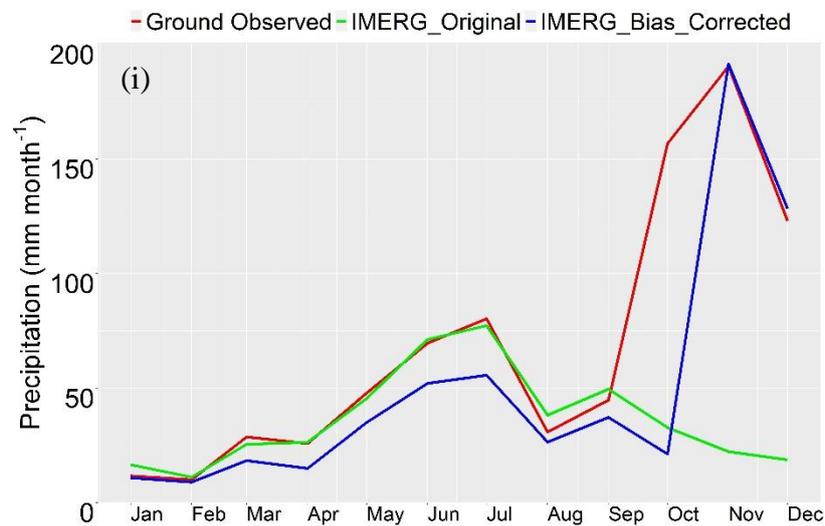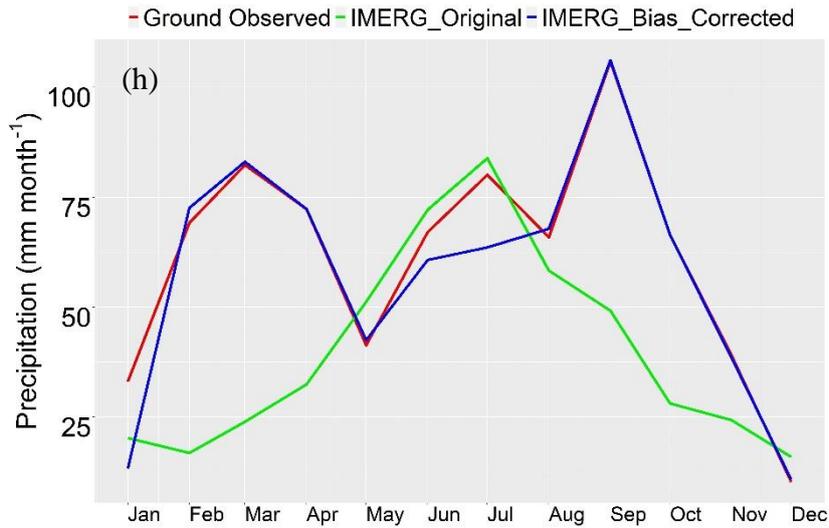
Figure 38. Spatially averaged monthly precipitation time series of gauge observation, IMERG-Ori, and IMERG-BC at validation pixels over the nine evaluation clusters ((a) represents cluster no.1 and (i) represents cluster no.9)

- **Statistical Evaluation**

Figure 39(a-i) compares the mean CC values of original and bias-corrected data calculated against the gauge observed data over different clusters for validation pixels. Figure 39(a) (cluster no.1) depicts the lower performance of BC in May, July, and August. For cluster no.2 (Figure 39(b)), also from May to September, the Ori data shows slightly higher CC values. In cluster no.3 (Figure 39(c)) just for Jun and August, the Ori data indicates slightly better values of CC. The worst performance is associated with the clusters no.5 and 6 (Figures 39(e) and 39(f)) where the value of CC related to BC is less than Ori over 6 and 5 warm months

respectively, while the best CCs of BC is obtained in clusters no.4 and 7 (Figures 39(d) and 39(g)) with just less value than Ori in one month. Clusters no.8 and 9 ((Figures 39(h) and 39(i)) with approximately the same performance indicate more reasonable improvements of BC than Ori. Generally, it is worth mentioning that the mean value of CC associated with BC over all clusters is higher than Ori CCs and the corresponding improvement rates are 36%, 15%, 35%, 39%, 27%, 18%, 5%, 40%, and 69%.
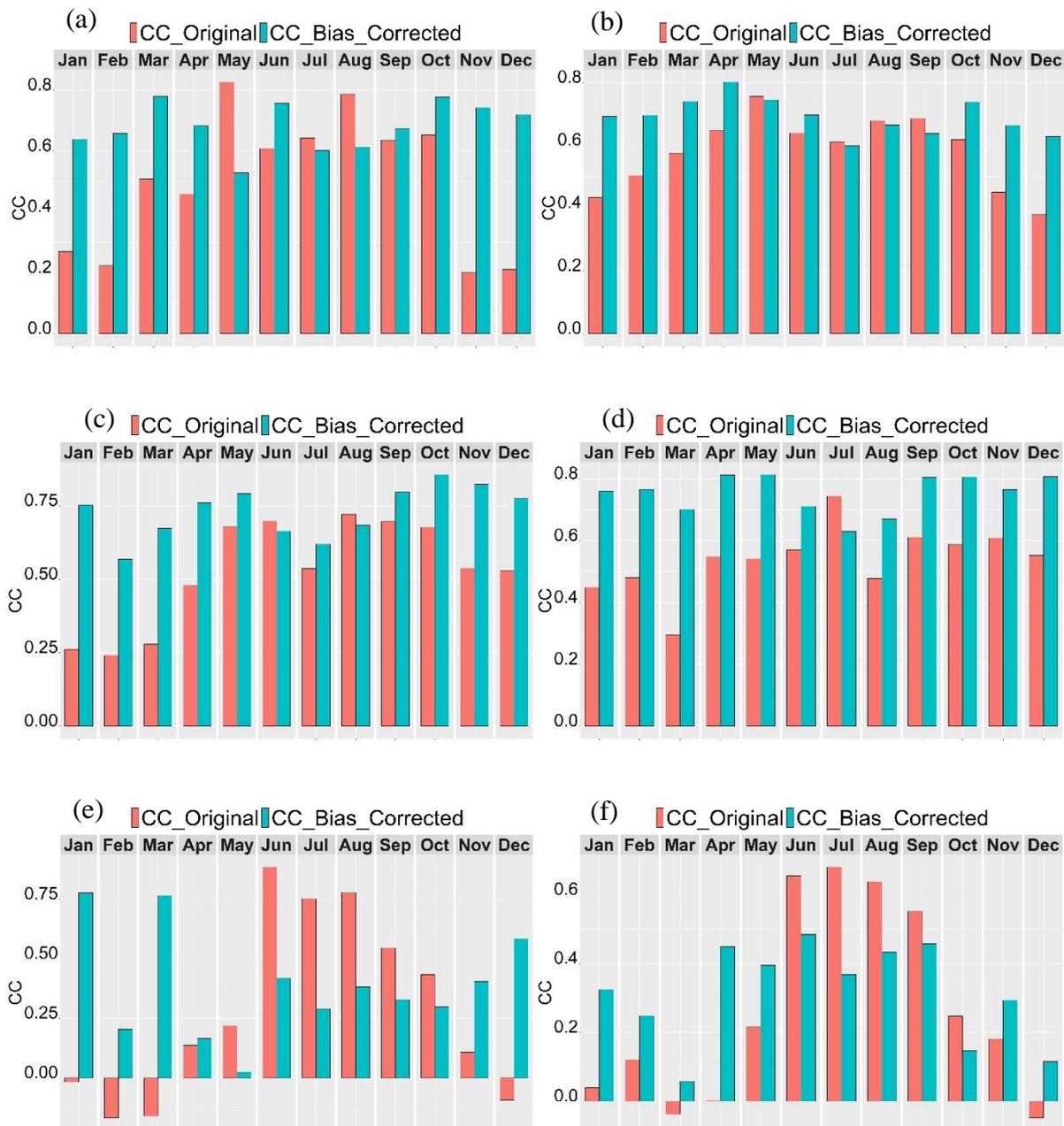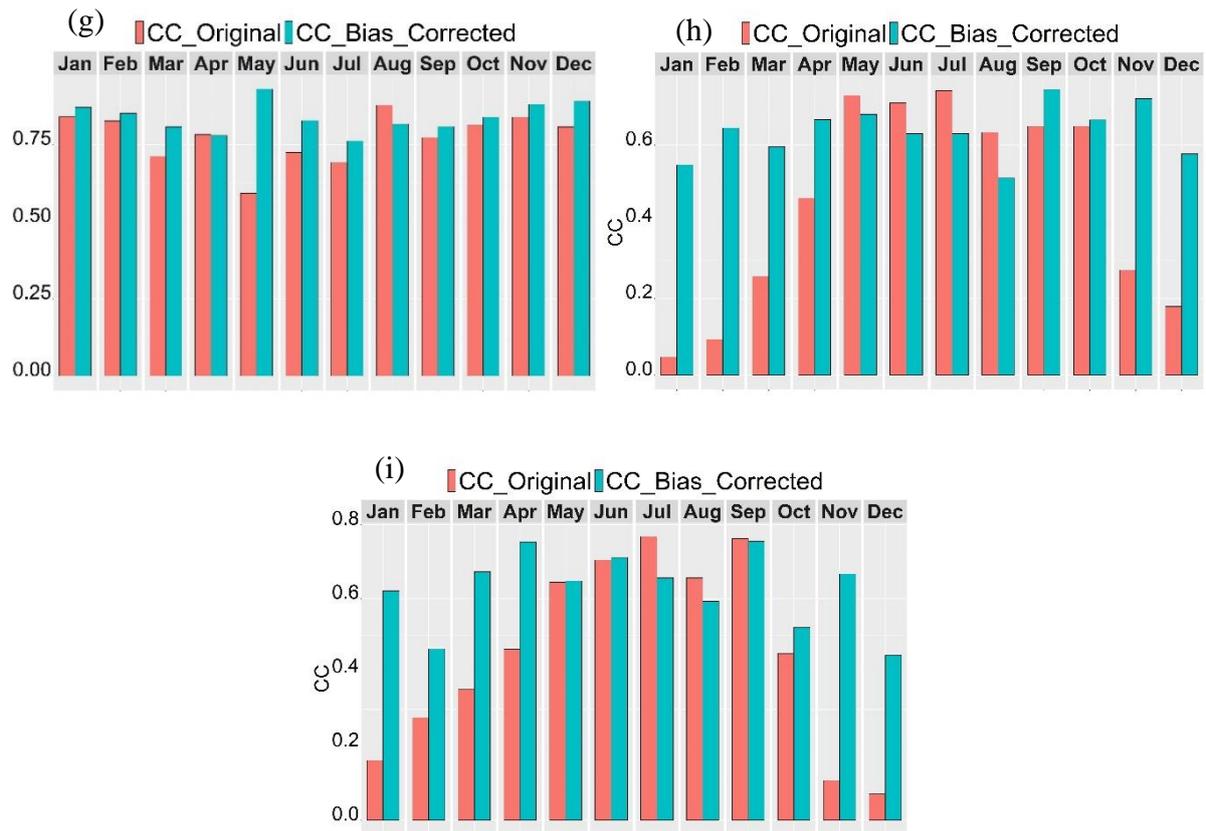
Figure 39. CC (correlation) spatially averaged daily of IMERG-Ori and IMERG-BC against gauge observation at validation pixels over the nine evaluation clusters ((a) represents cluster no.1 and (i) represents cluster no.9)

In terms of RBias, it can be seen in Figure 40 that the BC data outperforms Ori in most cases especially for the clusters no.1, 2, 5, 6, 7, and 9 where about 70% of months have less value of RBias related to BC. However, clusters no.3, 4, and 8 have a lower performance with higher RBias in almost 50% of months. There is not any consistency between CC and RBias as the underestimation/overestimation of both Ori and BC data leads to negative/positive value of RBias which may decrease/increase its mean value. Generally, 32% of improvement is obvious considering the entire regions and all months (mean RBias of BC data is 0.0211, and CC data is -0.065). In clusters no.1, 5, 6, 7, 8, and 9 the BC implies lower RBias of 0.035, 0.017, 0.0187, 0.107, -0.34, and 0.09, respectively, while for Ori the corresponding values are 0.22, -0.325, 0.325, 0.24, -0.4, and -0.14. In three other clusters of no.2, 3, and 4, the Ori shows better RBias with values of -0.017, 0.0008, and 0.16 against BC values of -0.03, -0.0275, and 0.32.
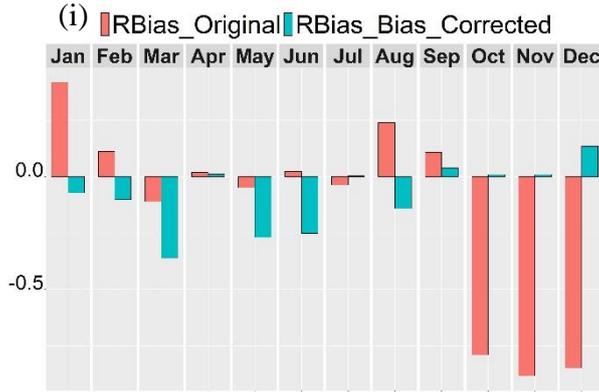
Figure 40. Spatially averaged daily RBias of IMERG-Ori and IMERG-BC against gauge observation at validation pixels over the nine evaluation clusters ((a) represents cluster no.1 and (i) represents cluster no.9)

