Reproducibility of Monsoon Precipitation over the Central Himalaya in Numerical Models

Doctoral Thesis

Reproducibility of Monsoon Precipitation over the Central Himalaya in Numerical Models

（数値モデルによる中央ヒマラヤにおけるモンスーン降水の再現性）

KADEL, Indira
（カデル・インディラ）

Department of Geophysics
Graduate School of Science
Tohoku University
September, 2017

Doctoral Thesis

Reproducibility of Monsoon Precipitation over the Central Himalaya in Numerical Models

（数値モデルによる中央ヒマラヤにおけるモンスーン降水の再現性）

KADEL, Indira
（カデル・インディラ）

Department of Geophysics
Graduate School of Science
Tohoku University

Thesis Committee Members

Associate Professor Takeshi Yamazaki (Chair, Supervisor)
Professor Toshiki Iwasaki
Professor Tadahiro Hayasaka
Professor Shuji Aoki
Professor Shinji Morimoto

September, 2017

Thesis Adviser
Assoc. Prof. Takeshi Yamazaki

Author
Indira Kadel

Reproducibility of Monsoon Precipitation over the Central Himalaya in Numerical Models

Abstract

Summer monsoon season (SMS) precipitation over the central Himalaya is a lifeline of huge population which supplies water for water resources, agriculture, energy production and industry. However, inhabitants are adversely affected by precipitation–induced natural hazards such as floods, landslides and droughts that cause huge losses of life and property, impact the Himalayan environment, and ultimately obstruct the socioeconomic development. Consequently, accurate prediction of precipitation at near and far future periods is essential to support in decision making process at various sectors. This study assesses the performance of state–of–the–art numerical models (global climate models, global operational weather models and a non-hydrostatic regional model) to predict monsoon precipitation and investigates the future projections under anthropogenic climate change.

Firstly, employing 38 Global Climate Models (GCMs) participating in the Coupled Model Intercomparison Project phase 5 (CMIP5), we assess the performance of state–of–the–art GCMs to reproduce SMS precipitation over the central Himalaya at present climate and investigate the projected changes in future under anthropogenic warming climate (Representative Concentration Pathways: RCP4.5 and RCP8.5) based on the systematically selected best models. GCMs are evaluated to reproduce annual cycles of the area averaged precipitation, the spatial pattern of SMS mean climatology and the interannual variability (IAV) of the seasonal mean precipitation at the present climate. Most of the models are reliable to reproduce the annual cycle of the mean precipitation over the central Himalaya. However, simulating the spatial distribution remains the major challenge to the models, only a few models are capable to reproduce the spatial pattern of seasonal mean climatology and IAV of seasonal mean precipitation. About 66% of the total models show wet bias. The selected best models and their multi-model means project an increase of SMS mean precipitation in mid and far future periods significantly under the both warming scenarios whereas the projected changes in variability are very uncertain. Unlike, the frequency of dry days which is projected to decrease, frequencies of moderate and heavy precipitation days are projected to increase in far future period under RCP8.5 scenario. Furthermore, the frequency and the length of active spells and the frequency of break spells are projected to increase in all future periods under RCP8.5 scenario. However, the projected length of dry spell is uncertain. The increase in mean monsoon precipitation and extreme indices can be attributed to intensified low level flow from the Arabian Sea, abundant availability of moisture and enhanced convective activities in a warmer atmosphere, and is partly balanced by weakening of upper level easterly wind over the India and the Arabian Sea. Our results suggest that precipitation–induced natural disasters will likely be increased in the future since extremes indices are projected to intensify.

Secondly, we assess the predictability of precipitation in five major global operational ensemble prediction systems (EPSs) namely the China Meteorological Administration (CMA), the Canadian Meteorological Centre (CMC), the European Centre for Medium–Range Weather Forecasts (ECMWF), United States National Centers for Environmental Prediction (NCEP), and the United Kingdom Meteorological Office (UKMO) which contribute in The Observing System Research and Predictability Experiment (THORPEX) Interactive Grand Global Ensemble (TIGGE) dataset. Daily ensemble mean precipitation forecasts for Nepal from individual EPS at short to medium range time scales are evaluated in deterministic well as probabilistic sense against the Asian Precipitation Highly Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE) project dataset during summer seasons of 2009–2012. Various verification metrics have been employed to evaluate the performance at different lead times and precipitation thresholds. Results indicate that the performance of EPSs to forecast precipitation in Nepal, a country with the complex terrain, is poor. Comparatively, the ECMWF exhibits the highest forecast skills for both quantitative precipitation forecast (QPFs) and probabilistic QPFs (PQPFs) followed by that of UKMO. Although, CMC has shown comparatively lower error, better spread-skill relationship and reliability diagram, forecasts are least sharp and the spatial correlation and discrimination ability is very poor. The skill scores to verify QPFs and PQPFs show that NCEP is the least reliable. Similarly, CMA shows large wet bias particularly for the first three lead days, and over estimates medium and heavy rain events. Furthermore, we investigate the performance of multi–model grand ensemble means by applying various methods in 2012. The results show that forecast errors are reduced in grand ensemble means than the individual EPS, with the highest improvement in bias removed grand ensemble. However, the performance of ECMWF is as good as the simple multi-model grand ensemble mean.

Thirdly, reproducibility of an extreme rainfall event that occurred on 13–15 August 2014 in Nepal in a non-hydrostatic regional model (Japan Meteorological Agency Non–Hydrostatic Model: JMA–NHM) is examined. Four experiments are performed with two sets of domain sizes (large and small) at 25km and 5km spatial resolutions. Results show that JMA–NHM is able to simulate extreme precipitation at 25km to some extent. However, the spatial coverage of the event is greatly diminished, and excessive wet bias is apparent in the eastern parts of Nepal. The large wet bias could be related to the topography, cumulus parameterization scheme, and bias inherited from the initial and boundary conditions. With the use of high resolution (5 km) model, the simulation of precipitation is not improved significantly, and even deteriorated in the case of small domain on 14 and 15 August. Additionally, the impact of domain size is insignificant in the case of 25km experiments compared to the extreme wet bias particularly in the eastern parts of Nepal. However, domain size has shown large impact in simulation in the case of 5km experiments.

In summary, precipitation over the central Himalaya is highly variable both in space and time. Most of the state–of–the–art numerical models are not sufficient to accurately reproduce the precipitation characteristics. Therefore, detail precipitation mechanisms should be understood, and numerical models need to be improved accordingly.

Table of Contents

List of Figures

List of Tables

Acknowledgements

First of all, I would like to express my sincere appreciation to Prof. Takeshi Yamazaki for his invaluable supervision, guidance, patience and unremitting support throughout this research period. He continually and convincingly conveyed the spirit of research which brought me in this stage.

I am extremely grateful and indebted to Prof. Toshiki Iwasaki for his brilliant advices and support. I am deeply thankful to Assoc. Prof. Sha Weiming and Asst. Prof. Takahiro Sasai for their insightful suggestions and encouragements. I would like to acknowledge Dr. Shin Fukui for his suggestions, guidance and technical support. I express many thanks to Mr. Muhammad Rais Abdillah for his help in programming codes. I also like to thank Mr. Yuki Kanno, Mr. Kenji Aono, Ms. Irina Melnikova, Ms. Madam Taqiyya, and all members of the Atmospheric Science Laboratory for their support. A special appreciation goes to Ms. Saeko Hasebe for her administrative support.

I would like to thank IGPAS program of Tohoku University for selecting me as a graduate student under School of Science, and MEXT (Ministry of Education, Culture, Sports, Science and Technology), Japan for the entire financial support. I am thankful to Meteorological Research Institute and Japan Meteorological Agency for providing permission to use climate model developed by them. I am grateful to Department of Hydrology and Meteorology, Ministry of Population and Environment, Government of Nepal for accepting my study leave and providing necessary meteorological data.

Finally, I would like to express special thanks to my husband, Dr. Anant Babu Marahatta for standing beside me throughout my research period. I also thank my wonderful child, Aarush, for always making me smile and his patience. I am greatly indebted to my parents and family for their continuous support and encouragement to follow my ambition. At last but not least, I acknowledge all of professors, teachers, relatives, colleagues and friends for their wonderful support during my study in Japan.

Indira Kadel

September, 2017

Chapter 1

General Introduction

1.1. Motivation and background

Precipitation, a highly variable weather parameter, serves human beings as a primary source of water on Earth. Its benefits and shortcomings vary among regions. In the central Himalaya, people depend to a great degree on precipitation for water resources *(Chalise 2002)*, agriculture *(Menon 2009)*, energy, and industrial supplies. Frequently, the residents are adversely affected by precipitation-induced natural hazards such as floods *(Mirza 2011)*, landslides *(Dahal and Hasegawa 2008; Dahal 2012)*, and drought *(Sigdel and Ikeda 2010; Wang et al. 2013)*. Furthermore, the Himalayas are regarded as a water tower for housing three major river basins namely Indus, Ganges, and Brahmaputra, which supply water to the huge population of South Asia *(Chalise 2002; Mishra 2015)*. In addition, the Himalayas are a storehouse of biological diversity, are home to endangered species and are an extremely important part of the global ecosystem *(Beniston 2003)*. Consequently, it is clear that any variability in the precipitation amount, intensity, and frequency might impart remarkable economic, social, and ecological costs in the Himalayas as well as the Indian subcontinent.

According to International Panel on Climate Change (IPCC), warming of the climate system is unequivocal. It is linked predominantly with increased anthropogenic greenhouse gases from human activities *(IPCC 2013)*. Anthropogenic greenhouse gases contribute to global mean surface warming, to change in the global water cycle, to reduction in snow and ice, and to changes in climate extremes. Moreover, anthropogenic influence is expected to contribute to observed increases in atmospheric moisture contents and changes in precipitation *(IPCC 2013)*.

Studies related to climate change must be conducted in the most vulnerable region from an environmental, social, and economic perspective. The Himalayas, a unique geographical mountainous region of the world, is extremely sensitive to climate change *(Xu et al. 2009)*. The consequences of so-called climate change have already been noticed, such as an increased number of climatic extremes *(Baidya et al. 2008; Karki et al. 2017a)*, reduced in snow and ice *(Ageta et al. 2001; Shrestha and Aryal 2011)*, as well as changing precipitation trends *(Palazzi et al. 2013, 2015; Wang et al. 2013; Panthi et al. 2015; Roxy et al., 2015)*, and variability *(Duan et al. 2006)*, particularly over the central Himalaya, Nepal. Such impacts of climate change and associated vulnerabilities are known to be further exacerbated by the global warming, and certainly hinder the pace of socio-economic development by imposing extra stresses to vulnerable mountainous communities. Moreover, precipitation-induced natural disasters such as floods, landslides and droughts cause loss of many lives and huge properties; damage crop and agricultural land and water resources every year in Nepal all of which ultimately threaten the food and water security. According to the Germanwatch Global Climate Risk Index, Nepal is ranked as 14th most affected country by extreme weather events during 1993–2012 *(Kreft and Eckstein 2013)*. Consequently, it is eminent to have the information about the future evolution of precipitation at various time scales such as short and medium range forecasts and long term future climate projections under warming climates, to assist policymakers and planners to ensure proactive adaptation measures and sustainable economic development.

In recent decades, global climate models are widely used to study the long term impact of anthropogenic climate change, and weather models are used in weather forecasting. The performances of these models vary from region to region and season to season due to various reasons such as resolutions of the global models, topography and climate of the region, and dynamics and configurations of the models *(Turner and Annamalai 2012; Rajendran et al. 2013; Sperber et al. 2013; Ramesh and Goswami 2014; Su et al. 2013)*. Many previous studies have documented the performances of these models for different regions *(e.g. Turner and Annamalai 2012; Rajendran et al. 2013; Sperber et al. 2013; Ramesh and Goswami 2014; Su et al. 2014, Sharmila et al. 2015)*. However, studies on model evaluations are limited for the central Himalayan regions. A few previous studies have documented performances of global models to simulate precipitation and assessed future projections *(Palazzi et al. 2013, 2015; Panday et al. 2015; Rajbhandari et al. 2016)*. However, these studies are either incomplete or focused only over the catchment scales. So far, no reports on the performance of the global operational models to forecast precipitation at short to medium range timescales are documented for the central Himalaya. However, there is increasing number of studies on the performances of limited area models to reproduce the climatological features of spatiotemporal variability of the precipitation *(Rupa Kumar et al. 2006; Kulkarni et al. 2013; Menegoz et al. 2013; Collier and Immerzeel 2015; Karmacharya et al. 2016; Norris et al. 2016; Karki et al. 2017b)*. Similarly, numerical studies on extreme events are also very limited *(Das et al. 2006; Shrestha et al. 2015)*. In this regards, additional studies investigating the performance of numerical models to reproduce precipitation are highly demanded.

1.2. Objectives and research design

The main purpose of this study is to investigate the capabilities of the state-of-the-art numerical models to reproduce summer monsoon season precipitation over the central Himalaya in and around the Nepal. The summer monsoon season (June–September) contributes about 80% of the total annual precipitation and accounts the highest extreme events, that is why we choose this season for the study. To achieve the purpose, we have divided this study into three main parts addressing a main research question in each. The first part answers the questions "How well the global climate can simulate the precipitation over the central Himalaya at present climate? And what is the projected changes of the precipitation characteristics under warming climate?" and is presented in chapter two. As a second part of this study, an analysis on the predictability of precipitation at short to medium range time scale in major global operational ensemble prediction systems is performed and described in chapter three. Finally, as the third part, we examine the reproducibility of an intense precipitation event in a regional climate model, and is presented in chapter four. The concise objectives of this study are:

(1) To investigate the performances of global climate models to reproduce summer monsoon season precipitation at present climate and assess the changes in future projections.

(2) To assess the predictability of precipitation on short to medium range time scales in global operational ensemble prediction systems.

(3) To investigate the reproducibility of an intense precipitation events in a regional climate model.

Chapter 2

Evaluation of global climate models in CMIP5 to reproduce precipitation at present climate and assessment of future projections under warming climate

2.1 Introduction

In recent decades, GCMs are widely used to study the response of the global climate system and its components to enhanced anthropogenic gas forcing. Precipitation is an awkward atmospheric variable. For the underlying numerical models, its treatment presents many challenges. Performance of GCMs is credible at the global scale. In fact, GCMs can elucidate large-scale climate features. Nevertheless, GCM performance at the regional scale remains variable *(Mehran et al. 2014; Ramesh and Goswami 2014; Sperber et al. 2013)* because of various factors such as topography, land-cover change, aerosols, and thermodynamic forcing *(Turner and Annamalai 2012; Rajendran et al. 2013)*. The model deficiency is more evident over the mountainous region because of inadequate representation of the topographic details and other climate-relevant features such as land cover, which are important determinants for modulating climate in the mountains *(Beniston 2003; Duan et al. 2013; Palazzi et al. 2015)*. Therefore, performances of a GCM should be evaluated before using its output for regional applications such as the study of climate change effects.

The Himalaya region receives about 80% of its annual total precipitation from the South Asian monsoon system during the months of June–September: the Summer Monsoon Season (SMS). The South Asian monsoon has been regarded as a complex, fully coupled ocean–land–atmosphere system *(Turner and Annamalai 2012)*. Precipitation is highly variable both in space and time. Most GCMs show difficulty in simulating the present South Asian monsoon precipitation climate correctly. Consequently, simulated precipitation is subjected to a substantial degree of uncertainty *(Webster et al. 1998; Turner and Annamalai 2012; Sperber et al. 2013)*. Therefore, the assessment of climate projection reliability poses a major challenge *(Ramesh and Goswami 2014)*. One widely used approach in handling model uncertainty is to apply a multi-model mean for projection analysis. However, multi-model means of all available models still carry uncertainty associated with the models that are incapable of simulating present climate and its variability. Another possible approach to treat these models' uncertainty is to choose models that can simulate the present precipitation climate and its spectrum of variability rather than a simple ensemble of all available models *(Turner and Annamalai 2012; Sperber et al. 2013; Ramesh and Goswami 2014)*.