Figure 41 displays the RMSE as a powerful metric for evaluating the performance of BC against Ori data. In cluster no.1 the mean value of RMSE over all months is improved by 25% (from 1.4 in Ori to 1.05 in BC). Also, for clusters no.2, 3, 4, 6, and 7 the rate of improvements by BC are respectively 23%, 25%, 22%, 7%, and 34%. In contrast, clusters no.5, 8, and 9 indicate higher RMSE of BC with rates of 28%, 10%, and 19% respectively. Overall, it can be concluded that RMSE is improved by 9% for all clusters.
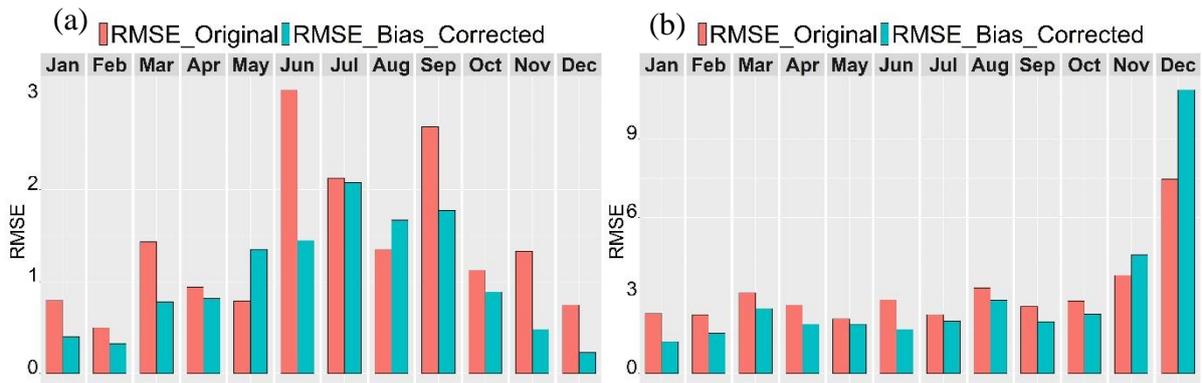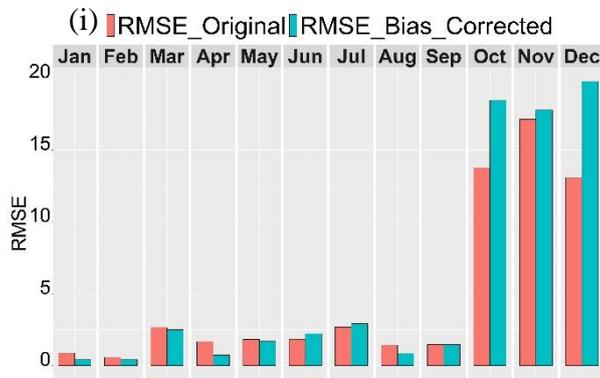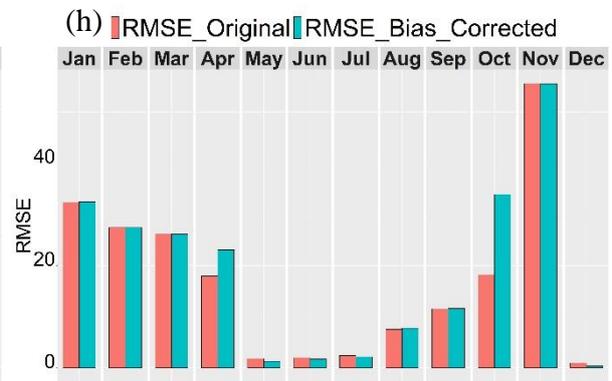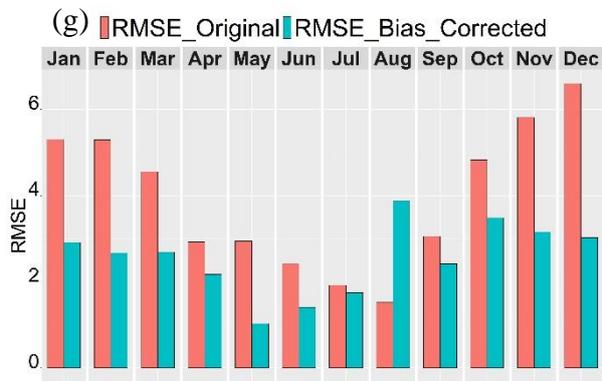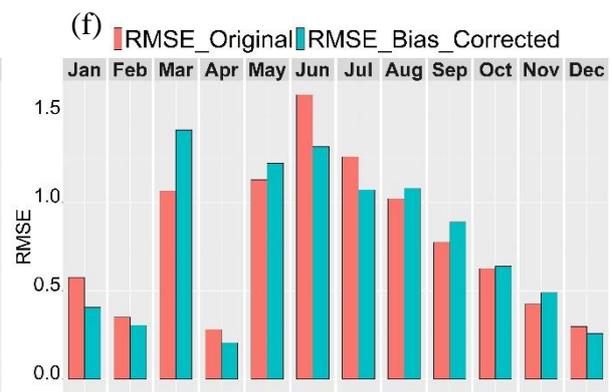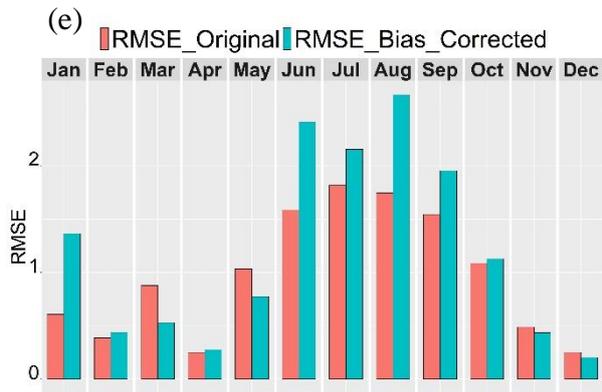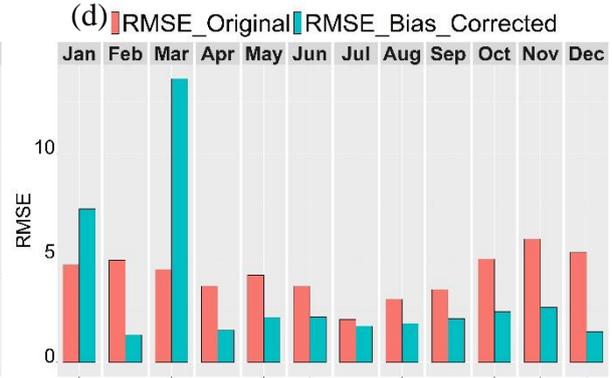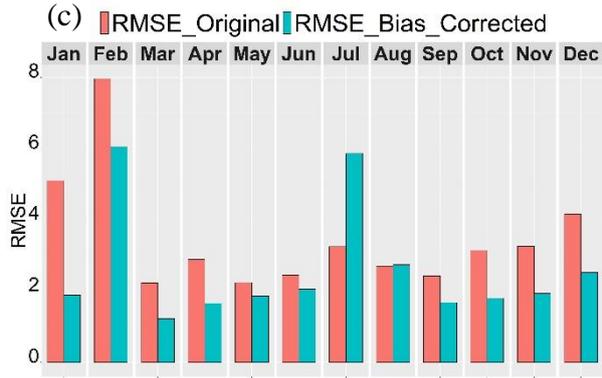
Figure 41. Spatially averaged daily RMSE (mm/day) of IMERG-Ori and IMERG-BC against gauge observation at validation pixels over the nine evaluation clusters ((a) represents cluster no.1 and (i) represents cluster no.9)

Table 7 lists the mean values of three above-explained metrics at validation pixels for each cluster separately which provides more detailed information quantitatively. In addition to the statistical metrics for evaluating the performance of the bias correction method developed in this study especially over the ungauged pixels, in the following the comparison of quantiles related to Ori and BC data against gauge observed precipitation is presented. In terms of quantile values as depicted in Figures 42(a)- (i), at all clusters after bias correction the satellite estimates become in agreement with the observed data perfectly despite the underestimations/overestimations of the raw data. These findings are obvious, as the model of bias correction is based on the quantile mapping which fits different quantile levels of satellite with the gauge observed data.
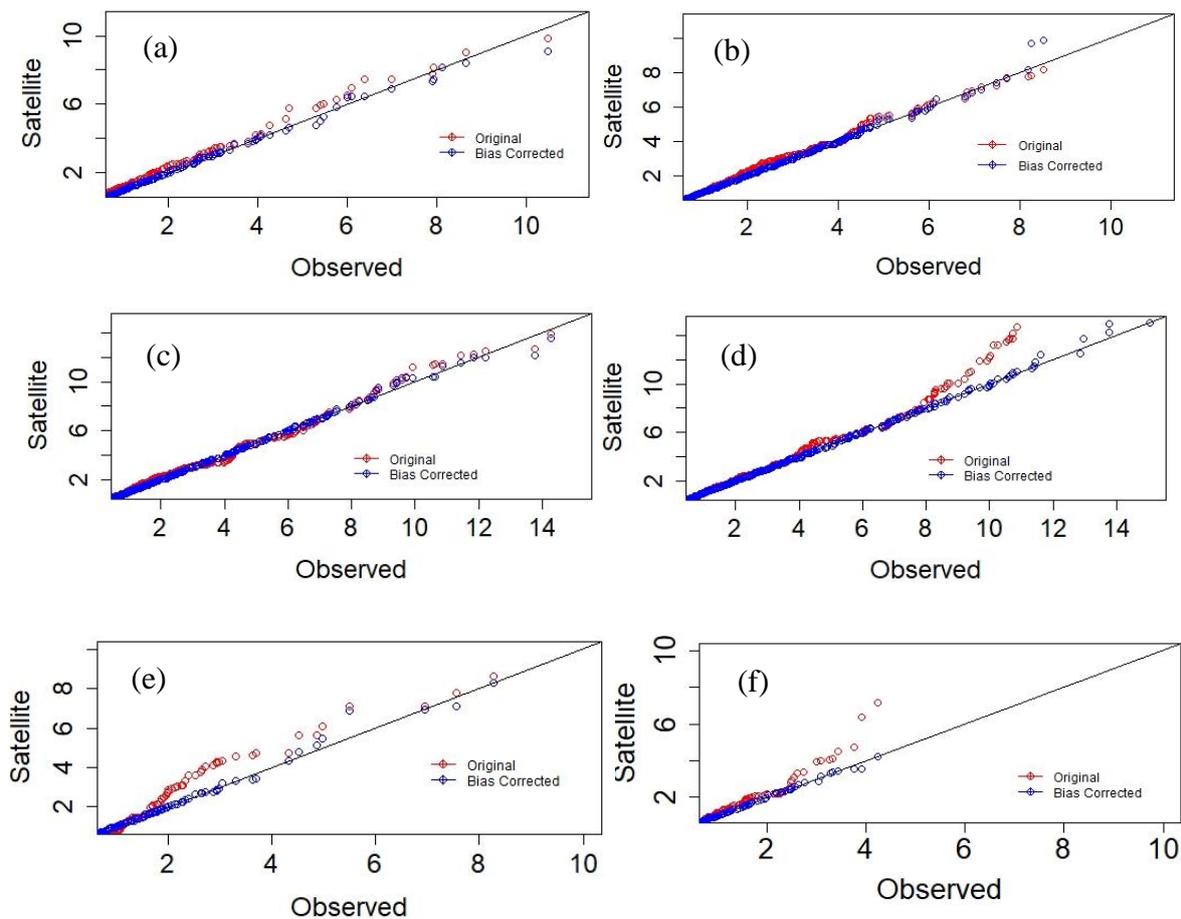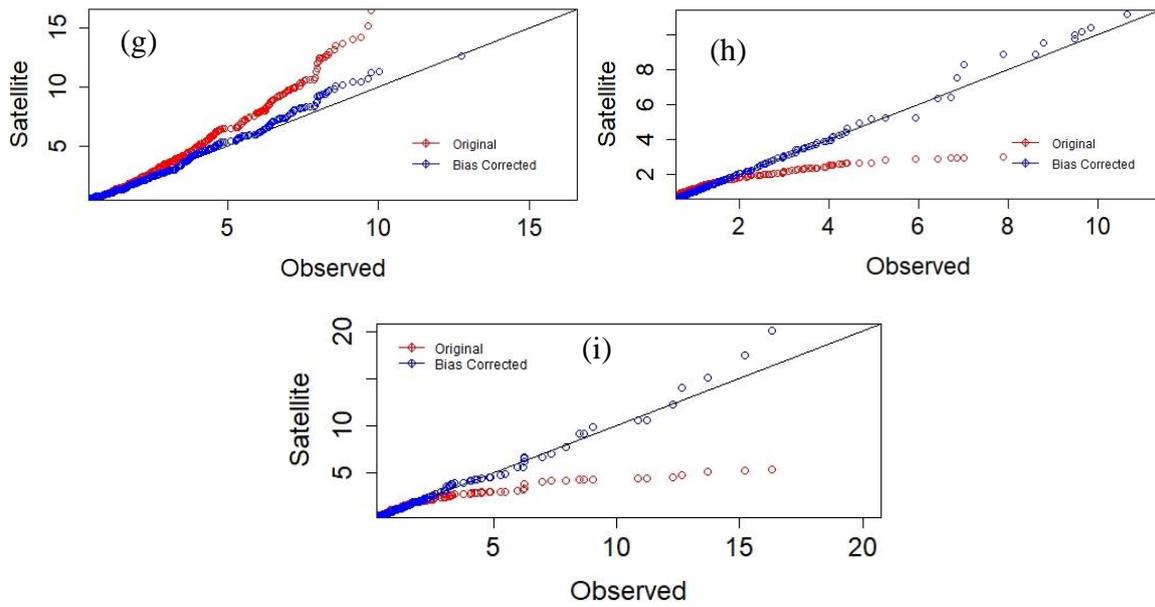
Figure 42. Spatially averaged daily Q-Q plot of IMERG-Ori and IMERG-BC against gauge observation at validation pixels over the nine evaluation clusters ((a) represents cluster no.1 and (i) represents cluster no.9) (unit is mm/day)

Table 7. Mean statistics value at validation pixels over each cluster

Time	Statistics	Cluster 1		Cluster 2		Cluster 3		Cluster 4		Cluster 5		Cluster 6		Cluster 7		Cluster 8		Cluster 9	
		Ori	BC	Ori	BC	Ori	BC	Ori	BC	Ori	BC	Ori	BC	Ori	BC	Ori	BC	Ori	BC
January	CC	0.48	0.88	0.1	0.68	0.15	0.76	0.28	0.83	0.4	0.87	0.34	0.87	0.53	0.92	0.07	0.66	0.03	0.61
	RMSE	12.08	4.09	2.9	1.85	3.73	1.83	3.4	1.84	8.36	4.6	9.2	4.6	10.4	3.23	1.18	0.8	5.0	4.6
	RBias	0.77	-0.16	0.43	-0.43	0.7	-0.3	0.49	-0.24	0.37	-0.2	0.73	-0.18	0.39	-0.1	-0.8	-0.4	-0.83	-0.42
February	CC	0.52	0.9	0.13	0.81	0.21	0.78	0.25	0.77	0.36	0.84	0.28	0.82	0.9	0.45	0.03	0.57	0.07	0.78
	RMSE	10.56	3.17	2.7	1.53	3.96	1.49	2.55	0.45	10.4	5.55	9.4	5.3	9.96	3.17	1.57	1.36	7.29	5.46
	RBias	0.75	-0.14	0.25	-0.28	0.35	-0.29	0.72	-0.31	0.14	-0.23	0.1	-0.18	0.35	-0.1	-0.88	-0.56	-0.88	-0.24
March	CC	0.5	0.89	0.35	0.85	0.31	0.91	0.5	0.87	0.5	0.88	0.28	0.87	0.42	0.9	0.003	0.76	0.02	0.64
	RMSE	8.5	2.7	2.24	1.08	4.5	2.02	2.2	0.85	9.02	6.0	5.25	2.25	8.4	2.77	0.95	0.59	4.35	3.9
	RBias	0.58	-0.14	0.29	-0.23	0.37	-0.16	0.35	-0.23	0.1	-0.2	0.22	-0.2	0.23	-0.1	-0.82	-0.33	-0.75	-0.42
April	CC	0.55	0.92	0.58	0.88	0.48	0.89	0.5	0.87	0.57	0.92	0.52	0.92	0.56	0.92	0.015	0.75	0.02	0.7
	RMSE	7.0	2.59	2.53	1.3	1.92	0.84	1.55	0.68	7.34	4.31	6.15	2.13	8.1	3.3	0.59	0.39	3.4	3.02
	RBias	0.54	-0.13	0.2	-0.2	0.24	-0.25	0.18	-0.27	0.11	-0.14	0.12	-0.09	0.2	-0.1	-0.6	0.34	-0.75	-0.35
May	CC	0.64	0.92	0.69	0.89	0.59	0.9	0.74	0.91	0.66	0.9	0.65	0.92	0.6	0.95	0.31	0.85	0.35	0.75
	RMSE	5.4	1.8	3.27	1.87	2.54	1.15	2.37	1.08	6.0	3.62	5.1	2.18	7.18	2.05	1.59	0.62	1.78	1.04
	RBias	0.47	-0.18	0.2	-0.2	0.19	-0.2	0.2	-0.17	0.16	-0.17	0.09	-0.11	0.22	-0.07	1.98	-0.33	0.04	-0.36
Jun	CC	0.63	0.88	0.7	0.89	0.66	0.9	0.7	0.89	0.66	0.9	0.6	0.91	0.64	0.93	0.53	0.91	0.61	0.9
	RMSE	5.83	2.65	4.6	2.35	3.65	1.92	4.7	2.55	7.03	4.0	5.15	2.15	6.7	2.5	3.15	1.05	3.52	1.42
	RBias	0.44	-0.22	0.12	-0.2	0.12	0.21	0.17	-0.19	0.07	-0.16	0.17	-0.13	0.16	-0.1	2.5	-0.17	0.46	-0.21
July	CC	0.61	0.84	0.67	0.85	0.69	0.88	0.64	0.86	0.63	0.88	0.6	0.92	0.58	0.9	0.62	0.92	0.64	0.9
	RMSE	6.4	3.04	4.47	2.8	3.54	2.1	5.3	3.55	8.1	4.0	4.94	2.3	6.1	3.04	3.21	1.21	2.9	1.47
	RBias	0.4	-0.24	0.17	-0.27	0.2	-0.29	0.13	-0.23	0.125	-0.21	0.16	-0.17	0.19	-0.19	0.45	-0.18	0.37	-0.26
August	CC	0.71	0.88	0.7	0.86	0.64	0.85	0.7	0.83	0.64	0.9	0.65	0.9	0.6	0.91	0.65	0.92	0.65	0.9
	RMSE	5.17	2.89	3.56	2.6	2.5	1.69	3.15	2.0	8.3	4.2	6.23	3.05	9.0	4.0	3.03	1.4	2.54	1.26
	RBias	0.65	-0.23	0.17	-0.28	0.17	-0.29	0.26	-0.3	0.13	-0.2	0.26	-0.17	0.16	-0.15	0.53	-0.21	0.32	-0.2
September	CC	0.66	0.92	0.7	0.92	0.75	0.94	0.7	0.88	0.68	0.9	0.66	0.93	0.66	0.94	0.73	0.95	0.52	0.84
	RMSE	7.04	2.83	3.37	1.73	2.63	1.1	4.4	2.41	6.17	3.4	5.9	2.58	9.5	2.5	1.8	0.61	1.96	0.98
	RBias	0.48	-0.13	0.17	-0.16	0.26	-0.16	0.18	-0.18	0.17	-0.17	0.17	-0.09	0.23	-0.1	0.08	-0.14	0.46	-0.3
October	CC	0.7	0.92	0.63	0.87	0.61	0.91	0.57	0.89	0.63	0.93	0.68	0.95	0.61	0.96	0.23	0.79	0.29	0.78
	RMSE	10.67	3.15	2.63	1.58	4.07	1.94	2.43	1.13	8.6	4.03	6.46	2.32	9.9	2.4	2.05	1.21	1.45	0.68
	RBias	0.48	-0.10	0.08	-0.2	0.1	-0.18	0.23	-0.1	0.13	-0.2	0.17	-0.07	0.38	-0.06	-0.37	-0.28	0.003	-0.3
November	CC	0.58	0.91	0.32	0.84	0.17	0.85	0.17	0.86	0.52	0.87	0.52	0.91	0.53	0.94	0.007	0.75	0.16	0.7
	RMSE	13.41	3.62	3.05	1.65	5.6	1.64	1.8	0.59	9.9	5.85	6.83	2.63	10.3	2.5	1.08	0.71	4.9	4.4
	RBias	0.5	-0.1	0.1	-0.23	0.34	-0.23	0.52	-0.22	0.13	-0.11	0.19	-0.08	0.27	-0.07	-0.78	-0.28	-0.77	-0.26
December	CC	0.51	0.91	0.12	0.74	0.07	0.74	0.19	0.76	0.36	0.84	0.59	0.91	0.6	0.94	-0.02	0.63	0.015	0.68
	RMSE	13.31	2.87	2.41	1.37	4.58	2.04	1.34	0.53	9.6	5.7	8.7	3.9	9.5	3.15	0.67	0.49	0.64	0.42
	RBias	0.93	-0.12	0.4	-0.38	0.76	-0.35	0.44	-0.32	0.27	-0.23	0.085	-0.09	0.32	-0.07	-0.75	-0.4	-0.75	-0.3

- **CDF Evaluation at Ungauged Pixels**

In this section of validation, the bias correction method, the Empirical Cumulative Distribution Function (ECDF) associated with the observed data is compared with their corresponding Ori and BC IMERG estimates. Figures 43(a)-(i) illustrate the ECDFs extracted from the mean daily precipitation during five years of the study period at cluster no.1 to 9, respectively. It is presented that the CDFs of the two observed and BC data sources for all clusters except in cluster no.5 which BC shows underestimation are relatively close to each other, indicating improvement of the systematic bias in Ori estimation after bias correction. However, Ori tends to systematically overestimate daily precipitation over clusters no.1 to 5 and 8, and underestimate it over clusters no.6, 7, and 9 respectively.
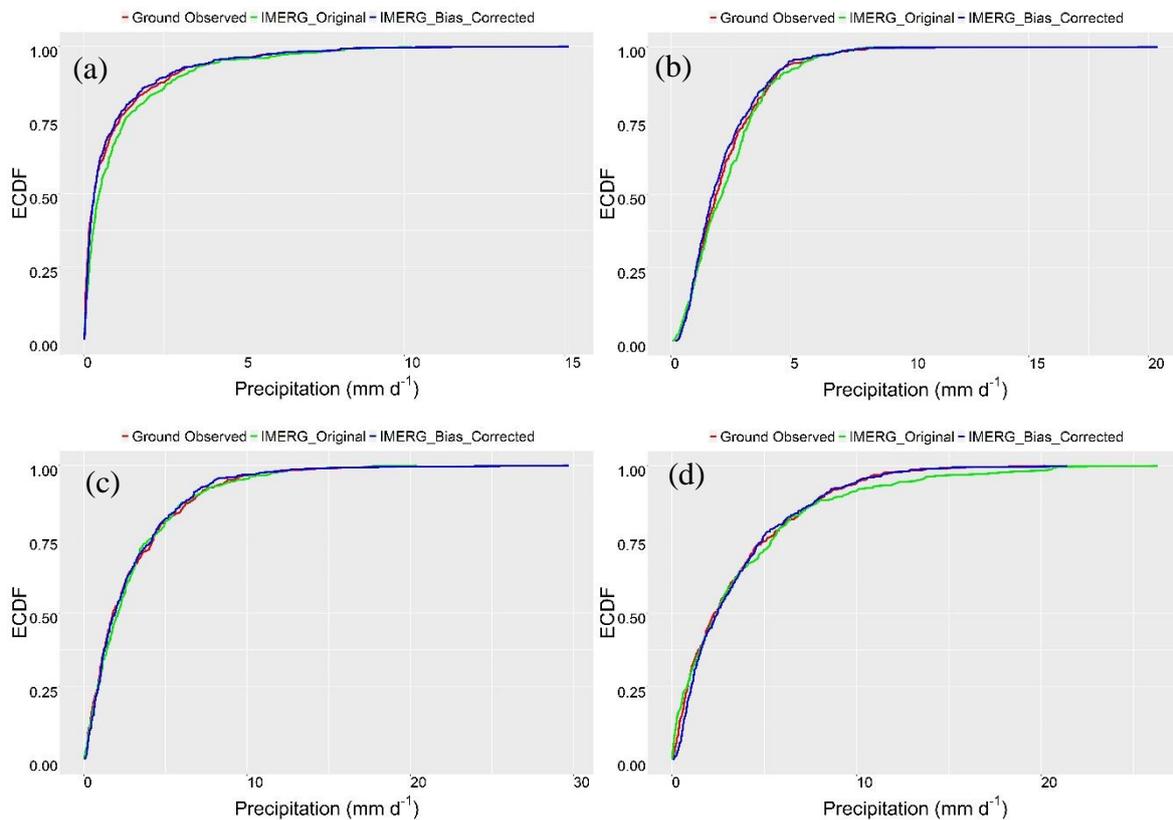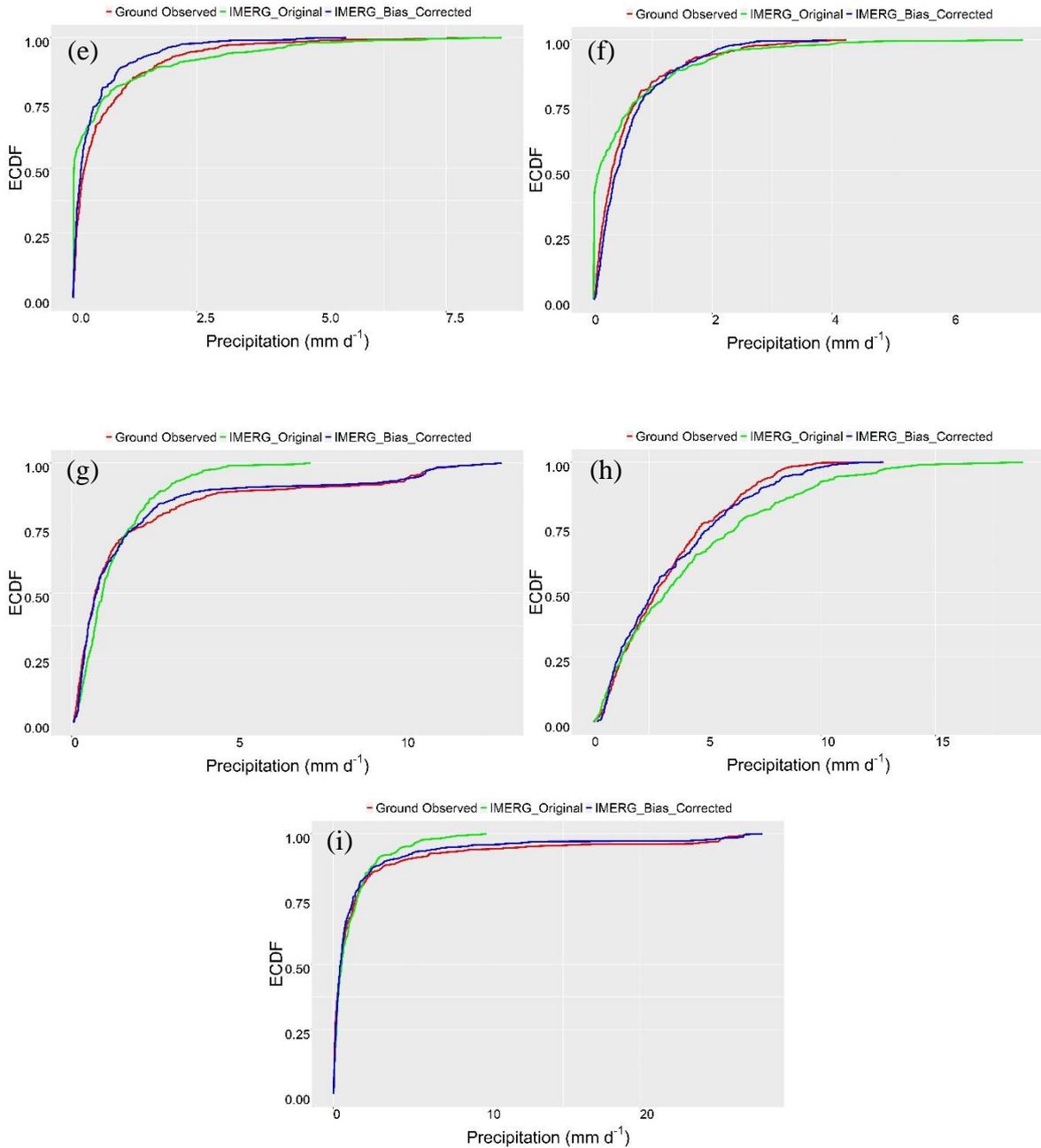
Figure 43. Spatially averaged daily CDFs of gauge, original, and bias-corrected IMERG estimations at validation pixels for the nine evaluation clusters ((a) represents cluster no.1 and (i) represents cluster no.9)

4.5. Summary and Conclusion

In this chapter, a method for bias correction of satellite IMERG PrCal product that showed the best performance among those other products (PrUncal, PrIR, and PRHQ) was developed aiming to introduce a more reliable remotely-sensed of high spatiotemporal resolution precipitation data over Canada. For this purpose, in addition to IMERG and ground-based

precipitation data, a set of invariant and time-varying reanalysis predictors resampled over the satellite pixel spatial resolution (0.1°) and was considered as independent variables for constructing the regression-based quantile mapping model. K-means and hierarchical clustering techniques were further used for categorizing the study area based on the climatic similarities between the satellite pixels as well as the spatial distribution of rain gauges. First, the method of bias correction was implemented at each cluster separately and the related parameters obtained over the gauged pixels for different quantile levels of precipitation data. Then, by using IDW, the parameters interpolated over the ungauged pixels inside the corresponding cluster. In the IDW technique here, the distance was defined as the climatic distance, not geographical distance. In the end, by applying the reanalysis covariates at each ungauged pixel in the interpolated model parameters of that pixel the bias-corrected satellite data was obtained.