The Indian subcontinent, including the Himalayas, receives monsoon season precipitation from the southwest (South Asian) monsoon system. However, central Himalayan monsoon precipitation is not well correlated with Indian monsoon precipitation *(Shrestha et al. 2000)*. Consequently, with regard to regional climate change perspectives, isolated studies of the future evolution of precipitation over the Himalayas must be conducted. Many earlier studies have examined state-of-the-art GCMs' capabilities to simulate historical monsoon precipitation and analyzed the future evolution in warming scenarios for the Indian monsoon *(e.g. Menon et al. 2013; Sperber et al. 2013; Ramesh and Goswami 2014; Sharmila et al. 2015)*. However, few studies have addressed the future evolution of Himalayan monsoon precipitation in warming scenarios *(Palazzi*

et al. 2013, 2015; Panday et al. 2015; Rajbhandari et al. 2016). Sharmila et al. (2015) studied performances of 20 GCMs that participated in the Coupled Model Intercomparison Project phase 5 (CMIP5) *(Taylor et al. 2012*) to simulate daily to interannual variability of Indian subcontinent monsoon precipitation assigning priority to precipitation variability over India. *Palazzi et al. (2013)* investigated the increase in monsoon season mean precipitation projection with enhanced heavy rainfall days over the Himalayas using single-model simulation. Other studies of temperature and precipitation extreme indices over the western and eastern Himalayas projected increases in wet extremes over the eastern Himalayas in future warming scenarios *(Panday et al. 2015).* Another study by *Palazzi et al. (2015)* assessed the capabilities of 32 CMIP5 GCMs to simulate the annual cycles of area-averaged precipitation over the western and central Himalayas. They found that the multi-model ensemble mean and most individual models exhibit a wet bias for all seasons. They classified the models further based on model capabilities, horizontal and vertical resolution, and aerosol representation in the model. Those results revealed that no single model or group of models best represents all precipitation characteristics considered in their study. They also analyzed the future evolution of seasonal mean precipitation, and revealed a wetter future with a gradual increase in summer monsoon rainfall over the central Himalaya in future warming scenarios *(Palazzi et al. 2015).* A recent study of climate change in a river basin in the Himalayas used statistically downscaled data from 8 CMIP5 GCMs and revealed a 14% increase during SMS precipitation by 2050 under *(Rajbhandari et al. 2016).* Studies on the evaluation of state-of-the-art GCM capabilities to simulate spatial and temporal variation of SMS precipitation over the central Himalaya as a whole remain insufficient.

In this chapter, we propose to assess the capability of CMIP5 GCMs to reproduce spatial and temporal variability of central Himalayan SMS precipitation in the present climate because we are concerned about the high spatiotemporal variability of precipitation and its consequent effects on the socio-economy of the Himalayas. We investigate future changes in the mean and variability of SMS precipitation under warming climates both at spatial and temporal scales based on systematically selected best models and their multi-model mean rather than a simple multi-model ensemble of all available models. To date, no report of the relevant literature has described a study examining the central Himalaya using models that can simulate spatiotemporal variability in present climate. This study is expected to yield valuable results that are useful to ascertain a model's ability to recreate statistics of the present condition and to elucidate the changes expected to occur in the future. Furthermore, knowledge related to model ability is valuable for researchers in selecting a suitable global model to use in regional applications. For example, information related to GCM ability in simulating present climate helps a regional climate modeler in making choices of GCM models for the generation of climate projection on a regional scale. In addition, this study provides opportunities for model developers to build on inherent strengths or address deficiencies in successive model development.

Although many factors such as model resolution, physical parameterizations, affect model performance and many variables to modulate the monsoon rainfall variability of, we only evaluate simulated precipitation. Assessing the model shortcomings and capabilities to simulate all elements of monsoon systems is beyond the scope of this study. However, this study provides some valuable information related to the GCM performance over a mountainous region.

The remainder of this chapter is organized as follows. An overview of study area, observation and model data, and methodology are described in section 2.2. Results of model evaluation and future change assessment are presented in section 2.3. Finally, conclusions and discussions are presented in section 2.4.

2.2 Study area, data and methodology

2.2.1 Study area

The central Himalaya is the world's most complex region in terms of topography, where elevations range from below 100 m to the "top of the world" (8,848 m above sea level, Mt. Everest) (Fig. 2.1). The complex topography of the region plays a crucially important role in the change of local wind circulation and precipitation characteristics *(Lang and Barros 2002; Barros et al. 2004; Anders et al. 2006)*. Precipitation throughout the region varies both temporally and spatially. The SMS receives about 50–90% of total annual precipitation from the southwest monsoon system *(Nayava 1980)*. Comparison shows that SMS precipitation is greater in southeastern parts than in northwestern parts. Northwestern areas receive considerable amounts of precipitation during winter (December–February) from western disturbance systems *(Nayava 1980)*.

SMS precipitation over the central Himalaya, particularly over Nepal and surrounding areas, is highly correlated with the southern oscillation index (SOI) at an interannual time scale *(Shrestha 2000; Shrestha et al. 2000; Sigdel and Ikeda 2012)*. In this work, the study area is designated based on the correlation of SMS mean sea surface temperature (SST) averaged over Niño3.4 region (5°N – 5°S and 120°W – 170°W) in the Pacific Ocean, and circulation in the upper troposphere over the Arabian Sea and central India (18°N – 25°N and 60°E – 80°E) with the SMS mean precipitation in the Himalayan region. Correlation was found to be significant over 26°–32° North latitude and 78°–91° East longitude (Fig. 2.2). Therefore, it was selected for the study. It is referred to as CHR hereinafter.

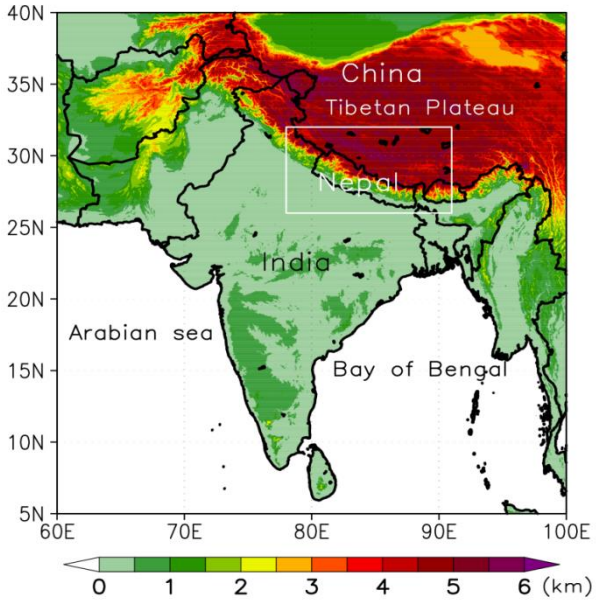


Figure 2.1. Map of South Asia. The white rectangle shows the study area. Shading shows the elevation in kilometers.

2.2.2 Observational datasets

Precipitation over the Himalayas exhibits significant spatial and temporal variations *(Lang and Barros 2002; Barros et al. 2004; Anders et al. 2006)*. The long-term pattern of precipitation over the CHR is limited because of the lack of measurements of representative spatial scales and the lack of measurements extending for more than a few years or decades. Most rain-gauge stations were established in the 1970s and 1980s *(Karki et al. 2017a)*. Existing rain-gauge networks are not sufficiently dense to reveal variability in a fine spatial scale. The gauge network density is low in the middle and high mountains because of the high elevation, severe weather, and remoteness of the regions.

Several gridded precipitation datasets derived from ground and/or satellite observations or reanalysis data are now available at varying temporal and spatial resolutions. A major difficulty related to evaluation of models is that these datasets are mutually inconsistent and are adversely affected by imperfect representation of observations *(Collins et al. 2013; Mishra 2015)*. The accuracy of the gauge-based data depends on the number of gauges, interpolation techniques, and climatology used *(Yatagai et al. 2012)*. These datasets are available for longer time periods. However, the accuracy of remotely sensed gridded data depends on the sensor instruments, algorithms used for retrieving data, sampling measurement error, and ground measurements used for calibration and adjustment of the remotely sensed precipitation *(Arakawa and Kitoh 2011; Yatagai et al. 2012)*. Satellite-based gridded data are available only from 1979 or later. Similarly, reanalysis datasets also have many sources of uncertainties. The uncertainty in reanalysis datasets derives from the atmospheric forecast model, assimilation algorithm, vertical and horizontal resolution of the model, and observational data *(Lin et al. 2014)*.

Many researchers have examined that gauge-based gridded precipitation datasets are better for analyses of the Himalayan regions than satellite and reanalysis datasets *(Andermann et al. 2011; Tong et al. 2014; Prakash et al. 2015)*. However, considerable discrepancies do exist among gauge-based precipitation datasets *(Collins et al. 2013; Palazzi et al. 2013, 2015; Mishra 2015)*. In this study, four gauge-based gridded precipitation datasets (Table 2.1) are presented to quantify the observational uncertainty over the study area. The main differences among these datasets are the number of rain gauges used, interpolation algorithms, and climatology used. The representativeness of the gridded datasets is determined largely by the number of rain gauges used. Himalayan regions have sparse gauge networks that certainly influence the quality of gridded precipitation datasets. Asian Precipitation Highly Resolved Observational Data Integration Towards the Evaluation of Water Resources (APHRODITE) have incorporated as many stations to create representative precipitation datasets *(Yatagai et al. 2012)*. The model performance is sensitive to the choice of reference data. For this study, APHRODITE is selected as a reference dataset because it is the best among the gridded datasets *(Andermann et al. 2011; Yatagai et al. 2012; Tong et al. 2014; Prakash et al. 2015)* for CHR and is available with a daily timescale. Many recent studies *(Jourdain et al. 2013; Toreti et al. 2013; Mishra et al. 2014)* have also used APHRODITE data to evaluate GCM performance.

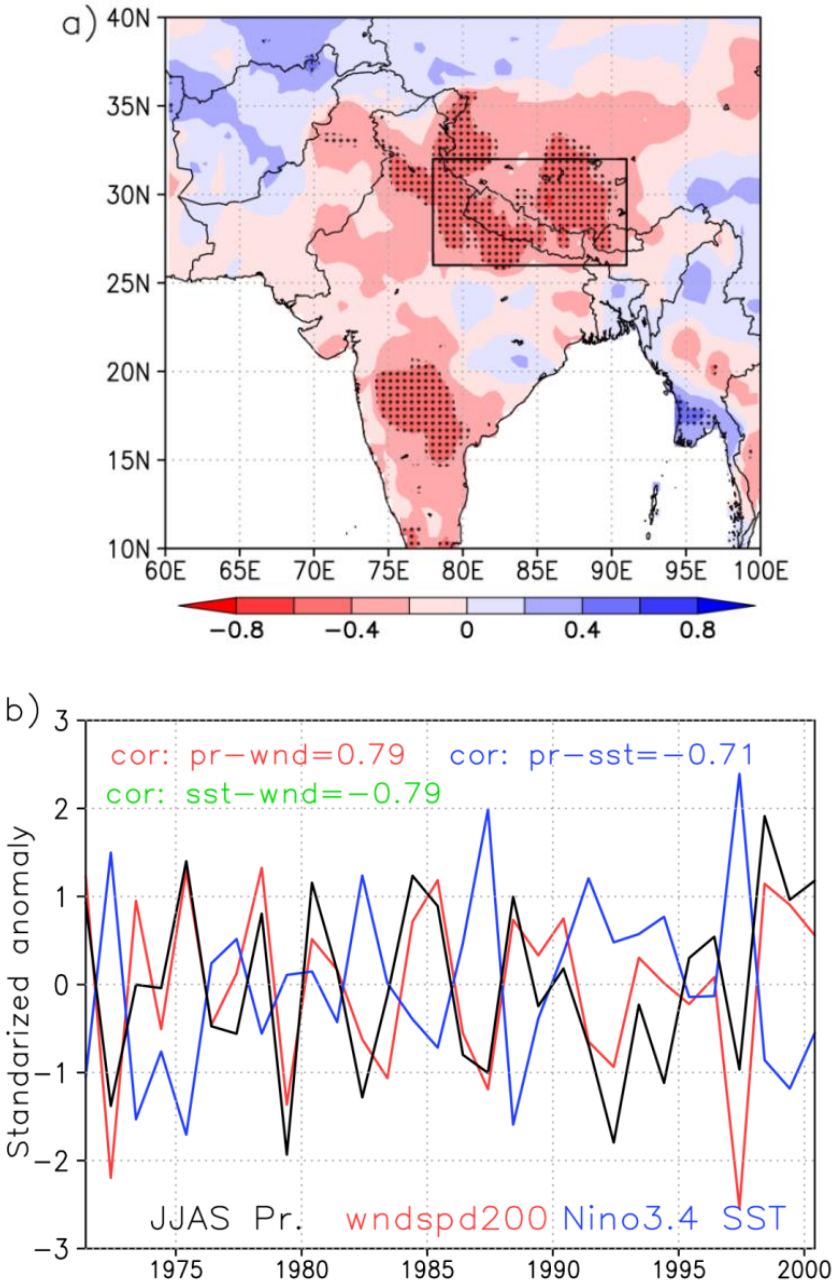


Figure 2.2. a) Spatial distribution of Pearson's correlation coefficient of SMS mean sea surface temperature (SST) over the Nino3.4 region (5 °N – 5 °S and 120 °W – 170 °W) with SMS mean precipitation anomaly, b) normalized time series of SMS mean precipitation over CHR (black), SMS mean wind speed at 200 hPa averaged over 18 °N – 25 °N and 60 °E – 80 °E (red) and SMS mean SST averaged over the Nino3.4 region (blue). Correlation coefficients of SST, wind, and precipitation are given at the top of the figure. The stipples in a) indicate significant areas at 95% confidence level. SST is derived from National Oceanic and Atmospheric Administration (NOAA) Extended Reconstructed Sea Surface Temperature version 3b (ERSST3b, Smith et al. 2008), Precipitation is from Asian Precipitation Highly Resolved Observational Data Integration Towards the Evaluation of Water Resources (APHRODITE, Yatagai et al. 2012), and wind data is from the Japanese 55-year Reanalysis (JRA55, Kobayashi et al. 2015).

Table 2.1: List of gridded precipitation datasets used for this study, resolution, source, available period and reference

Dataset	Resolution	Source	Available	Reference
APHRODITE v1101	0.5° × 0.5°/daily	Gauge	1951–2007	*Yatagai et al. 2012*
GPCC Full-data v6	0.5° × 0.5°/monthly	Gauge	1901–2010	*Schneider et al. 2014*
CRU Ts3.22	0.5° × 0.5°/monthly	Gauge	1901–2013	*Harris et al. 2014*
CPC v1.0	0.5° × 0.5°/monthly	Gauge	1948–prest.	*Chen et al. 2002*

2.2.3 Model datasets

Monthly mean precipitation simulated by 38 state-of-the-art GCMs participated in CMIP5 (Table 2.2) *(Taylor et al. 2012)* are used to evaluate the model performance in simulating current climatology. Although the simulation period is approximately 1850–2005 for historical runs, evaluations are performed during 1971–2000 corresponding to the establishment of observation network over the study region. To assess changes in the mean state and variability of precipitation in future periods, projections of precipitation under warming scenarios are analyzed for three tri-decadal periods, 2011–2040 (hereinafter near-future), 2041–2070 (hereinafter mid-future), and 2070–2099 (hereinafter far-future) of the twenty-first century. Projection data for two warming scenarios, Representative Concentration Pathways (RCP) 4.5 and RCP8.5, are used. Models are selected based on the availability of RCP4.5 and RCP8.5 scenario data. Because of the unavailability of the multiple ensemble runs for all models, we used monthly precipitation of each model from one ensemble member (r1i1p1) for all study periods. To investigate the associated change in atmospheric variables with the change in precipitation, we have also analyzed the wind, temperature, specific humidity, and vertical velocity.