For validation of the developed model at ungauged pixels, the bootstrap technique was used. By bootstrapping some gauged pixels (10% of total gauged pixels in each cluster) were selected randomly and assumed as the ungauged pixels in which have not been participating in the parameter estimating procedure of the proposed model. These pixels named as validation pixels and they were tested for assessing the model performance.

Overall the obtained results for both gauged and validation pixels indicated a major improvement in the IMERG estimates after bias correcting for both terms of precipitation intensity and occurrence especially over the regions where the original data had lower performances (west and coast parts of the country). In addition, the trend of precipitation time series can be captured by the bias-corrected data reasonably. However, at clusters no.5 and 6 (Figure 19), the RMSE and CC metrics were not improved after bias correcting. The reason is that due to the lower numbers and unevenly distributed of gauged pixels, the interpolation cannot be implemented perfectly over those ungauged pixels. Besides, the diversity in topographic and climatic conditions in these regions, makes it challenging to find an accurate climatic distance between the pixels. Further analysis needs to be performed by using more accurate covariates over such sparse-data areas.

Considering the errors in IMERG products discussed in Chapter 3 of this study, the findings of this chapter which correct the biases spatially and temporally are useful especially over the satellite pixels where there is no rain gauge for estimating the direct bias.

4.6. References

Amorim, R. C.; Hennig, C., 2015. Recovering the number of clusters in data sets with noise features using feature rescaling factors. Information Sciences. 324: 126–145. arXiv:1602.06989. doi:10.1016/j.ins.2015.06.039

Acharya N., Chattopadhyay S., Mohanty U.C., Dasha S.K., Sahooc L.N., 2013. On the bias correction of general circulation model output for Indian summer Monsoon. J Meteorol Appl 20(3):349–356

Barry, R.G. and Chorley, R.J. Atmosphere, weather and climate (7th ed) Routledge 1998 p. 80-81 ISBN 0-415-16020-0.

Carl de Boor, A Practical Guide to Splines, Published in Applied Mathematical Sciences 1978, Mathematics, Computer Science

Davis R.E., 1976. Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean. J Phys Oceanogr 6:249–266

Dee, D.P. et al., 2014. Toward a Consistent Reanalysis of the Climate System. Bull. Amer. Meteor. Soc., 95(8), 1235-1248

Diaz-Nieto J., Wilby R.L., 2005. A comparison of statistical downscaling and climate change factor methods: impacts on low flows in the River Thames, United Kingdom. Climat Change 69(2–3): 245–268

Gudmundsson, L., J. B. Bremnes, J. E. Haugen, and T. Engen Skaugen, 2012. Technical note: Downscaling RCM precipitation to the station scale using quantile mapping—A comparison of methods, Hydrol. Earth Syst. Sci. Discuss., 9(5), 6185–6201, doi:10.5194/ hessd-9-6185-2012.

Gumindoga W., Rientjes T.H.M., Tamiru Haile A., Makurira H., Reggiani P., 2019. Performance of bias-correction schemes for CMORPH rainfall estimates in the Zambezi River basin, Hydrol. Earth Syst. Sci., 23, 2915–2938, 2019, https://doi.org/10.5194/hess-23-2915

Grillakis M.G., Koutroulis A.G., K. Tsanis, I.K., 2013. Multisegment statistical bias correction of daily GCM precipitation output, Journal of Geophysical Research: Atmospheres, Vol. 118, 3150–3162, doi:10.1002/jgrd.50323, 2013).

Hastie, T., Tibshirani, R., and Friedman, J. H.: The Elements of Statistical Learning, Springer, 2001. 6190, 6191

Hay L.E., Wilby R.J.L., Leavesley G.H., 2000. A comparison of delta change and downscaled GCM scenarios for three mountainous basins in the United States. J Am Water Resour Assoc 36:387–397. doi:10.1111/j.1752-1688.2000.tb04276.x

Hofmann-Wellenhof, B. and Moritz, H. "Physical Geodesy", 2005. ISBN 3-211-23584-1.

Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P; Gao, X.; Ferreira, L.G (2002). "Overview of the radiometric and biophysical performance of the MODIS vegetation indices". Remote Sensing of Environment. 83 (1–2): 195 213. Bibcode:2002RSEnv. 83..195H. doi:10.1016/S0034-4257(02)00096-2.

Ines A.V.M., Hansen J.W., 2006. Bias correction of daily GCM rainfall for crop simulation studies. J Agric For Meteorol 138:44–53

Kantha, L.H., Carol C., 2000. Small scale processes in geophysical fluid flow. San Diego: Academic Press.

Kharin V.V., Zwiers F.W., 2002. Notes and correspondence climate predictions with multimodel ensembles. J Clim 15:793–799

Lafon T., Dadson S., Buys G., Prudhomme C., 2013. Bias correction of daily precipitation simulated by a regional climate model: a comparison of methods. Int J Climatol 33:1367–1381

Najafi M.R., Zwiers F.W., Gillett N.P., 2016. Attribution of the spring snow cover extent decline in the Northern Hemisphere, Eurasia and North America to anthropogenic influence. Climatic Change. 136(3-4):571-86.

Najafi M.R, Zwiers F.W., Gillett N.P., 2017. Attribution of observed streamflow changes in key British Columbia drainage basins. Geophysical Research Letters, 44(21):11-012.

Ramsay J.O., Hooker G., Spencer G., 2009. Functional data analysis with R and Matlab, Springer, New York.

Tang G., Behrangi A., Long D., Li C., Hong Y., 2018. Accounting for spatiotemporal errors of gauges: a critical step to evaluate gridded precipitation products. J. Hydrol. 559, 294–306. https://doi.org/10.1016/j.jhydrol.2018.02.057.

Ward J. H., Jr., 1963. Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association, 58, 236–244.

Zhang X. J., Tang Q.H., 2015. Combining satellite precipitation and long-term ground observations for hydrological monitoring in China, J. Geophys. Res. Atmos., 120, 6426–6443, doi:10.1002/2015JD023400.

Chapter 5

5. Hydrological Evaluation of Daily IMERG Data

5.1. Introduction

Hydrological modeling is being used to forecast streamflow and flood for years. Precipitation is the key variable in the hydrological models and accurate precipitation input is important to rainfall-runoff models for rationally simulating the hydrological processes at the regional and basin scales. Over the remote areas and complex terrain with sparse ground-based rain gauges, developing a hydrological model would be challenging and its output may not be reliable (Tang et al., 2016; Yuan et al., 2018). Three ways to measure precipitation include rain gauges, weather radars, and satellite-based sensors (Li et al. 2013). As mentioned before, rain gauges are not available sufficiently over most of the basins and they are also subject to the impact of topography on precipitation pattern, and the wind-induced undercatch, with an increasing fraction of solid precipitation (Schwarb, 2000; Sevruk, 1997; Sevruk et al., 2009). The weather radar can detect and estimate precipitation with relatively higher temporal and spatial resolutions, however, it is often subject to signal blockage, attenuation by rain, and vertical variability of reflectivity which reduces its quality of data (Dinku et al. 2002; Tian and Peters-Lidard, 2010). Currently, satellite-based precipitation products as an alternative source are able to provide high-quality data to employ in hydrological models especially over complex terrain and data-sparse or ungauged basins. However, due to the indirect nature of precipitation estimates by satellite algorithms as discussed in previous chapters of this study, they are subject to uncertainty and errors that need to be evaluated before using in hydrological processes.

By increasing the use of SPPs in recent years, assessment of the performance of these kinds of data in hydrological applications has been conducted in several studies (Hao et al., 2014; Jiang et al., 2016; Su et al., 2017; Meng et al., 2014). Among the most widely-used SPPs containing TRMM-TMPA, PERSIANN, CMORPH, and IMERG, the last one indicates higher performance as input data into different hydrological models over different basins (Wang et al., 2017; Yuan et al., 2018; Su et al., 2019). However, considering that IMERG was released only six years ago, few studies have focused on the hydrological utilities of IMERG products

at basin scales with no study over a basin in Canada. Therefore, in this study, the main objective is to evaluate the IMERG skills as enforcing data into a hydrologic model over a small basin in Canada. For this purpose, the last version (V06) of IMERG PrCal at daily scale during 15 years from 2001 to 2015 is applied in Raven as a robust and flexible hydrological modelling framework over Batchawana River Watershed (BRW) in central Ontario, Canada.

A few previous studies have focused on the hydrological simulations using IMERG products in other basins in the world. Tang et al, 2016 assessed the Day 1 IMERG Final Run product by using the CREST hydrological model over the midlatitude Ganjiang River basin in southeast China and found that IMERG is consistent with the gauge-observed data in daily streamflow simulations with NSE of 0.68. In the Beijiang River basin in China, Wang et al., 2017 expressed that daily streamflow simulations using the VIC model enforced by the Day 1 IMERG Final Run product present a reasonable hydrological performance with NSE of 0.742. Zubieta et al. (2017) indicated that IMERG is as useful as 3B42V7 and 3B42RT in modeling streamflows in southern regions of the Peruvian–Ecuadorian Amazon basin, but these data sets fail to properly simulate streamflows in northern regions. As a suitable product for hydrological applications, IMERG can use for simulating streamflow and flood in daily and sub-daily time scales over different basins where local rain gauge networks are sparse (Tang et al. 2016; He et al., 2017; Zubieta et al. 2017; Wang et al., 2017; Yuan et al., 2018). Therefore, this chapter is also pursuing the advantages of this most recent well-known precipitation data in a calibrated hydrological model as a first and preliminary study over a basin in the southern part of Ontario, Canada. It can be of great importance as flood damage has become a growing problem in Ontario (Environment and Climate Change Canada, Pirani and Najafi, 2019). Besides, the study region suffers from high quality of ground-based rain gauges which leads to unreliable streamflow/flood simulations.

It is worth mentioning that, in Chapters 3 and 4 of this thesis the comprehensive statistical evaluation of IMERG precipitation data followed by developing a spatiotemporal bias correction technique was presented. Then in this chapter by assessing the ability of IMERG-V06 in streamflow simulation, can provide a secondary check of this new product, particularly for hydrologic applications.

The chapter consists of 6 sections. Section 2 and 3 describe the study area and data respectively. Section 4 explains the hydrologic model. Section 5 outlines the results and discussion of findings. Finally, concluding remarks are presented in Section 6.

5.2. Study Area

The Batchawana River Watershed (BRW), which drains into Batchawana Bay on the eastern shore of Lake Superior, is approximately 1280 km^2 and is located in central Ontario, Canada (Figure 44). It is in a transition zone between the Great Lakes-St. Lawrence Forest to the south and the Boreal Forest to the north. Three distinct vegetation zones exist within the BRW. The climate in the BRW is continental, with precipitation being influenced by the lake effect caused by Lake Superior to the west of the basin, and local orographic effects in areas of high relief. The mean annual precipitation for the Montreal Falls meteorological station during 1977-2011 was 1180 mm, ranging from 763 to 1554 mm in any given year (Sanford et al., 2007).

5.3. Data

5.3.1. Precipitation Data

Precipitation data used in this chapter contains 1) satellite IMERG PrCal product that previously explained in "Data" section of Chapter 3 of this thesis, 2) daily rain gauge observation data. Based on the drainage area of the watershed (1280 km2) and also considering the spatial resolution of IMERG product (10×10 km), the study area is covered by a number of 16 satellite pixels that each of them has its precipitation amount aggregated from half-hourly to daily time scale during 2001-2018. In contrast, there is just one rain gauge (Montreal Falls) inside the watershed providing daily data from 1977 to 2011. Therefore, the similar time scale for comparing the two sources of precipitation is from 2001 to 2011 (11 years), however, as the daily streamflow station is available during 1967-2015 (next section), it is possible to calibrate the model based on the sufficient period of gauge observations (35 years). It also should be noted that all 16 satellite pixels are interpolated by using IDW over the basin to create one precipitation time series corresponding to the rain gauge data.