Additionally, we have analyzed daily precipitation simulation (r1i1p1) from the selected best models to investigate change in daily precipitation frequency, and active and break spells under RCP8.5. All model outputs are freely available at the Earth System Grid Federation (ESGF) archive through ESGF data portals (http://cmip-pcmdi.llnl.gov/cmip5/data_getting_started.html). Brief information of CMIP5 models used for this study is presented in Table 2.2. Some of these models share their components (atmosphere, ocean, sea-ice). The analysis of genealogy of the models is available in *Knutti et al. (2013)*. Additional details of models and experiments are available in a report by *Taylor et al. (2012)*, and online at http://cmip-pcmdi.llnl.gov/cmip5/guide_to_cmip5.html .

Table 2.2: List of CMIP5 models used in this study with their respective modeling institution, model abbreviation, horizontal and vertical resolution of atmosphere and ocean models.

Modeling Institution, country	CMIP5 Model	AGCM Resolution	OGCM Resolution
Commonwealth Scientific and Industrial Research Organization (CSIRO) and Bureau of Meteorology (BOM), Australia	ACCESS1.0	N96 L38	1° × 0.33-1° L50
	ACCESS1.3	N96 L38	1° × 0.33-1° L50

Beijing Climate Center, China Meteorological Administration, China	BCC-CSM1.1	T42 L26	1° × 0.33-1° L40
	BCC-CSM1.1m	T106 L26	1° × 0.33-1° L40
Beijing Normal University, China	BNU-ESM	T42 L26	1° × 0.33-1° L50
National Centre for Atmospheric Research, USA	CCSM4	1.25° × 0.9° L26	1.125° × 0.27-0.54° L60
	CESM1-BGC	1.25° × 0.9° L26	1.0° × 1.0° L60
	CESM1-CAM5	1.25° × 0.9° L30	1.125° × 0.27-0.54° L60
Centro Euro-Mediterraneo sui Cambiamenti Climatici, Italy	CMCC-CM	T159 L31	2° × 0.5-2° L31
	CMCC-CMS	T63 L95	2° × 0.5-2° L31
Centre National de Recherches Meteorologiques and Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique, France	CNRM-CM5	T127 L31	1.0° × 0.33-1.0° L42
Commonwealth Scientific and Industrial Research Organisation in collaboration with Queensland Climate Change Centre of Excellence, Australia	CSIRO-Mk3.6.0	T63 L18	1.875° × 0.9375° L31
Canadian Center for Climate Modelling and Analysis, Canada	CanESM2	T63 L35	1.41° × 0.94° L40
EC-EARTH consortium published at Irish Centre for High-End Computing, Ireland	EC-EARTH	T159 L62	1.0° × 0.33-1.0° L42
LASG (Institute of Atmospheric Physics) and CESS (Tsinghua University), China	FGOALS-g2	2.8125° × 2.8125° L26	1° × 0.5-1° L30
The First Institute of Oceanography, State Oceanic Administration, China	FIO-ESM	T42 L26	1.125° × 0.27-0.64° L40
NOAA Geophysical Fluid Dynamics Laboratory, USA	GFDL-CM3	C48 L48	1.0° × 0.33-1.0° L50
	GFDL-ESM2G	2.5° × 2.0° L24	1.0° × 0.33-1.0° L63
	GFDL-ESM2M	2.5° × 2.0° L24	1.0° × 0.33-1.0° L50
NASA Goddard Institute for Space Studies, USA	GISS-E2-H	2.5° × 2.0° L40	1° × 0.2-1° L26
	GISS-E2-H-CC	2.5° × 2.0° L40	1° × 0.2-1° L26
	GISS-E2-R	2.5° × 2.0° L40	1.25° × 1° L32

	GISS-E2-R-CC	2.5° × 2.0° L40	1.25° × 1° L32
National Institute of Meteorological Research, Korea Meteorological Administration, Korea	HadGEM2-AO	N96 L60	1° × 0.33-1° L40
Met Office Hadley Centre, UK	HadGEM2-CC	N96 L60	1° × 0.33-1° L63
	HadGEM2-ES	N96 L38	1° × 0.33-1° L40
Institute of Numerical Mathematics, Russia	INMCM4	2° × 1.5° L21	1° × 0.5° L40
Institut Pierre Simon Laplace, France	IPSL-CM5A-LR	3.75° × 1.875° L39	2.0° × 0.5-2.0° L31
	IPSL-CM5A-MR	2.5° × 1.25° L39	2.0° × 0.5-2.0° L31
	IPSL-CM5B-LR	3.75° × 1.875° L39	2.0° × 0.5-2.0° L31
University of Tokyo, National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology, Japan	MIROC5	T85 L40	1.4° × 0.5-1.4° L50
	MIROC-ESM	T42 L80	1.4° × 0.5-1.7° L44
	MIROC-ESM-CHEM	T42 L80	1.4° × 0.5-1.7° L44
Max Planck Institute for Meteorology, Germany	MPI-ESM-LR	T63 L47	GR15L40
	MPI-ESM-MR	T63 L95	TP04 L40
Meteorological Research Institute, Japan Meteorological Agency, Japan	MRI-CGCM3	TL159 L48	1.0° × 0.5° L51
Norwegian Climate Centre, Norway	NorESM1-M	2.5° × 1.9° L26	1.125° × 1.125° L53
	NorESM1-ME	2.5° × 1.9° L26	1.125° × 1.125° L53

2.2.4 Methodology

2.2.4.1 Evaluation and selection of best models

To evaluate the model capabilities against the observations, all datasets are regridded to common resolution (0.5° × 0.5°) of the reference data using bilinear interpolation, which is a common regridding technique *(e.g. Lin et al. 2014; Mishra et al. 2014; Sharmila et al. 2015)*.

Ranking the GCM performance is an enormously difficult exercise. Several techniques and measures have already been practiced to evaluate GCM performance *(e.g. Chadwick et al. 2013; Kumar et al. 2013; Sperber et al. 2013; Mehran et al. 2014; Palazzi et al. 2015; Sharmila et al. 2015)*. No individual evaluation technique or performance measure has been regarded as a superior one. Many previous studies have used Taylor

diagrams *(Taylor 2001)* to evaluate GCM capabilities to simulate precipitation and other variables *(e.g. Inoue and Ueda 2011; Ogata et al. 2014; Sharmila et al. 2015)*. These diagrams are suitable for evaluating multiple aspects of climate models because they provide a succinct statistical summary of how closely a pattern of a test variable matches with reference variables in terms of their correlation coefficient (CC), centered root-mean-square difference (RMSD) and the ratio of their variations *(Taylor 2001)*. For this study, we used a normalized Taylor diagram in which the radial distance from the origin indicates the normalized standard deviation (NSD) of the variable and in which the azimuthal position gives the CC between test and reference variables. Similarly, the distance between the test and reference variable indicates the normalized centered RMSD (NRMSD). A model is said to be reliable if the test variable lies close to the reference variable in the diagram. For clear scientific analysis, we have defined criteria such that a model is regarded as reliable if CC ≥ 0.6, 0.5 ≤ NSD ≤ 1.5 and NRMSD < 1.0.

The mean monsoon precipitation and circulation influence the monsoon variability on all timescales. Such variability cannot be simulated correctly without correct representation of the mean state and seasonal cycle *(Sperber and Palmer 1996; Turner et al. 2005; Annamalai et al. 2007)*. Therefore, Taylor diagrams are constructed for three metrics: 1) the spatial pattern of SMS mean precipitation climatology, 2) the annual cycle of area averaged precipitation, and 3) the spatial pattern of interannual variability (IAV). IAV is measured by the interannual standard deviation of SMS mean precipitation. Finally, the best models are selected. We have regarded a model as superior if it is reliable for simulating all three metrics.

2.2.4.2 Assessment of future change

After selecting the best models, we constructed a time series of the multi-model mean (MMM) of monthly precipitation by assigning equal weight to each of the best models. Based on best models and their MMM, the future evolution of precipitation is analyzed.

In this study, changes in mean and IAV of SMS precipitation under two warming scenarios, RCP4.5 and RCP8.5, are assessed for near-future, mid-future, and far-future periods relative to the present climatology. The changes in mean ($\bar{P}_{change}$) and IAV (σ_{change}) are computed as

$$\bar{P}_{change} = 100 \times \left(\frac{\bar{P}_{Future} - \bar{P}_{Present}}{\bar{P}_{Present}}\right) \quad \text{and} \quad \sigma_{change} = 100 \times \left(\frac{\sigma_{Future} - \sigma_{Present}}{\sigma_{Present}}\right)$$

where,

$$\bar{P} = \frac{1}{N}\sum_{i=1}^{N} P_i \quad \text{and} \quad \sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - \bar{P})^2}.$$

In those equations, P_i is the JJAS mean precipitation in the i^{th} year, i=1,2 3,…,N, where N is the total number of years used for analysis. Present is the present period (1971–2000). Future signifies future periods (2010–2040, 2041–2070 and 2070–2099).

In addition, the change in annual cycle is investigated based on the daily climatological mean of first three harmonics under both warming scenarios. Changes in the frequency of light (1–10 mm day-1), moderate

(10–40 mm day-1), and heavy (>40 mm day-1) precipitation days in future periods under RCP8.5 scenario are also analyzed.

Reportedly, IAV of monsoon precipitation is related significantly to monsoon active and break spells *(Krishnamurthy and Shukla 2000)*. Long break spells are also known to affect agricultural production *(Gadgil and Joseph 2003)*. Therefore, changes in active and break spell characteristics are analyzed. To identify the active and break spells, a rainfall index (RI) is created by application of 6–90 day bandpass filter to the daily area averaged (26–30° N and 80–89° E) precipitation anomaly computed by subtracting the climatological mean of the first three harmonics. Spectrum analysis was conducted to ascertain the dominant periodicity of 6–90 days period used in filtering the detrended anomaly (Fig. A.1 and section A.1 in Appendix A). The standard deviation of RI is computed for each year. Then, the climatological mean of the standard deviation (RIstd) is computed. Finally, active (break) spells are identified such that RI ≥ +1.0 (≤ -1.0) RIstd for at least three consecutive days or more for all years (Fig. 2.3a). During active (break) spells of monsoon, the precipitation anomaly is positive (negative) in foothills of the Himalayas (Fig. 2.3b–c). The precipitation pattern agrees well with previous studies of active and break spells of the Indian monsoon *(Krishnamurthy and Shukla 2000; Gadgil and Joseph 2003; Rajeevan et al. 2010)* as these reports investigated the deficient precipitation in the foothills of the Himalaya during active periods in the central India and vice versa.

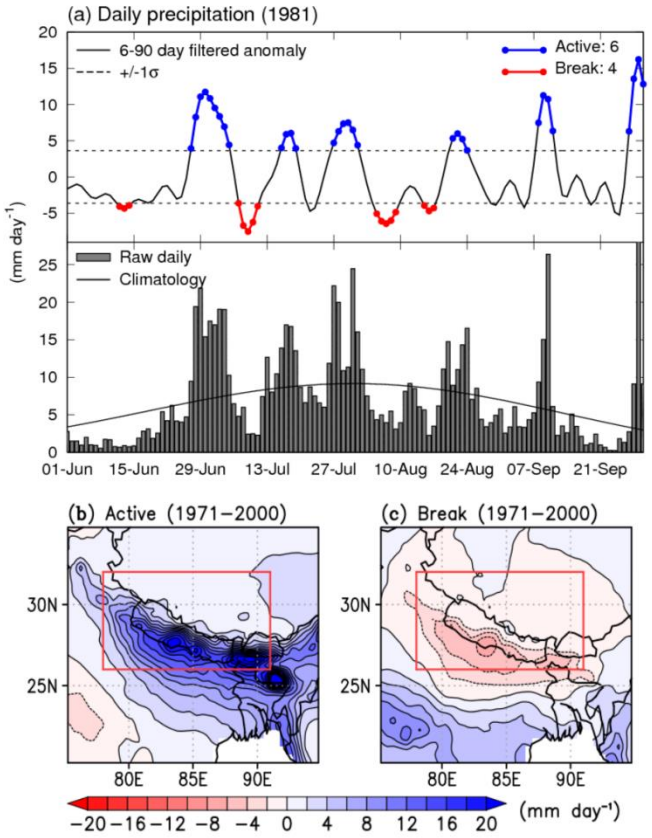


Figure 2.3. a) 6–90 day filtered precipitation anomaly (solid line in upper panel) and daily rainfall averaged over CHR (bar in lower panel) from 1 June to 30 September 1981. Dashed lines in the upper panel denote the climatological ±1 standard deviation during 1971–2000. The blue (red) line with dots in the upper panel shows active (break) spells. The solid line in the lower panel shows the climatological mean of the first three harmonics. b) and c) respectively show composites of active and break spells of monsoons precipitation. Red boxes in panels b) and c) show the study area.

2.3 Results

2.3.1 Evaluation of model performance and selection of the best models

Reliable representation of topography and physical processes in climate models are important for realistic simulation of regional precipitation. Available state-of-the-art GCMs run on different model configurations with different representations of topography and physical processes (Table 2.2, see chapter 9 in IPCC 2013). In this study, performances of 38 GCMs (Table 2.2) are evaluated for simulating CHR SMS precipitation for the present climate (1971–2000) by constructing Taylor diagrams for all three metrics: (1) annual cycle, (2) spatial pattern of SMS mean precipitation climatology, and (3) spatial pattern of IAV of SMS mean precipitation (Figs. 2.4a–c).

All the models reproduced annual cycles of mean precipitation over the CHR more or less accurately with CC greater than 0.6, except IPSL-CM5B-LR (-0.24) (Fig. 2.4a). However, only about half of the models (ACCESS1.3, BCC-CSM1.1, BCC-CSM1.1m, CMCC-CM, CMCC-CMS, CNRM-CM5, CSIRO-Mk3.6.0, CanESM2, EC-EARTH, FIO-ESM, GFDL-CM3, HadGEM2-ES, INMCM4, IPSL-CM5A-LR, IPSL-CM5A-MR, MIROC-ESM, MIROC-ESM-CHEM and MRI-CGCM3) are reliable in terms of the criteria described in section 2.2.4.1 for reproducing the annual cycle with reference to APHRODITE. Two models lying at the boundary, ACCESS1.0 and HadGEM2-CC, are also regarded as reliable models. CNRM-CM5 and the multi-model mean of all 38 models (MMM38) are very close to reference data. Skills of observational data are consistent to reproduce the annual cycle and are close to reference data values. However, these datasets have slightly higher NSD.

Results showed that simulating the spatial pattern of monsoon precipitation climatology over the complex terrain presented a major challenge for the models. They generally fail to capture observed topographically driven variability in the study region. Of the 38 GCMs, only 10 are reliable to reproduce reasonable spatial patterns of seasonal climatology (Fig. 2.4b): ACCESS1.0, CNRM-CM5, EC-EARTH, GFDL-ESM2G, GFDL-ESM2M, HadGEM2-AO, HadGEM2-CC, HadGEM2-ES, INMCM4, and IPSL-CM5A-LR.

Similarly to SMS mean precipitation climatology, simulation of the spatial pattern of IAV is found to be problematic. An inability of GCMs to represent IAV is evident *(Sharmila et al. 2015; Alves et al. 2016)*. A Taylor diagram of the spatial pattern of IAV revealed that only eight models performed well (Fig. 2.4c): ACCESS1.0, ACCESS1.3, CMCC-CMS, CNRM-CM5, EC-EARTH, HadGEM2-AO, HadGEM2-CC, and HadGEM2-ES.