5.3.2. Streamflow Data

There is just one hydrometric station operated by Water Survey of Canada (WSC) in the outlet of watershed (Figure 44) which provides daily streamflow records from 1967 till the end of 2015. Based on hydrologic records from the WSC gauging station, peak discharges usually occur in either the spring (April –May) or late fall (October –November) in response to snowmelt or stormfall, respectively.
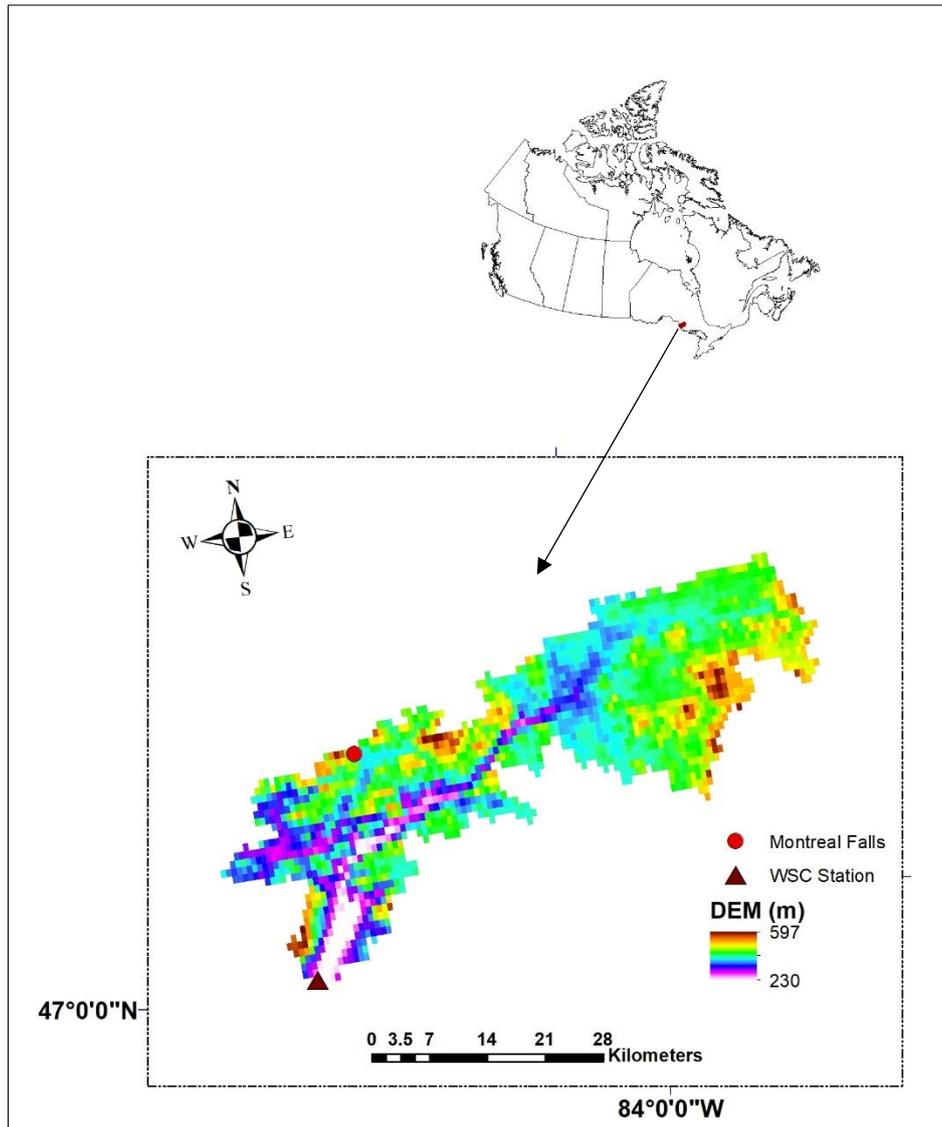
Figure 44. Batchawana River Watershed as the study area of hydrological modeling

5.4. Methodology

Evaluation of IMERG precipitation inputs as the objective of this chapter is performed by using the Raven model with respect to the ground-based streamflow observations at the watershed outlet.

5.4.1. Raven Hydrological Model

In this study, to identify the ability of IMERG PrCal in streamflow simulation the Raven hydrological model is utilized. Raven is a flexible hydrological modeling framework that allows for the development of lumped and semi-distributed models. The discretization of land into Hydrological Response Units (HRUs) and flexibility of using empirical models or physical systems to represent hydrological processes are some of the features of the Raven model (Shafii, 2017).

In this study, two different scenarios are used for hydrological evaluation of IMERG precipitation data. In the first scenario, the Raven model is calibrated by using the daily precipitation of one rain gauge located within the watershed from 1977 to 2000 (67% of the whole period (1977-2011)). The rest of the time series during 2001-2011 (33% of the whole period) is considered as the validation period. The model parameters are obtained based on the daily observed rain gauge data, then, the daily IMERG PrCal (aggregated from the original half-hourly data) during 2001-2015 is employed into the calibrated model without changing its parameters. In the second scenario, the model is calibrated based on the IMERG data from 2001 to 2015, then the performance of simulated streamflow is validated during 2001-2015. It is noted that due to the short term of having data (15 years) for IMERG, the total data is considered as the calibration period in this scenario.

The parameterization of the Raven model data is somewhat alleviated through the use of HRU classification schemes. Each HRU, in addition to being defined as having a unique representative area, centroid, slope, aspect, and elevation, is assumed to belong to a unique set of non-overlapping classifications. The benefit of this classification approach is that parameters need only be specified on a class-by-class basis rather than an HRU-by-HRU basis, simplifying the mechanics of calibration (each parameter shows up once in the model input files), the data storage (parameters are linked to class instances rather than directly to the HRU), and improving the portability of parameters to ungauged basins (Craig et al., 2020).

Forcing data (e.g., precipitation, incident radiation, etc.) is distributed to HRUs by using a generic form of interpolation between gauge stations. Raw forcing data required by Raven

includes, at the very least, time series of total precipitation and minimum/maximum or average daily temperature either in gridded netCDF format or at the set of gauges. Raven includes a suite of algorithms for determining related forcings (potential evapotranspiration, potential melt, shortwave radiation longwave radiation, snow/rain partitioning, cloud cover, relative humidity, wind speed, etc.) (Craig et al., 2020).

5.4.2. Evaluation Metrics

In this section, just the metrics used for the simulated streamflow evaluation are explained as the statistical evaluation metrics of precipitation have already presented in Chapters 3 and 4. The performance of the daily streamflow simulations is evaluated using four statistical indices: RMSE, Nash-Sutcliffe Efficiency coefficient (NSE), LogNSE, and Kling-Gupta Efficiency (KGE). The highest score of NSE is 1. Although the maximum of NSE is a commonly used objective function for the optimization of the hydrological model, it indicates better performance to high flows, and so the optimized model may not capture low flows accurately. Thus, for considering both high- and low-flow processes, the maximum sum of NSE and log-transformed NSE (LogNSE) are utilized here (Yuan et al., 2018).

$$RMSE = \sqrt{\frac{\sum_{i=I}^{n}(Q_i^o - Q_i^s)^2}{n}} \tag{5.1}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}(Q_i^s - Q_i^o)^2}{\sum_{i=1}^{n}(Q_i^o - \overline{Q^o})^2} \tag{5.2}$$

$$LogNSE = 1 - \frac{\sum_{i=1}^{n}\lfloor log(Q_i^s) - log(Q_i^o)\rfloor^2}{\sum_{i=1}^{n}\lfloor log(Q_i^o) - \overline{log(Q^o)}\rfloor^2} \tag{5.3}$$

Where n is the sample size of observed or simulated streamflow time series; Q_i^o is the observed streamflow time series at the hydrometric station (m^3/s); Q_i^s is the simulated streamflow time series by either satellite or rain gauge inputs (m^3/s); and Q_i^o and $\overline{log(Q^o)}$ represent the mean observed streamflow and mean log-transformed observed streamflow at the hydrometric station (m^3/s), respectively.

$$KGE = 1 - \sqrt{(r - 1)^2 + (\beta - 1)^2 + (\gamma - 1)^2} \tag{5.4}$$

KGE combines the Pearson correlation coefficient r, the bias ratio β, and the variability ratio γ into one metric that ranges from -∞ to 1 with 1 being the ideal score (Singh and Najafi, 2020).

5.5. Results and Discussions

The simulated streamflow using the IMERG precipitation inputs is compared with the observed streamflow to evaluate the hydrological utility of satellite precipitation products. First, Figure 45 compares the rain gauge daily observed precipitation used to calibrate the hydrological model with the IMERG precipitation data during 2001-2011. Based on the metric of Bias = - 0.44, IMERG indicates a slight underestimation in this watershed. Also, RMSE and CC values represent that the IMERG is not reasonably consistent with the ground-based data. So, it is expected that the biases associated with the IMERG propagate in the hydrological model and affect model output (streamflow). For exploring the extent of bias in the simulated streamflow by IMERG inputs, Figure 46 depicts the scatterplots of both simulated and observed streamflow against each other at the outlet of the watershed (hydrometric station).
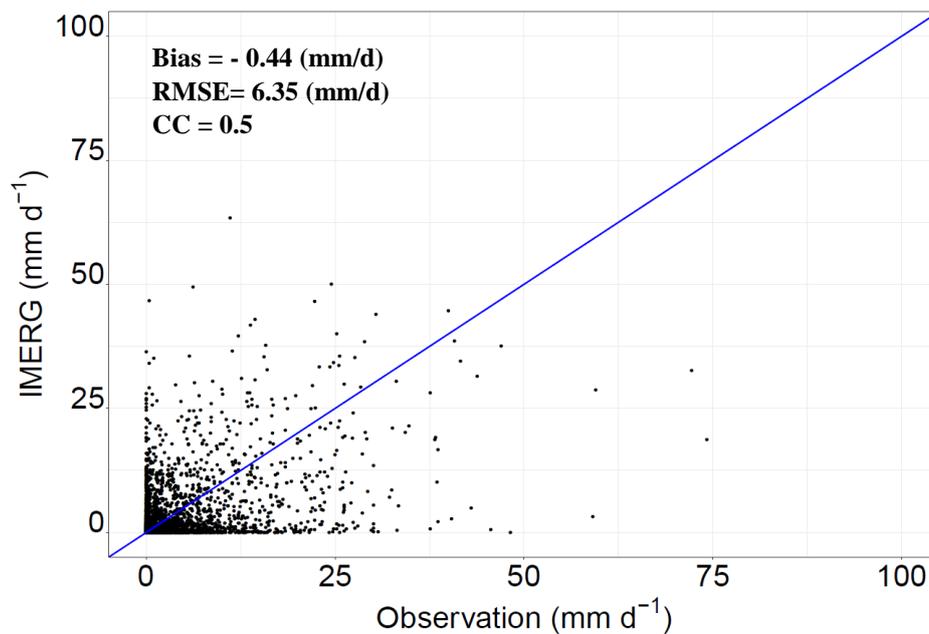
Figure 45. Comparison of the ground-observed (one rain gauge in the study area) and IMERG daily precipitation (average value of 16 pixels covered the study area) during 2001-2011

Comparing both simulated IMERG- and gauge-based streamflow in Figure (a) proves the underestimation of IMERG in streamflow forecasting, however, the correlation between two streamflow values (CC= 0.8) is higher than that between two precipitations (CC = 0.5 in Figure 45). Also with values of NSE = 0.58 and LogNSE = 0.7, IMERG performs better in capturing

low flow than high flow. By comparing the gauge-based simulated streamflow and observed flow at the hydrometric station in Figure 46 (b), it can be realized that the calibrated model forced by rain gauge data does not perform perfectly. Although the NSE and KGE values (0.67 and 0.74, respectively) are acceptable, they are still indicating the weakness of the calibrated model in streamflow simulation accurately. Nevertheless, with the LogNSE value of 0.74, the model displays better performance in capturing low flow. Considering both obtained results in Figures 46(a) and 46(b), two sources of uncertainties associated with the IMERG precipitation forced data and the calibrated model by gauge data are obvious. In this study, as we used the calibrated model, analyzing the model structure, its sources of uncertainty, and the performance of model parameters are not discussed here and will be considered in future studies.

The ability of IMERG as forced precipitation data in the hydrological model is further assessed in Figure 46(c) and (d), respectively for two scenarios of gauge- and IMERG- calibrated model. As see, in the first scenario the IMERG tends to underestimate the streamflow (Bias of - 4.3 (CMS)) which is in agreement with the precipitation evaluation (Figure 45). Besides, low values of NSE and KGE (0.38 and 0.4, respectively) demonstrate that the simulated streamflow forced by IMERG doesn't agree well with the observations. For the second scenario, where the model is calibrated based on the IMERG input data, although the simulated streamflow is slightly improved (Figure 46 (d)), it still cannot capture the observed streamflow reasonably.
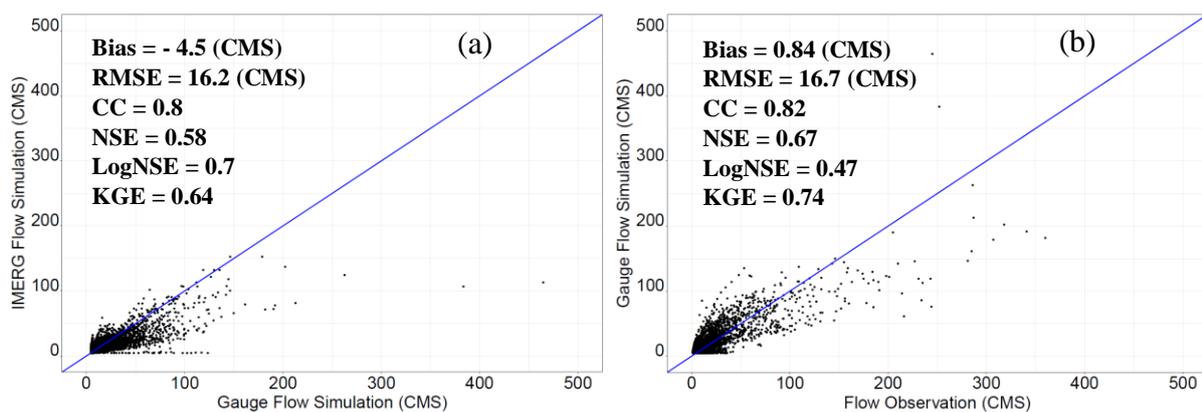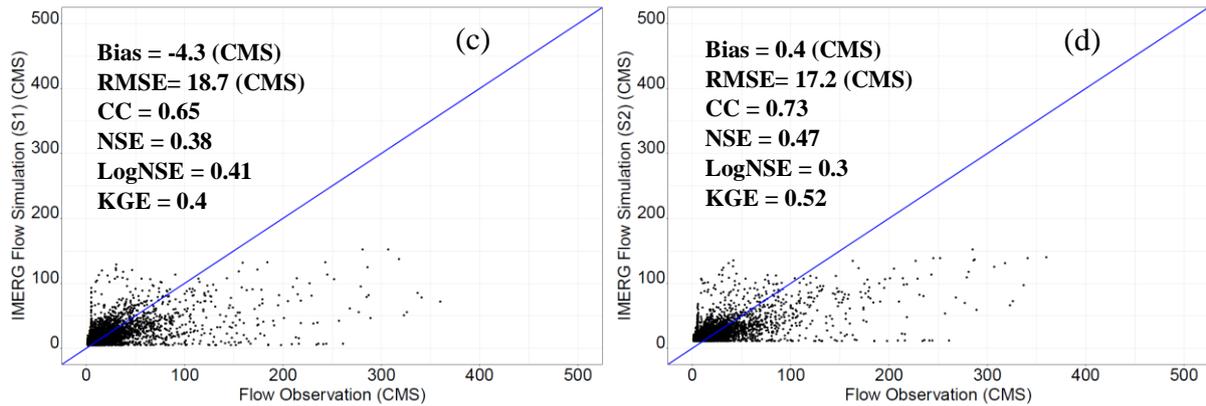
Figure 46. Comparison of (a): simulated streamflow using IMERG-based and gauge-based precipitation, (b): simulated streamflow using gauge-based precipitation and observed streamflow, (c): simulated streamflow using IMERG-based precipitation in scenario 1 and observed streamflow, (d): simulated streamflow using IMERG-based precipitation in scenario 2 and observed streamflow

Figure 47 depicts the daily time series of the observed and gauge-based simulation streamflow during 2001-201. As explained in Figure 46 (b) the simulated streamflow forced by rain gauge data with minor overestimation can capture the observed flow. The same plot is created for the IMERG-based simulation streamflow in Figure 48. This figure compares the simulated with the observed streamflow for scenario 1. Also, Figure 49 demonstrates the daily simulated streamflow for two scenarios. By comparing these two hydrographs with the observed one, it can be seen the improvement of simulation in scenario 2 rather than scenario 1.
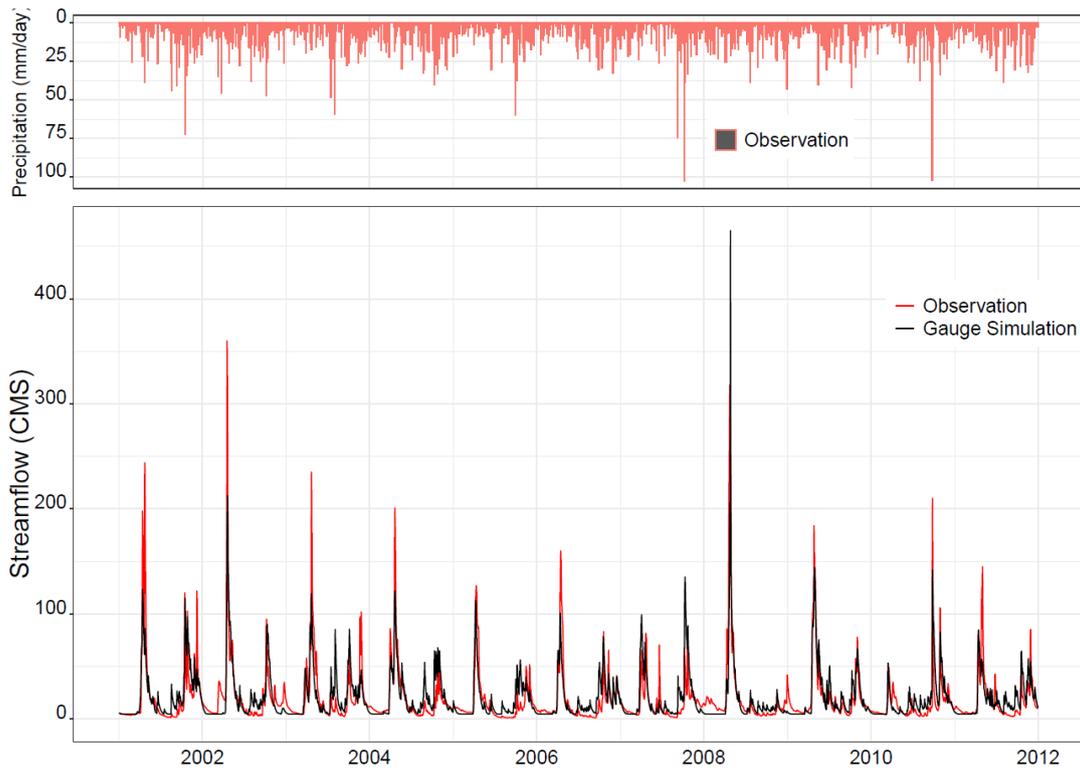
Figure 47. Comparison of the observed and gauge-based precipitation simulated daily hydrographs at the hydrometric station during 2001-2011
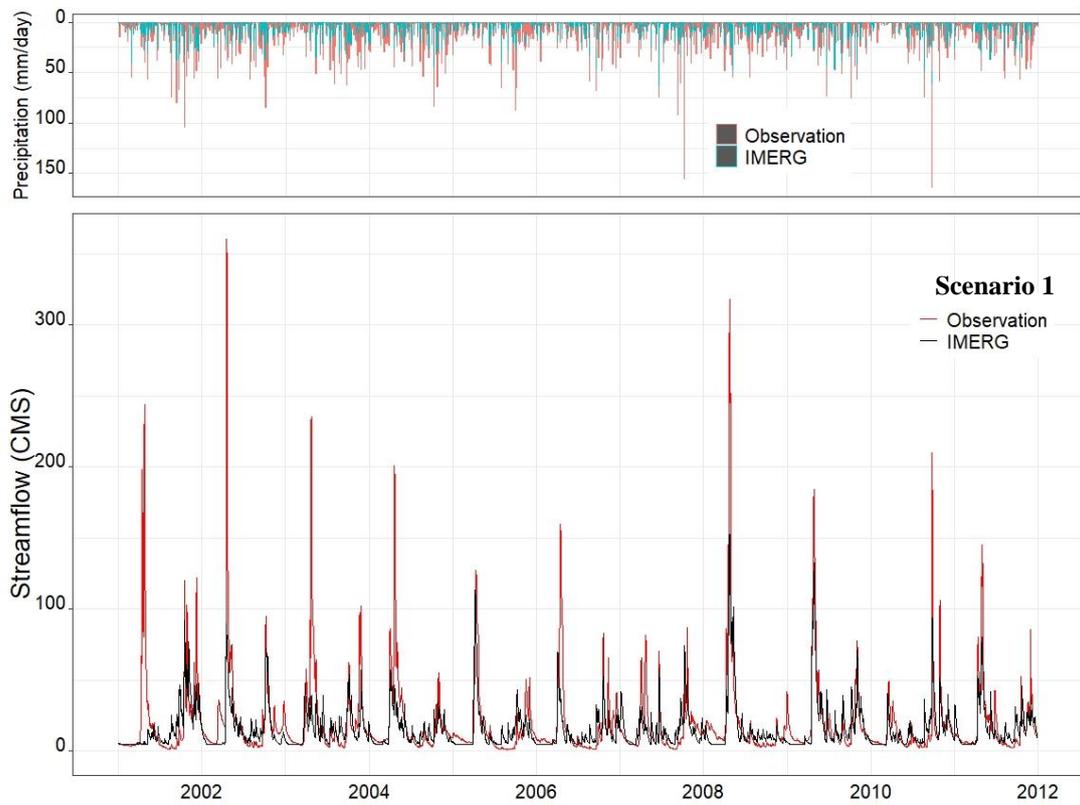
Figure 48. Comparison of the observed and IMERG-based precipitation simulated daily hydrographs in scenario 1 at the hydrometric station during 2001-2011
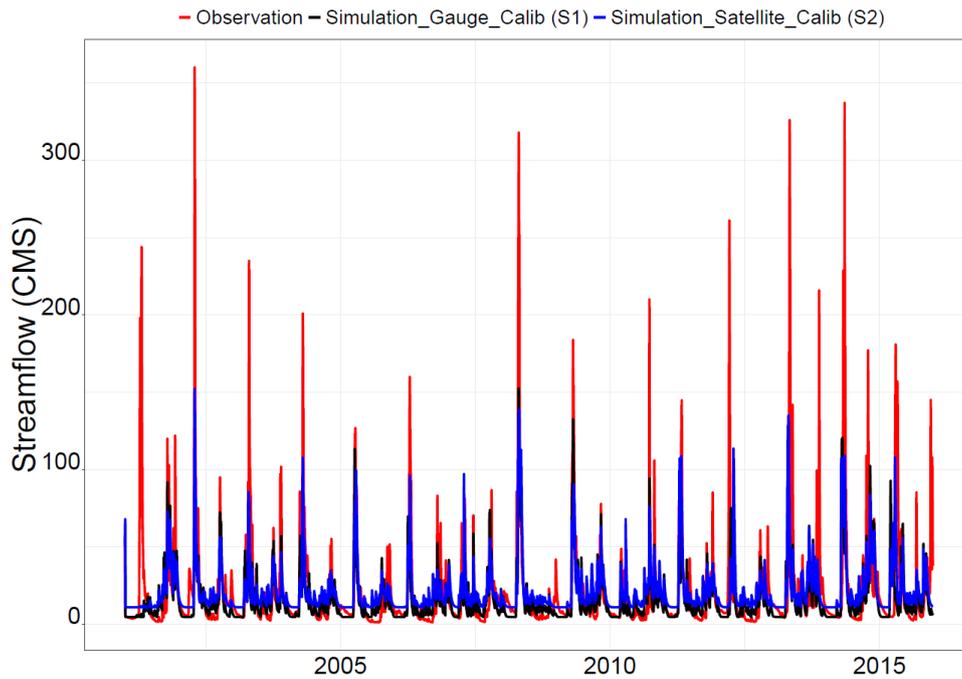
Figure 49. Comparison of the observed and IMERG-based precipitation simulated daily hydrographs in scenarios 1 and 2 at the hydrometric station during 2001-2015

The more explicit hydrographs are displayed in Figure 50 where the mean daily time series of simulated and observed streamflow during 15 years (2001-2015) is depicted. As seen the peak daily discharges for simulated and observed flows occur in the spring (April –May) in response to snowmelt. Both simulated scenarios can approximately capture the trend of discharge with higher performance related to scenario 2. However, they tend to underestimates/overestimates during spring/summer significantly.

The same results are obtained for monthly flow in Figures 51 and 52 which respectively depict the monthly hydrograph varying from 2001 to 2015 and the mean monthly hydrograph over the 15 years for the simulated and observed streamflows.
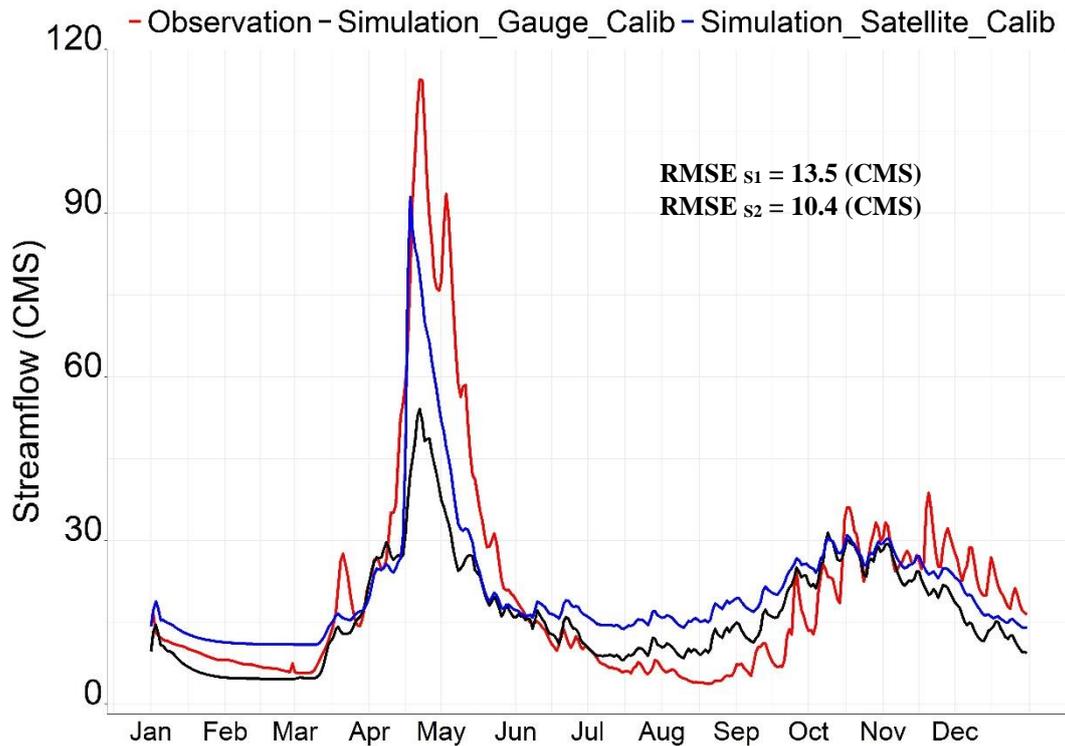
Figure 50. Comparison of the observed and IMERG-based precipitation simulated
mean daily hydrographs in scenarios 1 (black) and 2 (blue) at the hydrometric station