Taylor diagrams revealed that neither of the models performed equally when simulating all three metrics used for this study. For example, BCC-CSM1.1m is one of the closest models for reproducing the annual cycle, but its performance is drastically reduced in reproducing the spatial pattern of climatology and IAV. Similarly, INMCM4 is reliable for simulating the annual cycle and spatial patterns of SMS precipitation, but it failed to simulate the spatial pattern of IAV. CC of MIROC5 is almost identical to other observation datasets for simulating all three metrics, but it has extremely large NSD compared to reference data. In general, the models that are reliable for representing spatial patterns are reliable for representing the annual cycle.

Consistency among most models in reproducing annual cycles shows that models are able to reproduce large domain average features rather than a spatial pattern at small/grid scale in spite of the introduction and

improvements of complex physical processes and increasing resolutions in recent versions *(Taylor et al. 2012)*. This fact underscores the need for further understanding of underlying atmospheric processes, especially in mountainous regions and improvements of the models. Not only are the model performances lacking, but the observation datasets show weak equivalence to reference data in reproducing spatial patterns. Among the observations, GPCC is close to the reference dataset. These results emphasize the necessity of reliable observation data in the study area. To overcome such difficulty in the coming days, the establishment of proper observation networks is warranted.

Additionally, we have presented the mean, bias, and inter-model spread and range of SMS precipitation over CHR during the present climate (Table 2.3). The reference dataset, APHRODITE, shows mean precipitation of 5.13 mm day-1 for 1971–2000. The inter-model range is as large as two times the reference dataset, whereas the inter-model spread is 55% of the reference dataset. Two thirds of the total models and other observation datasets have shown a wet bias. Biases of multi-model means are almost identical to the biases of observation datasets.

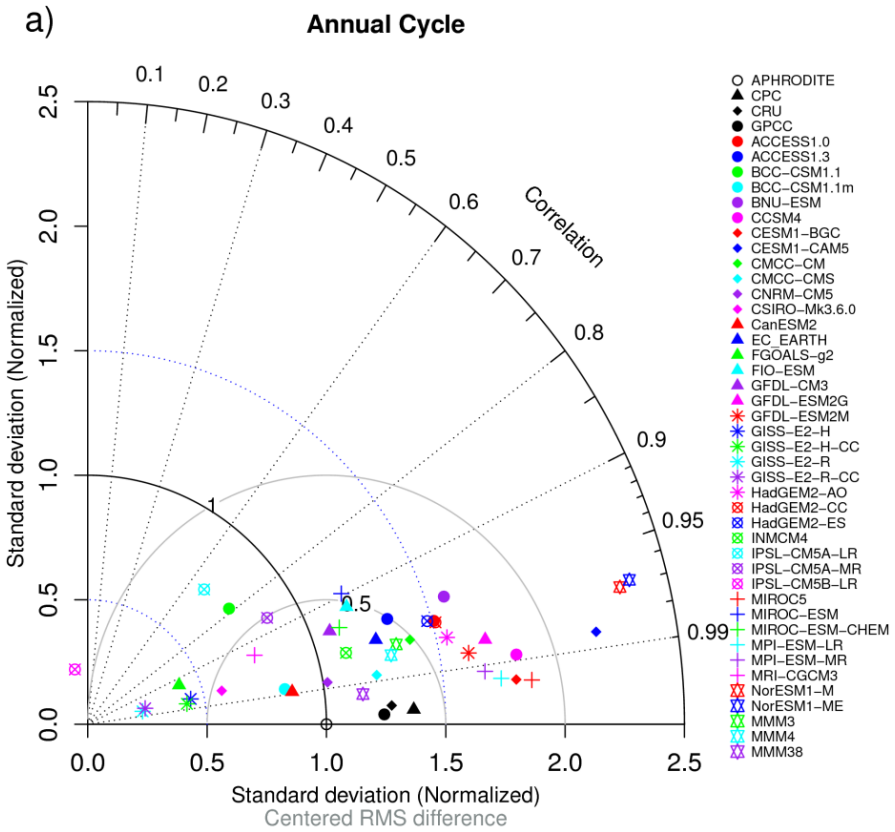
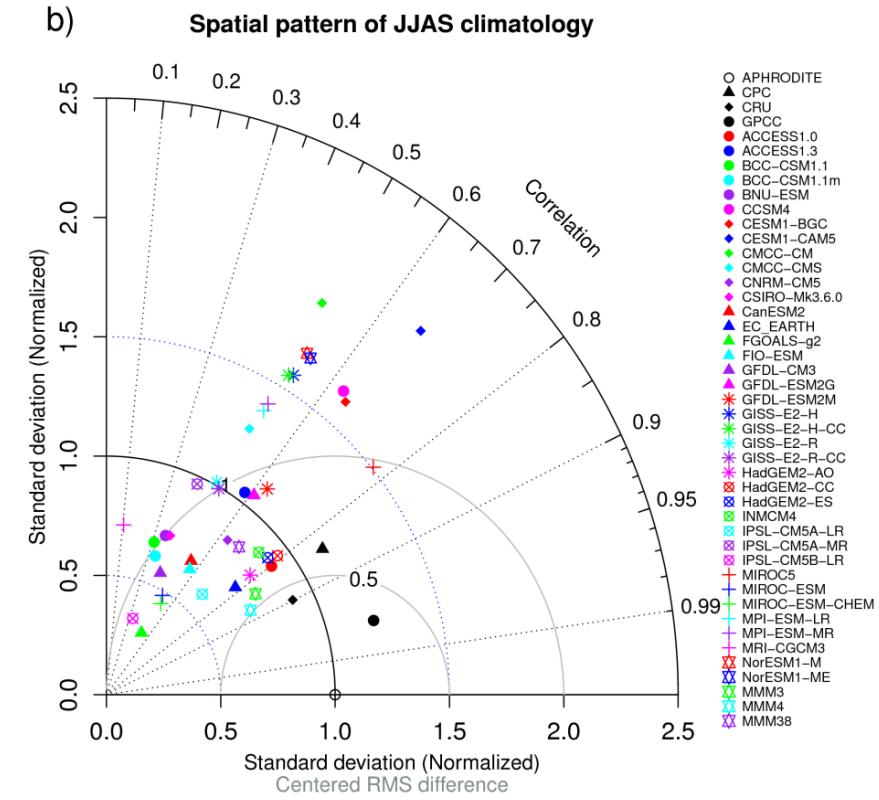


Figure 2.4. Taylor diagram of a) annual cycle of precipitation averaged over CHR, b) spatial pattern of climatological SMS mean precipitation, and c) spatial pattern of IAV measured by the standard deviation of SMS mean precipitation during 1971–2000.



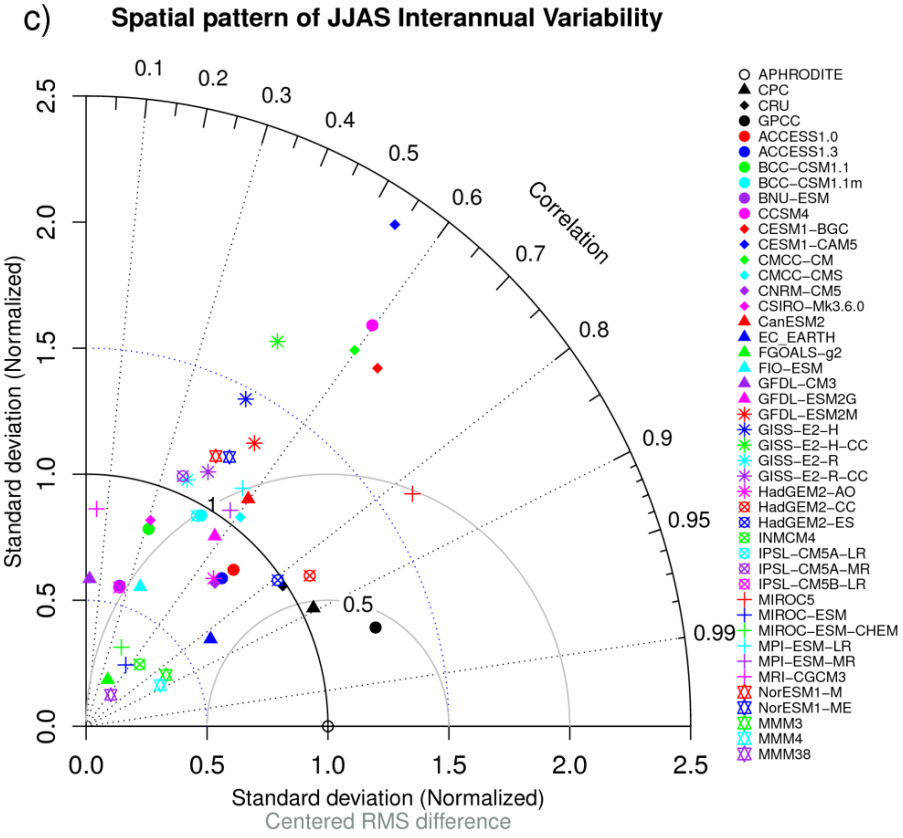


Figure 2.4. (continued).

With the background of model capabilities explained above, six models are identified as the best performing models based on the criteria defined in section 2.2.4.1: ACCESS1.0, CNRM-CM5, EC-EARTH, HadGEM2-AO, HadGEM2-CC, and HadGEM2-ES. Among them, HadGEM2 models are run with the same horizontal resolutions but with different complexities in terms of the inclusion of complex physical processes *(The HadGEM2 Development Team, 2011)*. Statistical capabilities of HadGEM2 models are almost identical (Figs. 2.4a–c) for present climatology. Consequently, only HadGEM2-ES is considered for additional study

Table 2.3: Climatological JJAS mean bias (mm day^{-1}) of precipitation averaged over CHR during historical period (1971–2000). Bias is calculated related to APHRODITE SMS mean precipitation (5.13 mm day^{-1}).

Dataset	Bias	Dataset	Bias
APHRODITE	×	GISS-E2-H	-2.50
CPC	2.01	GISS-E2-H-CC	-2.53
CRU	1.56	GISS-E2-R	-3.48
GPCC	1.36	GISS-E2-R-CC	-3.43
ACCESS1.0	2.75	HadGEM2-AO	2.93
ACCESS1.3	1.92	HadGEM2-CC	2.89
BCC-CSM1.1	-0.96	HadGEM2-ES	2.61
BCC-CSM1.1m	0.13	INMCM4	1.20
BNU-ESM	3.32	IPSL-CM5A-LR	-1.63
CCSM4	4.74	IPSL-CM5A-MR	-0.57
CESM1-BGC	4.62	IPSL-CM5B-LR	-3.39
CESM1-CAM5	6.02	MIROC5	4.92
CMCC-CM	2.07	MIROC-ESM	2.53
CMCC-CMS	2.01	MIROC-ESM-CHEM	2.43
CNRM-CM5	0.60	MPI-ESM-LR	4.06
CSIRO-Mk3.6.0	-2.11	MPI-ESM-MR	3.61
CanESM2	-0.46	MRI-CGCM3	-1.41
EC-EARTH	1.66	NorESM1-M	6.49
FGOALS-g2	-1.92	Nor-ESM-ME	6.61
FIO-ESM	1.42	MMM3	1.99
GFDL-CM3	1.09	MMM4	1.91
GFDL-ESM2G	3.47	MMM38	1.45
GFDL-ESM2M	3.52		

Inter-model standard deviation: 2.85 mm day^{-1}

Inter-model range: 10.09 mm day^{-1}

because it is better than the other two HadGEM2 models. Therefore, the four best models are selected for additional study: ACCESS1.0, CNRM-CM5, EC-EARTH, and HadGEM2-ES. Among those models, ACCESS1.0, CNRM-CM5, and HadGEM2-ES are recommended as the best models for south Asian *(Prasanna 2016)* as well as global monsoon *(Lee and Wang 2014)* assessment. EC-EARTH is also found as an excellent model for simulating central Himalaya monsoon rainfall *(Palazzi et al. 2013, 2015)*.

The four best models and their mean (MMM4) are used to study the future evolution of precipitation under RCP4.5 and RCP8.5 scenarios. However, RCP4.5 scenario datasets of EC-EARTH were not accessible at the time of analysis. Therefore, we have also constructed a multi-model mean of the three best models (MMM3) excluding EC-EARTH. Additionally, we have constructed multi-model means of all 38 models considered in this study. Skills of MMMs are portrayed in Figs. 2.4a–c. MMMs have shown better capability for the annual cycle and spatial distribution of SMS mean precipitation climatology. However, MMMs are unreliable for spatial patterns of IAV. Among MMMs, MMM38 is the best for the annual cycle but the poorest for the spatial pattern because of the inclusion of unreliable models.

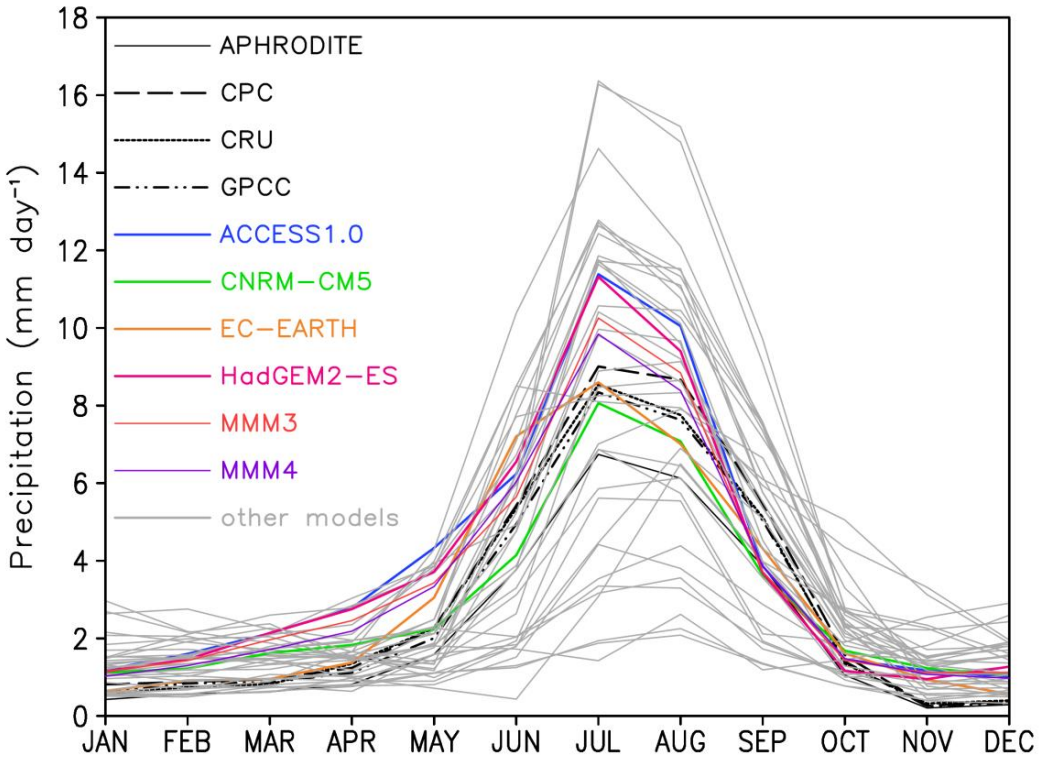


Figure 2.5. Annual cycle of precipitation (mm day^{-1}) in the CHR.

For more detail, the annual cycle of the best models and their MMMs is presented in Fig. 2.5. A large inter-model spread prevails in the shape and size of the annual cycle. The model uncertainty is greater when simulating SMS precipitation than it is for other seasons. It is noteworthy that APHRODITE precipitation is lower than those of other observation datasets as well as the best models, particularly during the monsoon season. The spatial distribution of SMS precipitation and its IAV are presented respectively in Fig. 2.6 and Fig. 2.7. The

spatial patterns of large-scale mean precipitation and IAV among the observations show remarkable consistency. Best models showed reliable capability in simulating spatial patterns of seasonal mean climatology (Fig. 2.6) and IAV to some extent (Fig. 2.7). However, local biases exist in many regions. ACCESS1.0 and HadGEM2-ES have shown the same climatology and bias as they share the atmospheric components from the same model family (Bi et al. 2013; Knutti et al. 2013). EC-EARTH has a noticeable dry bias (Fig. 2.6) and lower IAV over southeastern parts (Fig. 2.7). The MMMs of best models have shown improved capabilities for assessing mean precipitation (Fig. 2.6), but smooth out the IAV (Fig. 2.7) *(Sperber et al. 2013)*. Because IAV of MMMs is not reliable, further analysis of IAV is not performed based on MMMs data.

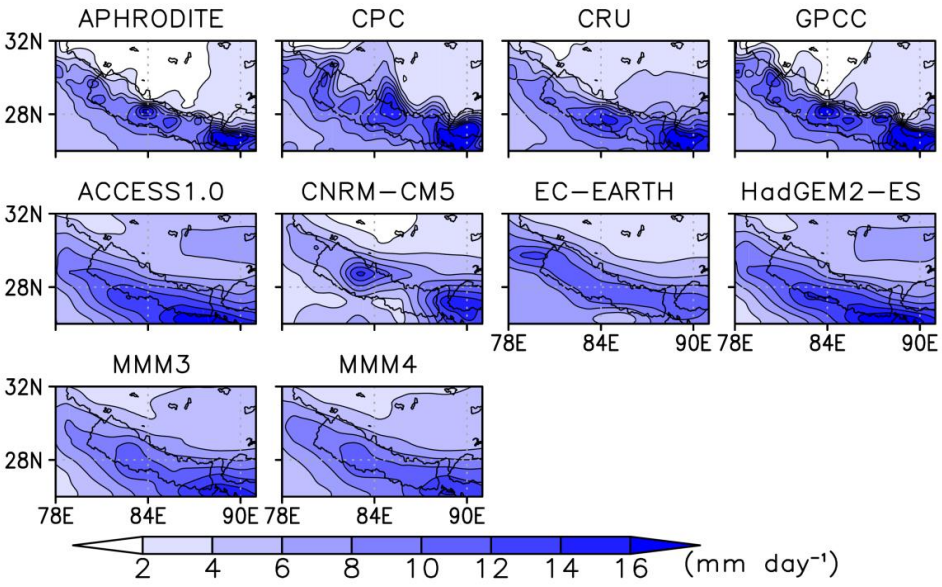


Figure 2.6. Monsoon (June–September) season mean precipitation climatology during 1971–2000.

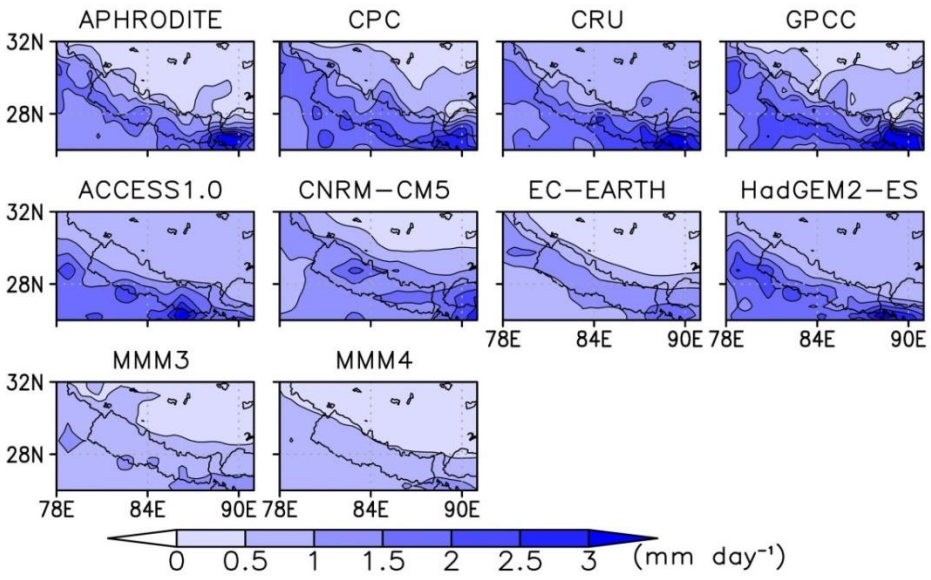


Figure 2.7. IAV of monsoon season mean precipitation during 1971–2000.

2.3.2 Projected change in future warming scenarios

2.3.2.1 Changes in SMS precipitation climatology

The annual cycle of daily precipitation (smoothed by removing the first three harmonics) of the best models and their multi-model means for the near-future (2011–2040), mid-future (2041–2070) and far-future (2070–2099) periods under RCP4.5 and RCP8.5 scenarios with their corresponding annual cycle at the present condition are depicted in Fig. 2.8. All models projected an increase in the amplitude of precipitation during the summer monsoon, June–September in all future time slices under both warming scenarios. However, no large difference exists in the amplitudes of precipitation projection between RCP4.5 and RCP8.5. The amplitude increased gradually from the near-future to far-future periods. The increase is large in July and August, with a value of about 3 mm day^{-1} in the far-future period. The winter precipitation (December–February) is projected to decrease slightly by multi-model means and member models (Fig. 2.8).

Table 2.4: Changes in mean and IAV of JJAS mean precipitation averaged over CHR. Confidence levels of 99% and 95% are denoted respectively by ** and *. The confidence level of mean and interannual variation change is tested respectively using Student's t-test and Levene's test.

Change in seasonal mean precipitation (%)

Models	2011–2040		2041–2070		2070–2099	
	RCP4.5	RCP8.5	RCP4.5	RCP8.5	RCP4.5	RCP8.5
ACCESS1.0	2.16	-0.83	9.38**	6.33**	12.67**	17.39**
CNRM-CM5	8.24	10.72**	9.84**	20.28**	17.71**	27.55**
EC-EARTH	×	7.34**	×	12.91**	×	15.45**
HadGEM2-ES	4.83	2.27	9.94**	7.68**	17.85**	20.58**
MMM3	4.76*	3.39	9.71**	10.56**	15.90**	21.27**
MMM4	×	4.35*	×	11.13**	×	19.87**

Change in IAV of seasonal mean precipitation (%)

Models	2011–2040		2041–2070		2070–2099	
	RCP4.5	RCP8.5	RCP4.5	RCP8.5	RCP4.5	RCP8.5
ACCESS1.0	-21.18	-12.20	17.72	-6.32	-1.69	36.28
CNRM-CM5	73.68*	44.84	19.61	35.77	48.10	63.72**
EC-EARTH	×	7.71	×	58.34*	×	55.71**
HadGEM2-ES	-12.61	-3.27	-1.84	15.06	-8.05	15.63

Table 2.4 presents the change in area-averaged SMS mean precipitation over CHR and its IAV for near-future, mid-future, and far-future periods under both warming scenarios relative to the present climatology. The area-averaged precipitation is projected to increase gradually from near-future to far-future periods. It is the highest for far-future period under RCP8.5. The increase is consistent during mid-future and far-future under both warming scenarios with a 99% confidence level. However, projection is not consistent for the near-future period. The change in mean precipitation is about -1% – 27%. Multi-model means projected an increase of about 20% of mean precipitation at the 99% confidence level.

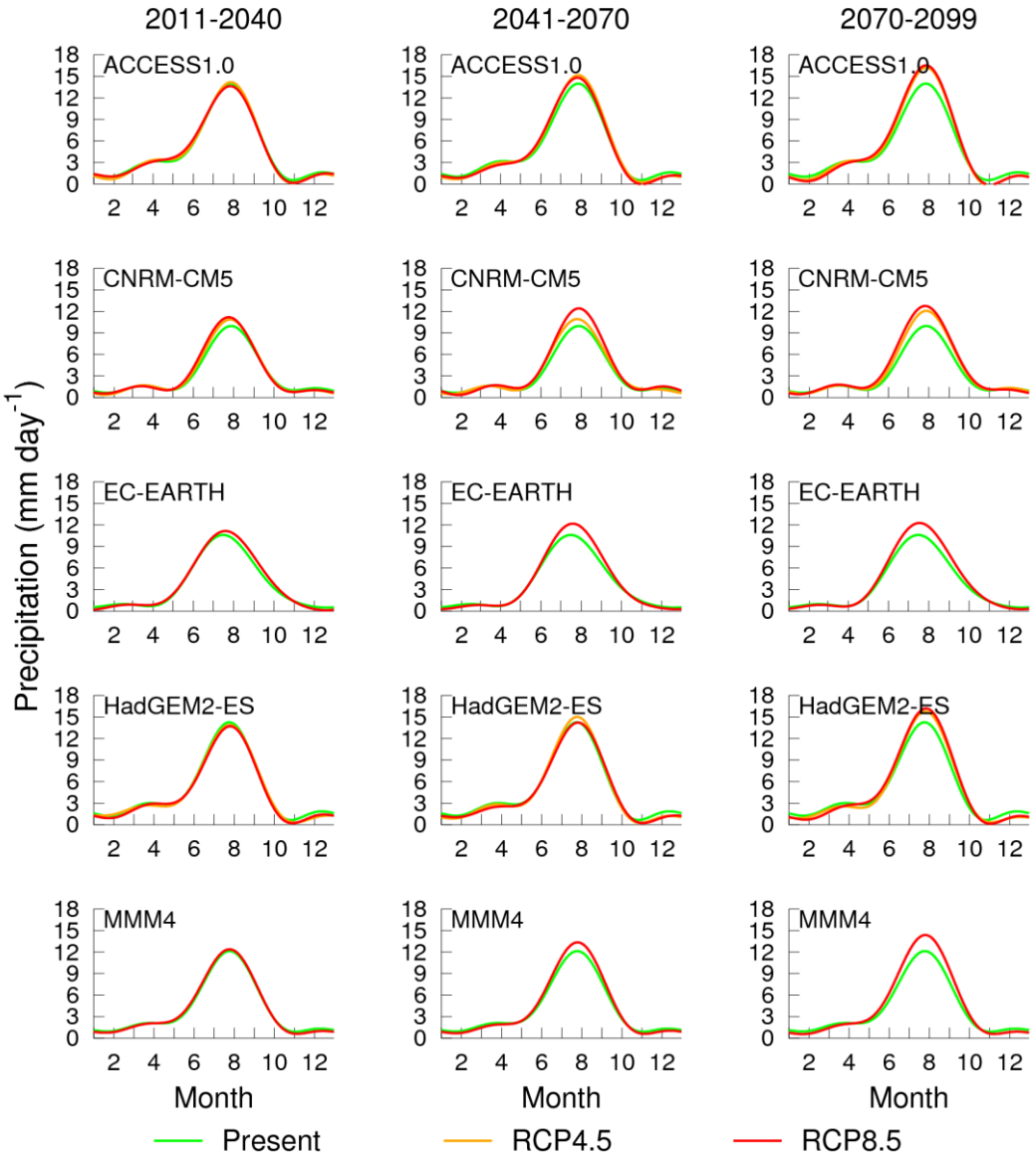


Figure 2.8. Annual cycle of best models and multi-model means of three (MMM3) and four (MMM4) best models. The left column shows 2011–2040. The middle column shows 2041–2070. The right column shows 2070–2099. The first three harmonics have been removed from each time series to facilitate a visual comparison of the phase and amplitude.

Different from mean precipitation, models projected mixed tendencies for changes in the IAV of SMS mean precipitation in all future time periods under both warming scenarios except for far-future periods under RCP8.5 scenario. Some models show a significantly increasing tendency. Some show a decreasing tendency even for the same periods under the same scenario (Table 2.4). However, models are consistent to reveal increasing variation during the far-future period under RCP8.5 scenario. Increased variation indicates irregular SMS precipitation, which strongly affects decision making in every sector from agriculture to water resource management. It is noteworthy that future projections of IAV are inconsistent. This result suggests that projection data must be considered carefully when applying to study the climate change impacts.

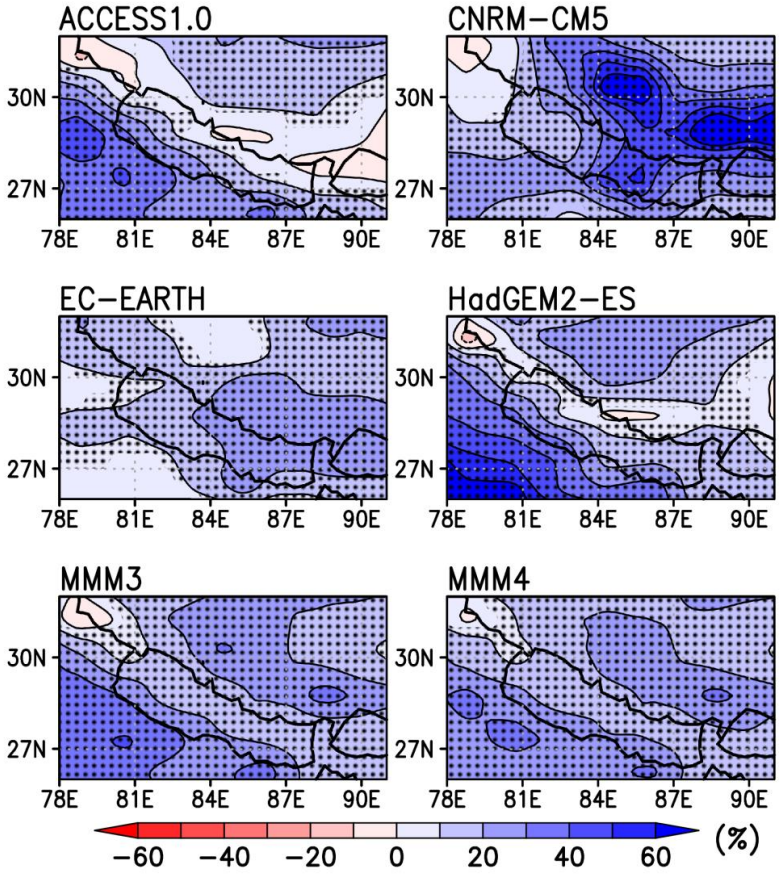


Figure 2.9. Spatial distribution of projected change in SMS precipitation during 2070–2099 under RCP8.5 relative to the present (1971–2000) climatology. Stippled areas are significant at the 95% confidence level. The level of significance was assessed using Student's t-test for equal and unequal variances based on Levene's test.

Fig. 2.9 presents relative changes in the spatial pattern of SMS mean precipitation climatology for the far-future period under the RCP8.5 scenario. Precipitation is projected to increase considerably over most parts of the CHR by all the best models and multi-model means. However, regional uncertainty exists. Most models have projected slight and not significant decreasing tendencies over northwestern parts of the study area. ACCESS1.0 and HadGEM2-ES projected similar spatial patterns of SMS mean precipitation climatology change, which possibly derives from common atmospheric models *(Bi et al. 2013; Knutti et al. 2013)*.