For a better understanding of the ability of IMERG data in extreme flow simulation, a comparison of the discharge-duration curves is performed in Figure 53. This indicates that scenario 1 of the simulated streamflow using the IMERG-based data generates the daily discharge frequency distributions that agree with the observed data for the quantile level of 0-75%. Nevertheless, both scenarios remarkably underestimate the quantiles higher than 75% and cannot capture the high flow quantiles $\geq$ 97%. On the other hand, scenario 2 tends to overestimate the quantile level of 0-75%.
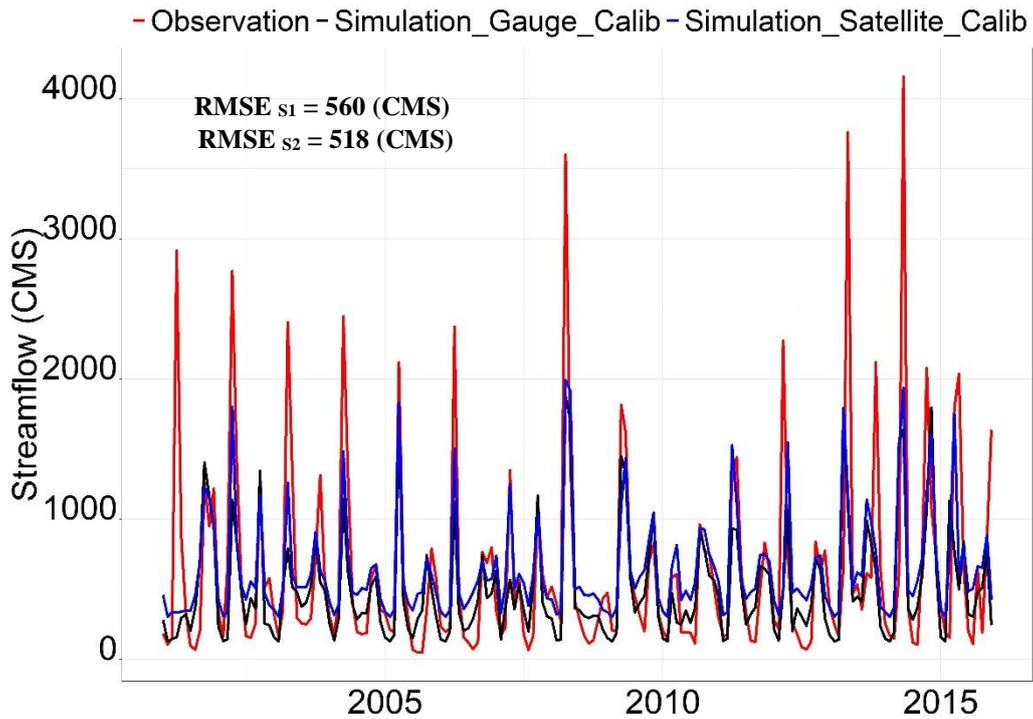
Figure 51. Comparison of the observed and IMERG-based precipitation simulated monthly hydrographs in scenarios 1 and 2 at the hydrometric station during 2001-2015
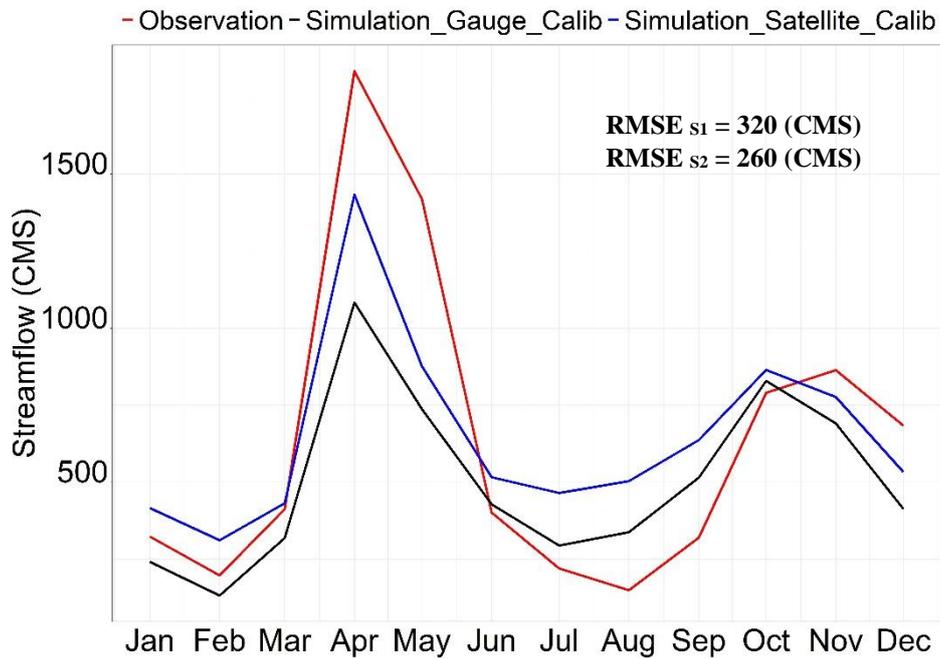
Figure 52. Comparison of the observed and IMERG-based precipitation simulated mean monthly hydrographs in scenarios 1 and 2 at the hydrometric station
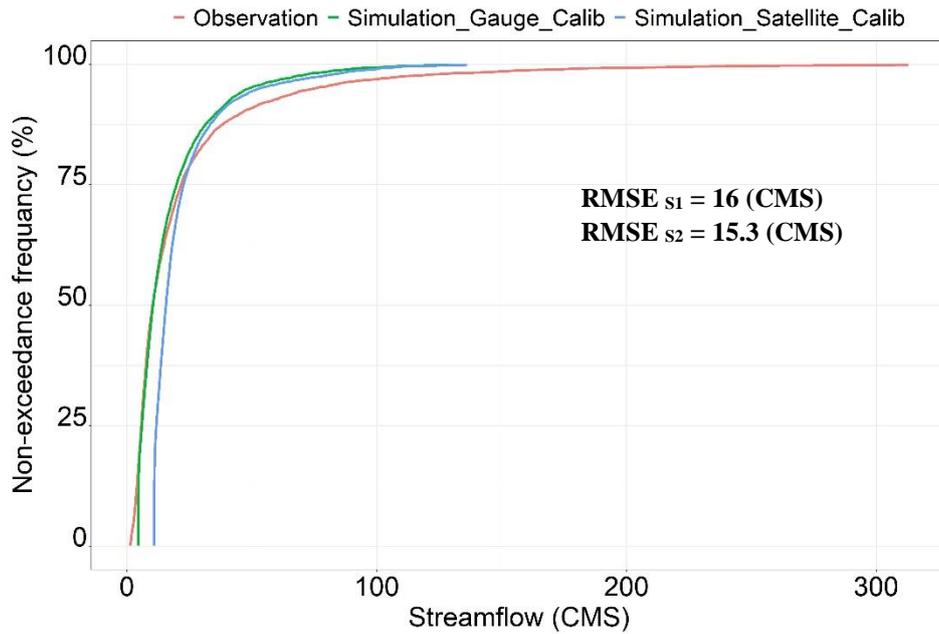
Figure 53. Comparison of the observed and IMERG-based precipitation simulated discharge-duration curves using the IMERG-based precipitation data in scenarios 1 and 2

5.6. Conclusions

This chapter evaluates the accuracy of the IMERG daily PrCal product in the hydrological application by forcing it in a calibrated model. Using the gauge-benchmarked model parameters, although IMERG-based data present unreliable simulations of daily streamflow over the watershed, it can capture the pattern of time series. Also, this study shows that calibrating the hydrological model based on satellite precipitation forcing data can increase the performance of simulated streamflow. Concerning the different levels of quantile, it can be concluded that the IMERG-based streamflow agrees with the observed one in quantile values less than 75% and notably underestimates the higher quantiles.

The considerable errors associated with the underestimates/overestimates of simulated streamflow are related to the systematic biases of satellite precipitation input data and the uncertainties of the hydrological model structure and parameters (Yuan et al., 2018). The systematic biases of IMERG products have been analyzed and discussed in Chapter 3 in more detail. However, at the study area of this chapter, the effects of IMERG systematic biases have been shown by capturing the frequency of events but not the precipitation magnitude.

Moreover, the gridded precipitation data set which is derived from 16 satellite pixels, then interpolated over the watershed might not be consistent with the only rain gauge data. Therefore, for conducting a more reliable evaluation of satellite gridded data, calibrating a distributed model over a denser basin with more ground-based data is proposed. Besides, each hydrological model has its own characteristics and procedure for computing the runoff, potentially influencing the simulation performance. Thus uncertainty analysis of the model structure is further needed to determine the portion of error related to the input precipitation data.

Overall, this chapter performs a preliminary assessment of IMERG precipitation products in daily streamflow simulation over a small watershed in the southern part of Ontario, Canada, and aims to explore the error propagation into the hydrological model. Indeed, using the IMERG data in the data-sparse or ungauged basins will have a more significant value. Nevertheless, the IMERG products still need to be improved for replacing the ground observations that provide the most accurate hydrological simulations (Su et al., 2019).

5.7. References

Craig, J.R., G. Brown, R. Chlumsky, W. Jenkinson, G. Jost, K. Lee, J. Mai, M. Serrer, M. Shafii, N. Sgro, A. Snowdon, and B.A. Tolson, 2020. Flexible watershed simulation with the Raven hydrological modelling framework, Environmental Modelling and Software, 129, 104728, doi:10.1016/j.envsoft.2020.104728.

Dinku, T., E. N.Anagnostou, andM. Borga, 2002: Improving radarbased estimation of rainfall over complex terrain. J. Appl. Meteor., 41, 1163, doi:10.1175/1520-0450(2002)041,1163:IRBEOR.2.0.CO;2.

Hao, Z., Tong, K., Liu, X., Zhang, L., 2014. Capability of TMPA products to simulate streamflow in upper Yellow and Yangtze River basins on Tibetan Plateau. Water Sci. Eng. 7 (3), 237–249.

He, Z., Yang, L., Tian, F., Ni, G., Hou, A., Lu, H., 2017. Intercomparisons of rainfall estimates for TRMM and GPM multisatellite products over the upper Mekong River basin. J. Hydrometeorol. 18, 413–430.

Jiang, S., Zhou, M., Ren, L., Cheng, X., Zhang, P., 2016. Evaluation of latest TMPA and CMORPH satellite precipitation products over Yellow River Basin. Water Sci. Eng. 9 (2), 87–96.

Li, Z., D. Yang, and Y. Hong, 2013: Multi-scale evaluation of highresolution multi-sensor blended global precipitation products over the Yangtze River. J. Hydrol., 500, 157–169, doi:10.1016/j.jhydrol.2013.07.023.

Meng, J., Li, L., Hao, Z., Wang, J., Shao, Q., 2014. Suitability of TRMM satellite rainfall in driving a distributed hydrological model the source region of Yellow River. J. Hydrol. 509, 320–332.

Pirani F.J., Najafi M.R., 2019. Recent Trends in Individual and Multivariate Compound Flood Drivers in Canada's Coasts. Water Resources Research:e2020WR027785.

Sanford S. E., Creed I. F., Tague C. L. , Beall F. D. , Buttle J. M., 2007. Scale-dependence of natural variability of flow regimes in a forested landscape. Water Resources Research, VOL. 43, W08414, doi:10.1029/2006WR005299, 2007.

Schwarb, M. (2000). The alpine precipitation climate: Evaluation of a high- resolution analysis scheme using comprehensive rain- gauge data. https://doi.org/10.3929/ETHZ- A- 004121274

Sevruk, B. (1997). Regional dependency of precipitation- altitude relationship in the Swiss Alps. Climatic Change, 36(3/4), 355–369. https:// doi.org/10.1023/A:1005302626066

Sevruk, B., Ondráš, M., & Chvíla, B. (2009). TheWMOprecipitation measurement intercomparisons. Atmospheric Research, 92(3), 376–380. https://doi.org/10.1016/j.atmosres.2009.01.016

Shafii, M. J.R. Craig, M.L. Macrae, M. C. English, S. Schiff, P. van Cappellen, and N. Basu, 2017. A diagnostic approach to constraining flow partitioning in hydrologic models using a multi-objective optimization framework, Water Resources Research, 53, doi:10.1002/2016WR019736.

Singh H., Najafi M.R., 2020. Evaluation of gridded climate datasets over Canada using univariate and bivariate approaches: Implications for hydrological modelling. Journal of Hydrology 584 (2020) 124673, https://doi.org/10.1016/j.jhydrol.2020.124673.

Su, J., Lü, H., Wang, J., Sadeghi, A.M., Zhu, Y., 2017. Evaluating the applicability of four latest satellite-gauge combined precipitation estimates for extreme precipitation and streamflow predictions over the upper yellow river basins in China. Remote Sens. 9, 1176. https://http://dx.doi.org/10.3390/rs9111176.

Su J., Lü H., Zhu Y., Cui Y., Wang X., 2019. Evaluating the hydrological utility of latest IMERG products over the Upper Huaihe River Basin, China, Atmospheric Research 225,17–29, https://doi.org/10.1016/j.atmosres.2019.03.025.

Tang, G., Ma, Y., Long, D., Zhong, L., Hong, Y., 2016. Evaluation of GPM Day-1 IMERG and TMPA Version-7 legacy products over Mainland China at multiple spatiotemporal scales. J. Hydrol. 533, 152–167.

Tian, Y., and C. D. Peters-Lidard, 2010: A global map of uncertainties in satellite-based precipitation measurements. Geophys. Res. Lett., 37, L24407, doi:10.1029/2010GL046008.