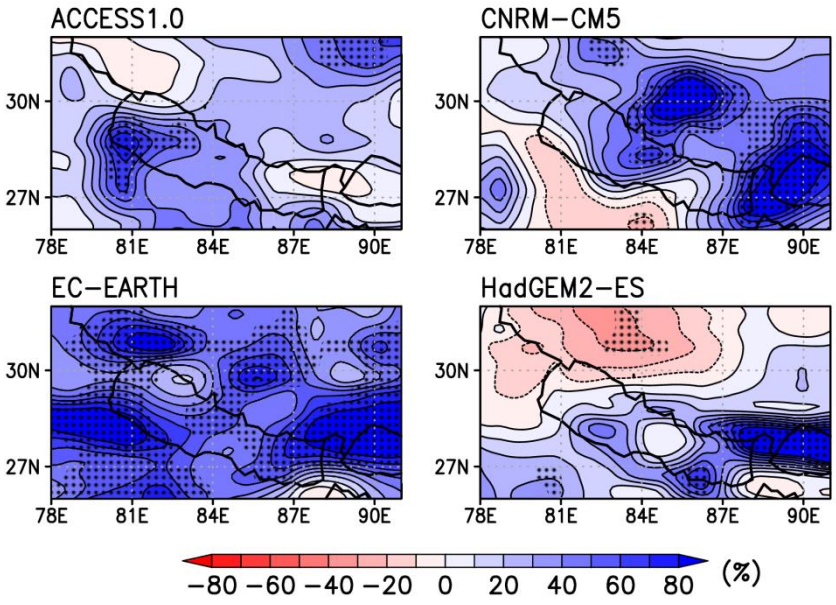


Figure 2.10. Same as figure 2.9 but for IAV. The level of significance in this case was tested using Levene's test. MMM3 and MMM4 are not shown here because these multi-model means are not reliable to simulate IAV for the present climate.

Projection of possible IAV change of SMS mean precipitation during the far-future period under RCP8.5 is presented in Fig. 2.10. The model consensus on changes in IAV spatial pattern is low, indicating considerable uncertainty. However, most of the best models consistently reveal increasing IAV over Nepal. It is noteworthy that not all increasing IAV area over Nepal is statistically significant in all models.

Projected changes in mean and variability are robust in the far-future period under the RCP8.5 scenario. Therefore, further analysis specifically addresses the projected changes only in the far-future period under the RCP8.5 scenario.

2.3.2.2 Changes in extreme indices

2.3.2.2.a Changes in daily precipitation characteristics

Changes in the frequency and intensity of extreme events are well known to impose stronger effects on socio-economy than seasonal mean changes. Mishra et al. (2014) investigated that the ensemble mean of best GCMs performed better than RCMs for simulating indices of extreme events over India. Wet extremes and daily intensity are projected to increase and rainy days are projected to decrease over the Himalayas in future warming scenarios *(Palazzi et al. 2013).* It is therefore worthwhile to analyze the changes in frequency of daily precipitation projected by the best GCMs and their MMM for the central Himalaya as well.

Fig. 2.11 portrays projected changes in frequency distribution (light, 1–10 mm day^{-1}; moderate, 10–40 mm day^{-1}; and >40 mm day^{-1}) of daily precipitation under the RCP8.5 scenario during SMS in the far-future

period. A multi-model mean of daily precipitation simulated by the four best models (MMM4) is created by assigning equal weight to each model, as with MMM4 of monthly precipitation. HadGEM2-ES has 360 (120) days in a year (SMS), with 30 days in each month. Consequently, MMM4 also has 120 days in SMS. Calculated frequencies of HadGEM2-ES and MMM4 are adjusted to 122 days, multiplying by a scaling factor (122/120) for treating all models equally. Best models and MMM4 consistently projected a decrease in light precipitation days over most parts of the study area. Moderate and heavy precipitation days are projected to increase. However, model convergence in spatial patterns of projected changes in the frequency of moderate precipitation over Nepal is low. ACCESS1.0 and HadGEM2-ES show decreased frequencies of moderate precipitation over Nepal. Decrease in light to moderate rainfall over the Himalayan foothills and Nepal might be attributable to the shift of light to moderate rainfall to heavy rainfall categories under warming scenarios. MMM4 projects an increase in moderate precipitation days by more than 6 days over large parts of the study area. The projected increase in moderate and heavy precipitation days is expected to be accompanied by a possible increase in precipitation-induced hazards such as floods and landslides.

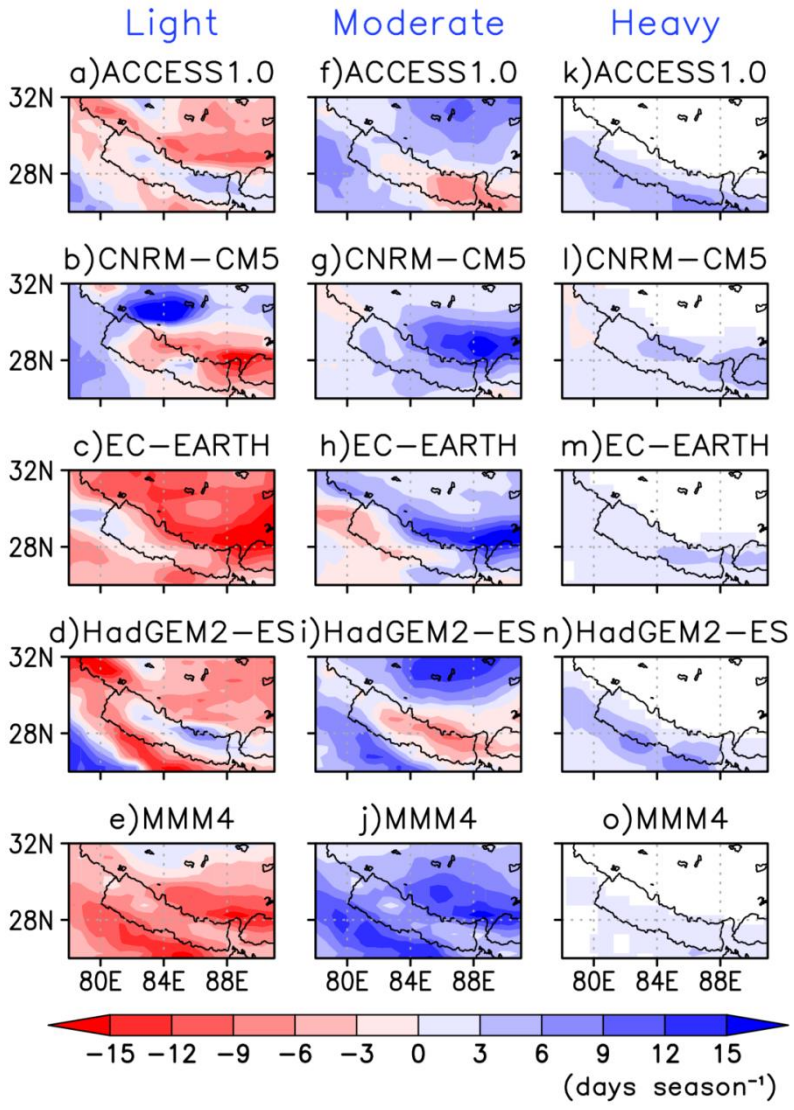


Figure 2.11. Projected changes in the frequency (days season^{-1}) distribution of daily precipitation (light, 1–10 mm day^{-1}; moderate, 10–40 mm day^{-1}; and heavy, >40 mm day^{-1}) under RCP8.5 during SMS of 2070–2099.

3.2.2.2.b Changes in active and break spells

Active and break spells of monsoon are known to affect the monsoon season total precipitation *(Rajeevan et al. 2010)*. IAV of rainfall during a monsoon season is characterized by the occurrence of active and break spells of monsoons *(Krishnamurthy and Shukla 2000)*. Break spells of long duration are known to have a strong effect on rain-dependent agricultural production *(Gadgil and Joseph 2003)*. Therefore, active and break spells of monsoons simulated by models for both the present and future under the RCP8.5 warming scenario are analyzed in this study. Best models and MMM4 are reasonable for simulating the frequency and length of active and break spells at the present climate (Figs. 2.12a–d). However, the spread in the length of break spells is greater than the length of active spells in the present climatology (Figs. 2.12b–d). Best models and MMM4 consistently projected a gradual increase in the frequency of active and break spells and the length of active spells. Frequencies of active (break) spells are projected to increase by 2 (1) in a season in the far-future period. However, the increase in the length of active spells is less than 1 day per season. Future changes in the lengths of break spells are uncertain (Fig. 2.12d).

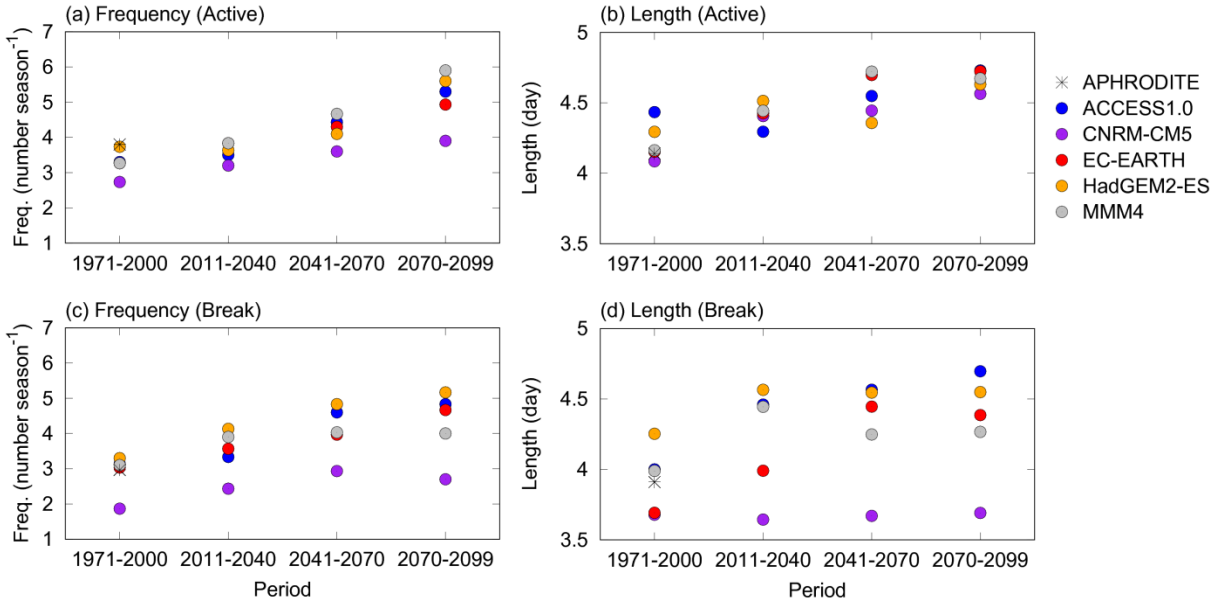


Figure 2.12. Active spell and break spell characteristics of present and future climates under the RCP8.5: a) and c) frequencies, and b) and d) length of active and break spell.

2.4 Conclusions and discussions

In this chapter, capabilities of 38 global climate models participating in Coupled Model Intercomparison Project phase 5 (CMIP5) to reproduce summer monsoon precipitation at the present climate are first evaluated systematically. An important obstacle hindering evaluation of model simulation over the Himalayas is the lack of a promising observation dataset *(Yatagai et al. 2012; Palazzi et al. 2013; Mishra 2015)*. Available gridded observed precipitation datasets show discrepancies (Figs. 2.4a–c, 2.5–2.7, Table 2.3) indicating uncertainty in the observation datasets, which further complicates model evaluation *(Ramesh and Goswami 2014)*. We have used the APHRODITE' precipitation dataset as a reference because this dataset is good for the Himalayan region

(Andermann et al. 2011; Yatagai et al. 2012; Tong et al. 2014; Prakash et al. 2015). However, the sparse observation network over the Tibetan plateau and the high Himalayas might cause reduced precipitation over the region *(Yatagai et al. 2012)*. A second issue is the problem of mapping the coarse resolution of model gridbox-average value to slightly fine resolution of APHRODITE, which is not considered here, and which would likely result in lower predictive capability *(McBride and Ebert 2000)*. Consequently, enhancement in observation network is suggested to have a representative precipitation dataset over the Himalayas.

The complex topography of the Himalayas plays a vital role in changing local wind circulation and precipitation characteristics. Consequently, reliable representation of topography and physical processes in climate models is important for the realistic simulation of regional precipitation *(Fennessy et al. 1994; Turner and Annamalai 2012; Rajendran et al. 2013)*. Most GCMs are credible to reproduce annual cycles of regional averaged precipitation (Fig. 2.4a) reasonably well, even though the inter-model spread is extremely large during the monsoon season. Realistic simulation of spatial distribution of the mean monsoon season precipitation and its interannual variability (IAV) is problematic (Fig. 2.4b–2.4c). The IAV of precipitation significantly depend on the horizontal resolution of the models *(Giorgi 2002)*. The inabilities of GCMs in reproducing the spatial distribution of precipitation at regional level are also described in reports of previous studies *(Turner and Annamalai 2012; Sperber et al. 2013; Su et al. 2013; Sharmila et al. 2015)*.

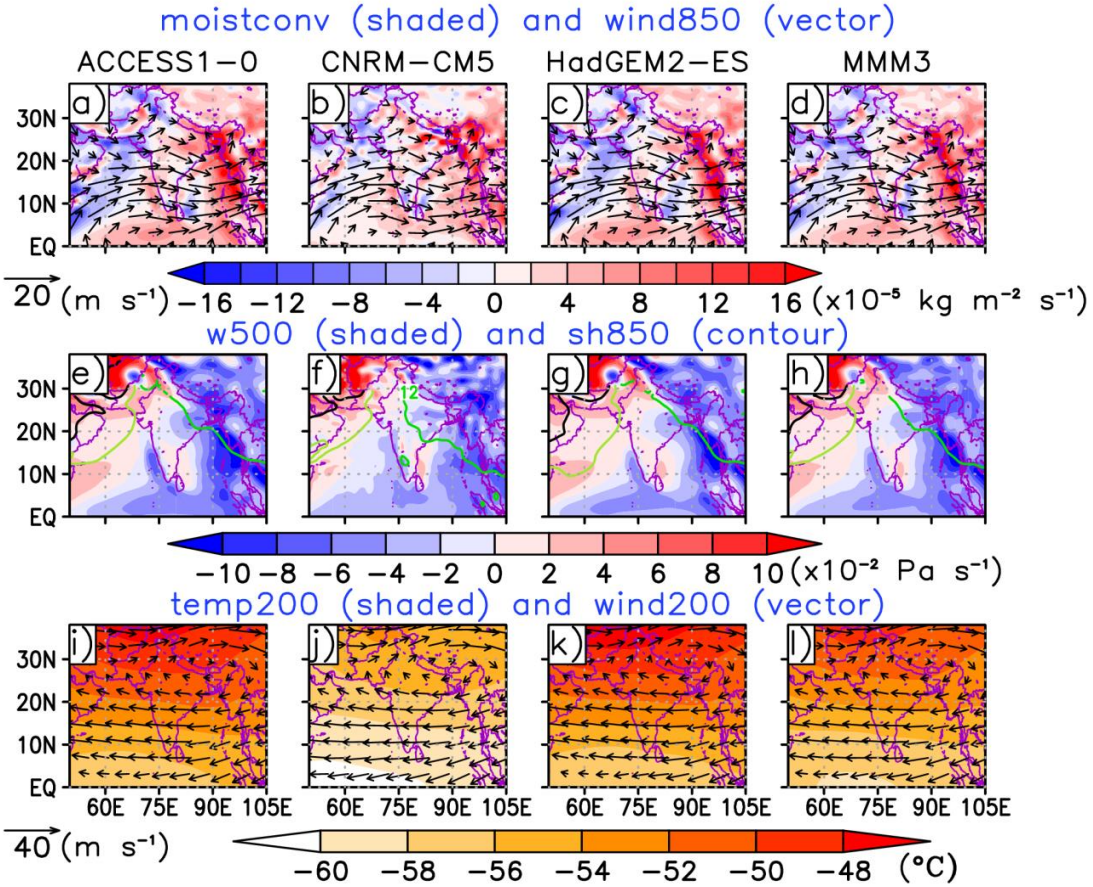


Figure 2.13. SMS mean state at present climate (1971–2000) a) – d) vertically integrated (surface to 300 hPa) moisture flux convergence (shaded) and wind at 850 hPa (vector); e) – h) omega at 500 hPa (shaded) and specific humidity at 850 hPa in g kg^{-1} (contour, 4 (black), 8 (light green) and 12 (green)); i) – l) temperature (shaded) and wind at 200 hPa (vector).

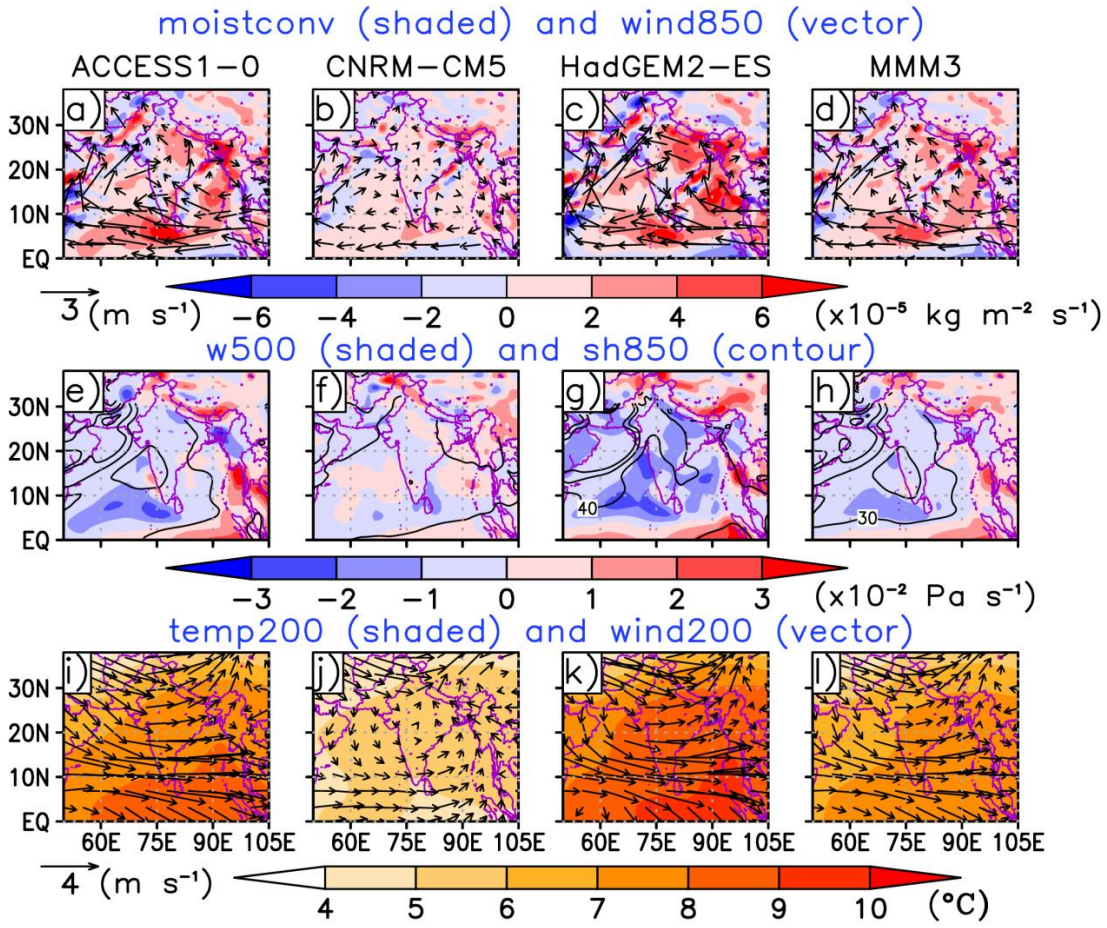


Figure 2.14. Projected changes in the SMS mean state during 2070–2099 under RCP8.5 a) – d) vertically integrated (surface to 300 hPa) moisture flux convergence (shaded) and wind at 850 hPa (vector); e) – h) omega at 500 hPa (shaded) and specific humidity at 850 hPa (contour, unit: %); i) – l) temperature (shaded) and wind at 200 hPa (vector). Projected changes are computed as ($Mean_{Future}$ - $Mean_{Present}$) except for the change in specific humidity. Changes in specific humidity are computed as (100 × ($Mean_{Future}$ - $Mean_{Present}$) / $Mean_{Present}$).

Most of the models (66%) show a wet bias over steep topography *(Mehran et al. 2014)*. The rest of the models show dry bias over CHR (Table 2.3). Large bias in the model is associated with the convective parameterization schemes *(Sabeerali et al. 2015)*. However, it is extremely difficult to find a single conclusion showing what makes a model good/bad in complex terrain (Figs. 2.4a–2.4c, Table 2.3) *(Su et al. 2013, Palazzi et al. 2015)*. Past evidence shows that high-resolution models perform better than coarse resolution models by representing topography and circulation *(e.g. Fennessy 1994; Rajendran et al. 2013)*. Our analysis shows that slight improvements in model resolution might not reduce model errors significantly. For example, resolutions of CCSM4, CESM1-BGC, CESM1-CAM5, and CMCC-CM are finer than those of HadGEM2 models, but their capability does not outperform that of HadGEM2 models (Table 2.2, Figs. 2.4a–2.4c). Nevertheless, it is noteworthy that the spatial pattern of SMS precipitation is poorly reproduced by coarse resolution (≥ 2.5° longitude) models. Similarly, it is apparent from three HadGEM2- (AO, CC, and ES) models that incorporation of complex physical processes does not improve the model capability to a marked degree. HadGEM2 models differ mutually in terms of the inclusion of complex physical processes *(The HadGEM2 Development Team*

2011). The improvement in bias is much less than the systematic mean bias. These results highlight that models are expected to have high resolution (a few kilometers) with improved physical parameterization schemes of physical processes to obtain a reasonable spatial distribution *(Rajendran et al. 2013)*.

Among the 38 GCMs, only 6 GCMs, ACCESS1.0, CNRM-CM5, EC-ERATH, HadGEM2-AO, HadGEM2-CC, and HadGEM2-ES fulfill our criteria of best models. This result suggests that further understanding of governing physical processes and model improvements must be undertaken to simulate precipitation over mountainous regions realistically. HadGEM2-AO and HadGEM2-CC are omitted to study the future projection of precipitation because their capability of reproducing precipitation characteristics is lower or nearly equal to HadGEM2-ES for the present climate. Our results show agreement with past studies *(Lee and Wang 2014; Prasanna 2016)*. The multi-model means of selected best models reproduce observed precipitation characteristics (Figs. 2.5–2.7) reasonably well. However, considerable biases do exist related to the observed precipitation and related to inter-model differences. It is particularly interesting that 3 of 4 of the best selected models: ACCESS1.0, CNRM-CM5, and HadGEM2-ES are recommended as the best models for South Asia *(Prasanna, 2016)* as well as for global *(Lee and Wang 2014)* monsoon assessment.

Best models and multi-model means project an increase in SMS mean precipitation in all future periods under RCP4.5 and RCP8.5. Model agreement and statistical significance are higher for mid-future and far-future periods. Unlike mean precipitation, model agreements in future projection of variability change are not consistent for all future periods. However, change in IAV in far-future periods (2070–2099) under RCP8.5 is consistent and is greater than the mean change. The multi-model mean of MMM4 projects an approximately 20% increase in the mean SMS precipitation. The increase in mean precipitation might be attributed to an increase in the moisture flux convergence attributable to the increase in the water holding capacity of the atmosphere in a warming environment *(Turner and Annamalai 2012; Sharmila et al. 2015; Prasanna 2016)*, enhancement of climatological low-level monsoon flow over the Arabian sea *(Ogata et al. 2014)* and convective activities but it is slightly offset by weakened easterly winds over the Arabian sea and central India at the upper troposphere because of warming of the Indian Ocean (Figs. 2.13 and 2.14). Actually, IAV of CHR precipitation is largely dominated by the Pacific Ocean sea surface temperature (ENSO). The ENSO–monsoon relation is projected to weaken over time *(Li and Ting 2015)*. However, simulation of ENSO by CMIP5 models is extremely uncertain *(Turner and Annamalai 2012; Jha et al. 2014)*.

Moderate to heavy precipitation days are projected to increase in the far-future period under RCP8.5, whereas light precipitation days are projected to decrease. However, regional uncertainty does exist. The projected increase of moderate to heavy rainfall might be related to increased flow from the Arabian sea, increased moisture attributable to warming and increased convective activity over the study region (Figs. 2.13 and 2.14) *(Houze et al. 2007)*. The decrease in light rainfall might be attributable to a shift of the daily amount to higher rainfall categories. Furthermore, the frequency and length of active monsoon spells is projected to increase consistently in future periods, which might be attributed to enhanced low level circulation, increased convective activities, and moisture attributable to warming (Fig. 2.14). Our results related to extreme precipitation show good general agreement with those of earlier studies *(Kitoh and Endo 2016; Sharmila et al. 2015)*.

Results show that increased mean and variability are largely the result of increased frequency of moderate to heavy precipitation days and active spells. Our analyses show that more frequent and prolonged intense precipitation might occur in the future, which might strongly and adversely affect society if suitable adaptation measures are not planned.

Thorough analyses are necessary, however, to investigate the physical processes responsible for spatial and temporal variability and to evaluate how those processes are reproduced in the climate simulation. As described herein, we did not perform analyses to ascertain the reasons behind the poor capability of models to simulate the spatial and temporal distribution of CHR SMS precipitation. In addition, possible causes of projected changes in precipitation characteristics might be related to numerous simulated variables. These matters are left as subjects for future work.

Chapter 3

Predictability of precipitation on short to medium range timescale in TIGGE ensemble prediction systems

3.1. Introduction

Precipitation-induced natural disasters such as floods, landslides and droughts are the major devastating and life threating disasters in the central Himalaya, Nepal. Accurate precipitation forecasts are of vital importance in preventing and mitigating natural disasters *(Fritsch et al. 1998)*. In addition, accurate precipitation forecast is of great value to devise proactive plans in various sectors such as agriculture, water resource management, hydropower operation and in daily life as well. In recent decades, numerical weather prediction (NWP) models have become the backbone of the weather forecasting worldwide as these models have acquired greater skills and are playing an increasingly important role in the weather forecasting *(Baur et al. 2015)*.

Despite the enormous improvement of NWP model, forecasts issued a few days ahead can frequently be quite wrong and even short-range forecasts, for instance, first day forecasts can sometimes have large errors. These errors are due to the chaotic dynamics of the atmosphere, which means that very small errors in the initial conditions can lead to an inevitable large growth of the forecast error, the so-called butterfly effect; and also due to the model imperfection itself. It is very difficult to have a perfect forecast system because we can never observe every detail of the initial state of the atmosphere and devise the model which perfectly represents the complex atmospheric processes. *Lorenz (1963, 1965)* investigated that the atmosphere has a finite limit of predictability even if the model is perfect and the initial conditions are known almost perfectly.

To cope with the uncertainty associated with the imperfect initial conditions and numerical model, leading operational forecast centers across the world are currently providing medium range ensemble weather forecasts through their ensemble predictions systems (EPSs) *(Tracton and Kalnay 1993; Molteni et al. 1996; Buizza and Palmer 1998; Kalnay et al. 1998)*. In an EPS, many forecasts are produced with either slightly varying initial conditions or slightly varying model or both to account the uncertainty in initial conditions and model formulations *(Charron et al. 2010)*. Ensemble forecasts are used to generate ensemble mean and probabilistic forecast. Probability forecasts derived from the EPS are of greater benefit than the deterministic forecasts produced by the same model *(Richardson 2000)*. Unlike a deterministic forecast, an EPS approach provides both the forecast of the most likely event and the uncertainty of the forecast.

As a major part of The Observing System Research and Predictability Experiment (THORPEX), which aims to accelerate the improvements in the accuracy of 1-day to 2-week high-impact weather forecasts, The International Grand Global Ensemble (TIGGE) integrates ensemble weather forecasts from 10 major centers in the world and facilitates research opportunities to systematically compare and validate the EPSs and multi-model grand ensemble *(Bougeault et al. 2010; Swinbank et al. 2016)*.

Many studies have been carried to evaluate TIGGE performances at global, hemispheric and regional levels. *Kipling et al. (2011)* analyzed the performances of forecasts from nine operational medium-range ensemble predictions systems (EPSs) available at TIGGE to investigate spatiotemporal dynamics of their perturbations using mean–variance of logarithms (MVL) diagram, and found that there is a divide between

ensembles based on singular vectors (SVs) or empirical orthogonal functions (EOF), and those based on bred vector (BV), ensemble transform with rescaling (ETR), or ensemble Kalman filter (EnKF) techniques. They further discussed how the MVL technique assists in selecting models for inclusion in a multi-model ensemble. *Matsueda and Endo (2011)* verified the performances of nine TIGGE EPSs to predict the Madden–Julian oscillation (MJO) using three years (2008–2010) forecasts, and investigated that European Centre for Medium-range Weather Forecasts (ECMWF) and United Kingdom Meteorological Office (UKMO) generally show the best performances in predicting the MJO with varying performance with the phase of MJO. *Hamill (2012)* verified the Probabilistic quantitative precipitation forecasts (PQPFs) from four TIGGE EPSs and ECMWF reforecast-calibrated PQPFs over the contiguous United States during July to October 2010, and showed that PQPFs from the Canadian Meteorological Centre (CMC) are most reliable but least sharp. Additionally, he investigated that the U.S. National Centers for Environmental Prediction (NCEP) and UKMO are least reliable but sharper. *Su et al. (2014)* evaluated ensemble mean quantitative precipitation forecasts (QPFs) and PQPFs from six TIGGE EPSs for tropics and mid-latitude of northern hemisphere (NH) during June to August 2008–2012, and found that ECMWF is generally the best EPS. They further revealed that the overall forecast skill is better in the NH mid-latitude than in the NH tropics, and CMC is relatively good for short-range QPFs and PQPFs at light precipitation thresholds.

Although EPSs have acquired greater skill, they still possess large forecast errors. In order to provide improved forecasting skills in an inexpensive way, different multi-model ensemble techniques have been practiced *(Krishnamurti et al. 1999, 2009; Johnson and Swinbank 2009; Roy Bhowmik and Durai 2010; Kumar et al. 2012; Durai and Bhardwaj 2014; Matsueda and Nakazawa 2015)*. In multi-model ensemble forecasting, the main issue relates to the removal of the collective errors of multi-models. Many studies have utilized simple multi-model ensemble of all available EPSs giving an equal weight to every model. The major drawback of assigning an equal is that it may degrade the overall skills by equally weighting the poor models. To address this problem, multi-model ensemble forecasting by assigning optimal weights based on recent past model performance is introduced. *Krishnamurti et al. (2009)* applied a multi-model super ensemble technique to forecast precipitation over china using five TIGGE EPSs, and showed a major improvement in the prediction skills compared to simple ensembles and individual EPSs. *Johnson and Swinbank (2009)* verified the multi-model ensemble means of ECMWF, UKMO and NCEP created by bias-correction, model-dependent weights and variance adjustment. They investigated that the multi-model ensemble gives an improvement over single-model ensemble, and the improvement is also dependent on variable. Furthermore, they explored that the improvements are small if the ensembles are similar to each other, and the improvement of further post-processing is relatively small. Several studies performed for India or Indian subcontinent monsoon precipitation forecasts revealed the benefit of multi-model ensemble techniques by assigning optimal weights to member models over the simple average of all member models and individual member models *(Roy Bhowmik and Durai 2010; Kumar et al. 2012; Durai and Bhardwaj 2014)*.