Wang, Z., Zhong, R., Lai, C., Chen, J., 2017. Evaluation of the GPM IMERG satellitebased precipitation products and the hydrological utility. Atmos. Res. 196, 151–163.

Wang, Z., Zhong, R., Lai, C., Chen, J., 2017. Evaluation of the GPM IMERG satellitebased precipitation products and the hydrological utility. Atmos. Res. 196, 151–163.

Zubieta, R., Getirana, A., Espinoza, J.C., Lavado-Casimiro, W., Aragon, L., 2017. Hydrological modeling of the peruvian-ecuadorian amazon basin using GPM-IMERG satellite-based precipitation dataset. Hydrol. Earth Syst. Sci. 21, 3543–3555.

Chapter 6

6. Concluding Remarks and Future Work

Precipitation plays a key role in hydrometeorological modeling, applications. For water resources management and accurate flood prediction, reliable quantification of precipitation estimation is crucial (Behrangi et al., 2011). However, obtaining high-quality precipitation is a challenging task especially in complex topographic and climatic conditions areas where ground-based measurement networks are either sparse or nonexistent (Caracciolo et al., 2018). Although conventional ground-based rain gauges can provide the most accurate precipitation data (Tapiador et al., 2012), alternatively satellite-based retrievals precipitation products have significantly considered at global and regional scales in recent decades for dealing with the challenges over data-sparse regions. The main advantages of satellite rainfall estimates are their global scale coverage, high spatiotemporal resolution, and short-latency which make them useful tor the requirements of flood forecasting (Su et al., 2019). Equipped by the latest Dual-Frequency Precipitation Radar (DPR) and a conical scanning multichannel, GPM Microwave Imager (GMI), and the joint application of these sensors have improved the accuracy of IMERG in detecting light and solid precipitation (Huffman et al., 2019a). In addition, for providing IMERG data, approximately 10 partner satellites are combined (Chapter2, section 2.1) as co-satellites that have effectively enhanced the spatiotemporal resolution of IMERG (Su et al., 2019). These skills lead to that the IMERG precipitation data be more accurate among other satellite precipitation products (e.g. TRMM, CMORPH, and PERSIANN). However, due to the indirect nature of such a remote sensing-based data estimators, the evaluation is needed before using them.

Similar to satellite, radar-based precipitation data provide real-time estimates of rain and snow rates at relatively fine spatial and temporal scales. Further, radar can measure the precipitation more directly by sending radio waves and receiving their reflectance back to the transmitting point than satellites that use cloud information for estimating precipitation. Nevertheless, radar precipitation estimates also have uncertainties due to the influence of ground clutter, beam blockage, and bright banding in the melting layer (Martinaitis et al. 2017). MRMS by

integrating several radars with a set of sensors can mitigate such deficiencies in the single-radar framework and provide more accurate diagnoses of atmospheric physical processes (Zhang et al., 2016).

In this study, in the first main chapter (Chapter 3), a comprehensive evaluation of four products (PrCal, PrUncal, PrIR, and PrHQ) of the IMERG latest version (V06) and MRMS precipitation data was performed over the entire country of Canada with diverse climatic conditions. To better understanding the capability of these products, we assessed them during different seasons by using a complete set of statistical and categorical metrics.

The hourly evaluations resulted in the better performance of PrCal and PrHQ than other IMERG products in estimating the precipitation amount and occurrence, respectively. The robustness of such sensors to detect precipitation suggests that they can be used to improve PrCal estimates. In daily evaluation compared to the hourly, IMERG performed better. For example, the obtained median Rbias and CSI of daily PrCal are 13% and 52% respectively, while they are 18% and 25% at the hourly scale. Considering spatially and temporally assessments, we concluded that IMERG represents better performance in terms of precipitation amount over most parts of the interior plains compared to the rest of the country with lower Rbias and RMSE values. However, in terms of precipitation occurrence, based on POD and FAR values the best capturing was obtained on the west and east coasts. Furthermore, according to the seasonal assessments, IMERG provides more reliable precipitation estimates during the warm months based on the correlation coefficients and categorical indices. We also realized that IMERG tends to overestimate the moderate to heavy precipitation events and shows relatively weak performance during the cold season, although it can keep the temporal and spatial variations of precipitation over most parts of Canada.

As the first study of such high-resolution radar-based precipitation estimates across Canada, MRMS indicated the overall promising performance of over its coverage area in the southern parts of the country. MRMS exhibited a higher average CC value (~ 0.6) than PrCal from IMERG (~ 0.4) and better CSI values over all regions. MRMS, however, tends to underestimate precipitation in the eastern and western parts of Canada and overestimates it in the interior

plains. Although MRMS can detect precipitation occurrence relatively well, it just covers the southern parts of Canada and can not be used for latitude higher than 55°N.

Considering the terms of the error in IMERG products, in the second main chapter of this study (Chapter 4), it was decided to propose and implement a framework for correcting the biases spatially and temporally. The method is based on the regression quantile mapping (RQM) to construct a relation between different quantile levels of precipitation and other dependent variables which is useful especially over the satellite pixels where there is no rain gauge for estimating the direct bias. This feature can result in bias correction spatially over the study area. Further, the developed method here is capable to adjust the different levels of rainfall intensities instead of considering just the mean or median values applied before in the conventional methods. Also, by using the time-varying covariates in RQM framework in this study, the technique can correct the biases temporally. The model parameters first were calculated by utilizing the gauged satellite pixels then interpolated over the entire study area.

The obtained findings of the bias correction method indicated significant improvement in IMERG estimates for both terms of precipitation intensity and occurrence especially over the regions where the original data had lower performances (west and coast parts of the country). The results expressed promising values where the average value of CC during the cold months from November to March was improved overall by 267% (from 0.3 to 0.8) and during the warm months from April to October by 150% (from 0.6 to 0.9). Capturing the trend of precipitation time series, consistency in lower and higher quantiles with the ground observed data, and estimating a reasonable data over ungauged regions are the main achievements of the opposed approach of bias correction in this research.

The last main chapter (Chapter 5) investigated the ability of daily IMERG precipitation as forced data in a calibrated hydrological model to further assess the performance of this widely-used remotely-sensed data in streamflow simulation. For this purpose, the Raven model calibrated by using the ground-based data over a small watershed (area of 1280 km^2) in the southern part of Ontario, Canada was selected. Although the simulated streamflow by using IMERG forced data was not reliable, it showed a reasonable capturing of the observed discharge trend as well as promising values for lower quantiles (less than 75%) of flow. In

addition to the systematic error of IMERG forced precipitation data, the uncertainty associated with the Raven hydrological model and its parameters resulted in the low quality of simulated streamflow.

It is worth mentioning that the limitations associated with the quality and sufficiency of the ground-based rain gauge data may be the main challenge of having more accurate evaluation, bias correction, and hydrological assessment of satellite IMERG precipitation products in this study. Therefore, by having a high spatiotemporal resolution network of rain gauges the method of bias correction can lead to a more accurate result. Also, the technique proposed in this study can be implemented over other regions around the globe as it is case sensitive.

Based on the findings in this research the following future works are suggested.

- Develop a high quality gridded combination precipitation product of both IMERG and MRMS remotely sensed data which can result in more accurate estimations especially for short duration events and in areas that have sparse rain gauges.

- Evaluate the ability of IMERG and MRMS in snowfall estimations by using a reliable ground-based network of snow data.

- Hydrological evaluations of IMERG and MRMS forced precipitation data in a well-calibrated distributed hydrological model over different watersheds equipped with a dense network of rain gauges and assess the skills of these high-resolution data in flood forecasting in Canada.

- Evaluate other global satellite precipitation products (e.g. CMORPH, TRMM, PERSIANN, and GsMap) over Canada and compare their results with findings of this study.

- Perform uncertainty analysis of hydrological models forced by the IMERG data to further characterize the error propagation of this input data into the hydrological applications.

References

Behrangi, A., Khakbaz, B., Jaw, T.C., AghaKouchak, A., Hsu, K., Sorooshian, S., 2011. Hydrologic evaluation of satellite precipitation products over a mid-size basin. J. Hydrol. 397, 225–237.

Caracciolo D., Francipanec A.,Violab F., Valerio Notoc L., Deidda R., 2018. Performances of GPM satellite precipitation over the two major Mediterranean islands, Atmospheric Research 213 (2018) 309–322, https://doi.org/10.1016/j.atmosres.2018.06.010.

Tapiador, F.J., Turk, F.J., Petersen, W., Hou, A.Y., García-Ortega, E., Machado, L.A.T., Angelis, C.F., Salio, P., Kidd, C., Huffman, G.J., de Castro, M., 2012. Global precipitation measurement: methods, datasets and applications. Atmos. Res. 104–105,70–97.

Su J., Lü H., Zhu Y., Cui Y., Wang X., 2019. Evaluating the hydrological utility of latest IMERG products over the Upper Huaihe River Basin, China, Atmospheric Research 225,17–29, https://doi.org/10.1016/j.atmosres.2019.03.025.

Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E.J., Sorooshian, S., Tan, J., Xie, P., 2019a. Algorithm Theoretical Basis Document (ATBD) Version 06, NASA Global Precipitation Measurement (GPM) Integrated Multi satellitE Retrievals for GPM (IMERG), National Aeronautics and Space Administration (NASA).

Huffman, G.J., Bolvin, D.T., Nelkin, E.J., Tan, J., 2019b. Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation, IMERG Tech Document, National Aeronautics and Space Administration (NASA).

Martinaitis, S.M., Cocks, S.B., Qi, Y., Kaney, B.T., 2015. Understanding Winter Precipitation Impacts on Automated Gauge Observations within a Real-Time System, Journal of OF Hydrometeorology, DOI: 10.1175/JHM-D-15-0020.1

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., Grams, H., Wang, Y., CoCKs, S., Martinaitis, S., Arthur, A., Cooper, K., Brogden, J., Kitzmiller D., 2016. Multi-Radar Multi-Sensor (MRMS) Quantitative Precipitation Estimation: Initial Operating Capabilities", Bulletin of the American Meteorological Society. DOI:10.1175/BAMS-D-14-00174.1.

Appendix: List of Acronyms

CC	Pearson Correlation Coefficient
CMORPH	Climate Prediction Center (CPC) MORPHing technique
CNbias	Correct Negative bias
CONUS	CONterminous United States
CSI	Critical Success Index
DPR	Dual-frequency Precipitation Radar
DJF	December January February
DPR	Dual-Frequency Precipitation Radar
ECCC	Environment and Climate Change Canada
ECMWF	European Centre for Medium-Range Weather Forecasts
FAR	False Alarm Ratio
Fbias	False bias
GEO-IR	GEOstationary InfraRed
GPCC	Global Precipitation Climatology Center
GPM	Global Precipitation Measurement
GPROF	Goddard PROFiling algorithm
GSMaP	Global Satellite Mapping of Precipitation
Hbias	Hit bias
HQprecipitation sensors	Precipitation extracted from merging High-Quality passive microwave
IMERG-V06 Version 6	Integrated Multi-satEllite Retrievals for Global precipitation measurement
IR	InfraRed
IRprecipitation	InfraRed geostationary satellite precipitation data
JJA	June July August
JAXA	Japan Aerospace and Exploration Agency

KGE   Kling-Gupta Efficiency

NASA   National Aeronautics and Space Administration

NCEP   National Centers for Environmental Prediction

NOAA   National Oceanic and Atmospheric Administration

NSE   Nash-Sutcliffe Efficiency coefficient

NWP   Numerical Weather Prediction

MAE   Mean Absolute Error

MAM   March April May

Mbias   Miss bias

MRMS   Multi-Radar Multi-Sensor

MW   MicroWave

PERSIANN   Precipitation Estimation from Remotely Sensed Information using Artificial
Neural Networks

PMW   Passive MicroWave

POD   Probability of Detection

PrecipitationCal   Calibrated Precipitation of IMERG

PrecipitationUncal   Uncalibrated Precipitation of IMERG

PrCal   PrecipitationCal

PrHQ   HQprecipitation

PrIR   IRprecipitation

PrUncal   PrecipitationUncal

QPE   Quantitative Precipitation Estimation

Rbias   Relative bias

RCS   Reference Climate Stations

RMSE   Root Mean Square Error

RSE   Remotely Sensed Estimate

SHSR   Seamless Hybrid Scan Reflectivity

SON   September October November

SPP   Satellite Precipitation Product

SPR   Surface Precipitation Rate

SPT   Surface Precipitation Type

TBRG   Tipping Bucket Rain Gauge

TMPA   TRMM Multi-satellite Precipitation Analysis

TRMM   Tropical Rainfall Measuring Mission

UTC	Coordinated Universal Time
VPR	Vertical Profile of Reflectivity
WSR-88D	Weather Surveillance Radar-1988 Doppler

Curriculum Vitae

Name: Saber Moazamigoodarzi

Post-secondary Education and Degrees:
School of Engineering, Ghazvin International University
Ghazvin, Iran
1999-2003 B.A.

K.N.Toosi University of Technology
Tehran, Iran
2004-2007 M.S.

Western University
London, Ontario, Canada
2019-2020 M.S.

Honours and Awards:
Scholarship for Excellent Student
2010-2011, 2012-2013

Related Work Experience
Teaching Assistant
Western University
2019-2020

Publications:
Moazami S. Najafi, M.R. (2020). A Comprehensive Evaluation of GPM-IMERG V06 and MRMS with Hourly Ground-Based Precipitation Observations across Canada. Journal of Hydrology, (Under Review).

Moazami S. Najafi, M.R. (2020). Spatiotemporal bias Correction of IMERG satellite precipitation data across Canada, (In preparation).

Conference:
Moazami S., Najafi, M.R., Struminski-Bodden, A. (2019). Evaluation of Remotely Sensed Quantitative Precipitation Estimations over Canada. 27th IUGG General Assembly, July 8-18, 2019, Montreal, Canada.

Moazami S. Najafi, M.R. (2020). Evaluation of the high-resolution spatiotemporal IMERG V06 and MRMS precipitation estimates over Canada. 54th CMOS Congress, May 26 to June 15, Ottawa, Canada.