Models performance may be different for different weather situations *(Lorenz 1965)*. These differences may depend on the geographical location particularly if there is presence of land-sea boarder or mountains *(Gofa et al. 2013)* where physiography (surface features such as topography, land–water boundaries, vegetation, and soil moisture) controls the precipitation. Forecast verification is a key tool for guiding model development and helping users to understand the suitability of outputs for different applications. To date, no systematic

verification of ensemble mean QPFs and PQPFs from TIGGE models over the central Himalaya is performed to fully assess the relative strengths and weaknesses of the models. In this chapter, we provide verification of summer monsoon season (June–September, 2009–2012) precipitation forecasts in short to medium range time scale (day+1 to +7 days) based on daily forecasts against the observed analysis. Apart from comparing their behavior as independent ensembles, we have assessed the improvement of forecast using multi-model grand ensemble techniques. The findings of this study are expected to be useful to the model developer and user of model outputs particularly, the operational forecasters of the concerned regions, and the regional modelers who are seeking information about skill of the global models.

The remainder of this chapter is organized as follows. Section 3.2 describes the datasets and verification methods, and section 3.3 presents the skill of individual EPSs and multi–model grand ensemble. Finally, discussions and conclusions are presented in section 3.4.

3.2 Datasets and verification methods

3.2.1 Verification period and dataset

In this chapter, forecasts from TIGGE EPSs are verified only over Nepal (Fig. 3.1). For the verification, we have selected four monsoon seasons (June–September, 2009–2012) based on the area averaged mean monsoon season precipitation over Nepal and El Niño/La Niña conditions (Fig. 3.2). In terms of the Oceanic Niño Index (ONI), 2009 (dry year), 2010 (normal year), 2011 (excess year) and 2012 (dry year) are respectively El Niño, La Niña, La Niña and normal years.

ONI is the sea surface temperature (SST) anomalies in the Niño 3.4 region (5°N-5°S, 120°-170°W) (http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml). Generally, Nepal receives below normal rainfall in El Niño year and excess rainfall in La Niña year *(Shrestha 2000; Shrestha et al. 2000)*.

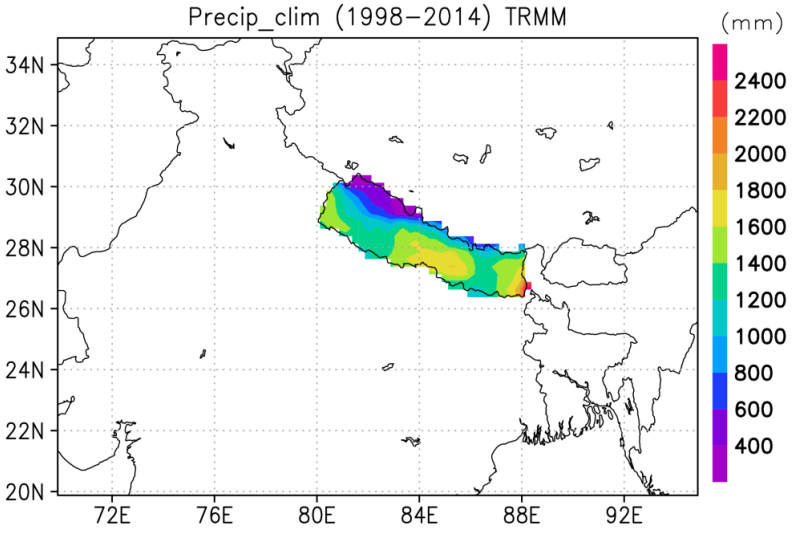


Figure 3.1. Map of study area (shaded area: Nepal). Shading indicates monsoon season precipitation climatology (1998-2014). Data source: TRMM3B42v7 (Huffman et al. 2007).

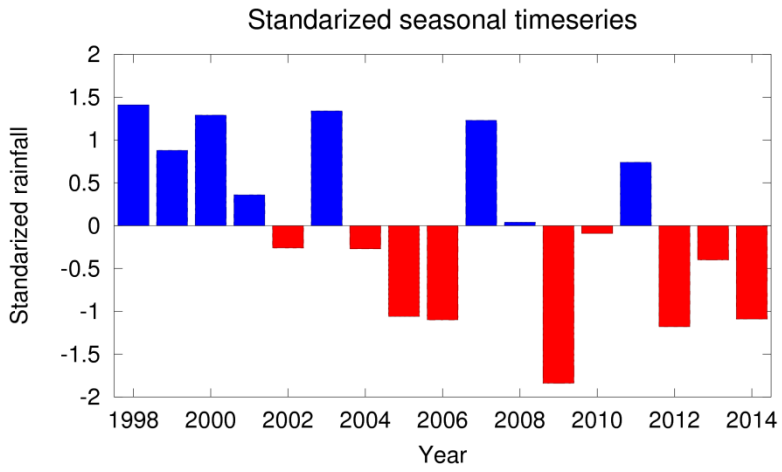


Figure 3.2. Monsoon season precipitation anomaly (standardized) of area averaged precipitation with reference to climatology (1998-2014). Data source: TRMM3B42v7 (Huffman et al. 2007).

APHRODITE's daily gridded precipitation *(Yatagai et al. 2012)* at 0.05° spatial resolution has been used to validate the ensemble precipitation forecasts. APHRODITE dataset is publicly available at 0.25° and 0.5° spatial resolutions only for the period of 1951-2007. APHRODITE team is going to release the updated version to public soon (http://st.hirosaki-u.ac.jp/~aphrodite2/english/downloads.html), and nationwide high resolution (0.05°) dataset has already been distributed only to the nations who had shared stations data with the project. High resolution dataset (0.05°) used in our study is acquired from the Department of Hydrology and Meteorology, Nepal.

Considering the spatial resolutions of the global operational models, APHRODITE's precipitation dataset at fine spatial resolution (0.05°) is interpolated to 0.25° using simple box averaging method *(He et.al., 2016)*, and a total number of 293 grids of 0.25° lies over Nepal. It is worthy to note that APHRODITE daily precipitation is the 24–hr accumulated precipitation from 03 Coordinated Universal Time (UTC) of previous day to 03 UTC of the day consistent with the synoptic observation time in Nepal *(Yatagai et al. 2012)*. Hence, in this study, we have used APHRODITE daily precipitation dataset ending at 03 UTC to validate daily accumulated precipitation forecast ending at 00 UTC. Since the contribution of late morning time (Nepal time is 5:45 hours ahead of UTC) precipitation to daily accumulated precipitation is less *(Shrestha and Deshar 2015)*, the error caused by three hours timing mismatch is usually negligible compared to model forecast errors *(Durai and Bhardwaj 2014)*.

3.2.2 Model forecast datasets

Precipitation forecasts from five operational models (Table 3.1, Appendix B.1): China Meteorological Administration (CMA) *(Wang et al. 2008)*, Canadian Meteorological Center (CMC) *(Charron et al. 2010)*, European Centre for Medium–Range Weather Forecasts (ECMWF) *(Buizza et al. 2007)*, National Centers for Environmental Prediction (NCEP) *(Cui et al. 2012)* and United Kingdom Meteorological Office (UKMO) *(Bowler et al. 2008)*, initialized at 00 UTC used in this study were acquired from the TIGGE archive hosted

online at the ECMWF TIGGE data portal (http://apps.ecmwf.int/datasets/) for June–September, 2009–2012. Daily accumulated precipitation forecasts for one to seven days lead time (day+1 to day+7) from both control run and perturbed members are extracted, and are interpolated to 0.25°×0.25° spatial resolution of APHRODITE dataset using bilinear technique. Consequently, verification time is different for different lead times. For instance, day+1 forecast verification period covers 2June–1October, 2009–2012 (122 × 4 = 488 days) whereas 7+day verification period covers 8June–7October, 2009–2012. For QPF verification, ensemble mean of all available forecasts including control forecast and only perturbed forecasts of each centre are computed giving equal weight to each member (Appendix B.2.1). Additionally, for PQPF verification, probabilities of precipitation for three different thresholds (1, 10, 25 mm day^{-1}) are computed for all lead times (day+1 to +7) using all available members including control forecast (Appendix B.3.2).

Table 3.1: Brief descriptions of the TIGGE EPSs employed in this study (as of September, 2012).

NWP centre	Forecast model resolution	Ensemble size	Base time (UTC)	Forecast length (days)	Perturbation method
CMA	TL213 L31	14 + 1	00/12	10	BV
CMC	600 × 300 L40	20 + 1	00/12	16	EnKF
ECMWF	TL639 L62 (0–10day) TL319 L62	50 + 1	00/12	15	SV- EDA
NCEP	T254 L42 (0–8day) T190 L62	20 + 1	00/06/12/18	16	BV-ETR
UKMO	N216 L85	23 + 1	00/12	15	ETKF

BV, Bred Vector; EnKF, Ensemble Kalman Filter; SV, Singular Vector; EDA, Ensemble of Data Assimilations; ETR, Ensemble Transform with Rescaling; ETKF, Ensemble Transform Kalman Filter

As EPSs are upgraded (e.g., higher resolution, improved initialization procedures, and improved parameterizations) during the course of TIGGE analysis (https://software.ecmwf.int/wiki/display/TIGGE), this study does not pay any attention to the change of performance after model upgrade because not all the EPSs provide reforecast datasets using current configuration of the model.

3.2.3 Multi-model grand ensemble mean

To investigate the relative improvement in forecast skills of multi-model grand ensemble means, we have used four types of techniques to create multi-model ensemble mean for June–September, 2012 (Appendix B.2.2). Two sets of multi-model grand ensemble means are created by simply averaging the ensemble mean forecasts from all five models with and without bias corrections, and the remaining two sets are created based on the model's performance (root mean square error: RMSE and correlation coefficient: CC) for the past three

consecutive seasons (JJAS, 2009–2012). To create ensemble mean of bias removed EPSs, best easy systematic bias (BES) *(Durai and Bhardwaj 2014)* is computed based on the performances for the past three consecutive seasons. The detail procedure to create multi-model grand ensemble means is provided in Appendix B.2.2.

3.2.4 Verification methods

A deterministic forecast gives the definite information on the occurrence (magnitude) of an event while a probabilistic forecast gives the information on the uncertainty of the prediction. In this study, multiple verification metrics which are widely used in forecast verification *(Wilks 2006; WMO 2008)* have been used to verify both deterministic and probabilistic precipitation forecasts.

3.2.4.1 Verification of quantitative precipitation forecasts (QPFs)

Deterministic forecasts can be verified as categorical events or continuous variables, with different verification scores appropriate for each view. Root mean square error (RMSE), correlation coefficient (CC), bias and ensemble spread are computed to verify continuous variable: precipitation amount (Appendix B.3.1). Only forecast error statistics (RMSE, CC, bias and spread) are not sufficient scores to investigate the skill of the models in case of precipitation forecast, which values more the statistical scores based on the frequency of occurrences in various classes. Therefore, it is important to examine the skill of precipitation forecast in terms of precipitation amounts in different categories and thresholds. Standard skill scores of dichotomous events (Frequency bias: bias; False alarm ratio: FAR; Probability of detection: POD; and Equitable threat score: ETS) are investigated using 2×2 contingency table (Appendix B.3.1). Additionally, we have used discrimination diagrams to investigate how well the EPSs can discriminate observed precipitation categories. In a discrimination diagram, ordinate shows the forecast relative frequencies of observed precipitation categories against the rainfall categories of forecast as abscissa. In this study, observed precipitation is divided into three categories (light: 1-10, moderate: 10-25, and heavy: 25-50 mm day^{-1}) whereas forecast precipitation is divided into five categories (no rain: <1, light: 1-10, moderate: 10-25, heavy: 25-50, and torrential: >50 mm day^{-1}).

3.2.4.2 Verification of probabilistic quantitative precipitation forecasts (PQPFs)

To evaluate the PQPF, reliability diagram and the Brier Skill Score (BSS) are used (Appendix B.3.1). The reliability diagram, a plot of the observed frequencies against the forecast probabilities where the range of forecast probabilities is divided into *K* bins, demonstrates how well do the predicted probabilities of an event correspond to their observed frequencies. In this study, forecast probabilities are divided into 10 bins with the interval of 0.1. When the forecast probabilities agree well with the frequencies of the events for a particular probability, the distribution curve should lie along a diagonal line. The deviation from the diagonal gives the conditional bias. The curve below (above) the line indicates that forecast system is overforecasting (underforecasting). The flatter the curve in the reliability diagram indicates less resolution. The frequency of forecasts in each probability bin (shown in the histogram) shows the sharpness of the forecast.

BSS, a widely used score for categorical probabilistic forecasts which uses Brier Score (BS) as the underlying accuracy measure, quantifies the improvement of the probabilistic forecast relative to a reference forecast. In this study, reference BS is computed using the sample climatology frequency on each grid *(Su et al. 2014)*. The values of BSS ranges between 0 and 1; 0 indicates a forecast comparable to the climatological forecast and 1 indicates a perfect forecast. BSS is not suitable for small samples. The BS can be decomposed into reliability (REL), resolution (RES) and uncertainty *(Murphy, 1973)* (Appendix B.3.2). The reliability measures the closeness of the forecast probability with the observed frequency whereas the resolution measures the difference of the forecast probability with the climatological probability of the event.

3.3 Results

3.3.1 Verification of ensemble mean QPFs

3.3.1.1 Precipitation climatology and forecast errors

Seasonal mean climatology of ensemble mean QPFs (JJAS, 2009–2012) from five TIGGE EPSs and APHRODITE are compared for all lead times. UKMO, ECMWF and CMA reproduce major observed spatial pattern with high spatial correlation coefficient consistently in all lead days with slightly decreasing wet bias (Fig. 3.3 and Figs. B.4.1.A1-7 in Appendix B). However, CMA shows very large wet bias throughout the central Nepal while UKMO shows wet bias in the southern parts. Reproducibility of spatial pattern in CMC is poor with dry bias over the western parts. NCEP shows even negative spatial correlation (Fig. 3.3 and Appendix B.4.1) which could be related to the northeastward shift of precipitation belt (Fig. 3.4). However, area averaged daily precipitation forecast shows skill close to the other models. For instance, daily forecasts of all EPSs for day+3 for JJAS, 2010 is comparable with other models as shown in Fig. 3.5.

The skills of control and ensemble mean QPFs are compared through the seasonal mean of daily RMSE, spatial correlation coefficient (CC) and bias (Fig. 3.6) during JJAS, 2009–2012. CMC shows lowest RMSE for control QPFs for all lead times, which is also lower than the RMSE of the mean forecasts of perturbed members of all TIGGE EPSs except ECMWF, and for ensemble mean QPFs for the first three lead days. CMA shows the largest RMSE for control QPFs for all lead times, and for ensemble mean QPFs for the first three days (Fig. 3.6a). The highest RMSE of CMA is due to the excessive wet bias while small RMSE of CMC is contributed from small dry bias (Fig. 3.6b). NCEP shows the lowest CC followed by CMC (Fig. 3.6c). Some of the TIGGE EPSs show different skills for control and ensemble mean QPFs. For instance, ECMWF is the second best model for the ensemble means QPFs and the second worst model for the control QPFs for the first three lead days in terms of RMSE. However, ECMWF is the best in terms of CC and bias. Based on the all metrics (RMSE, CC and bias), UKMO is one of the best performing EPSs for both control and ensemble mean QPFs. It is noticed that the improvements in skill of QPFs using ensemble configuration in the ECMWF and CMA EPSs are relatively higher than in other EPSs which could be related to the number of ensemble members and perturbations techniques used to initialize model and physical parameterization schemes.

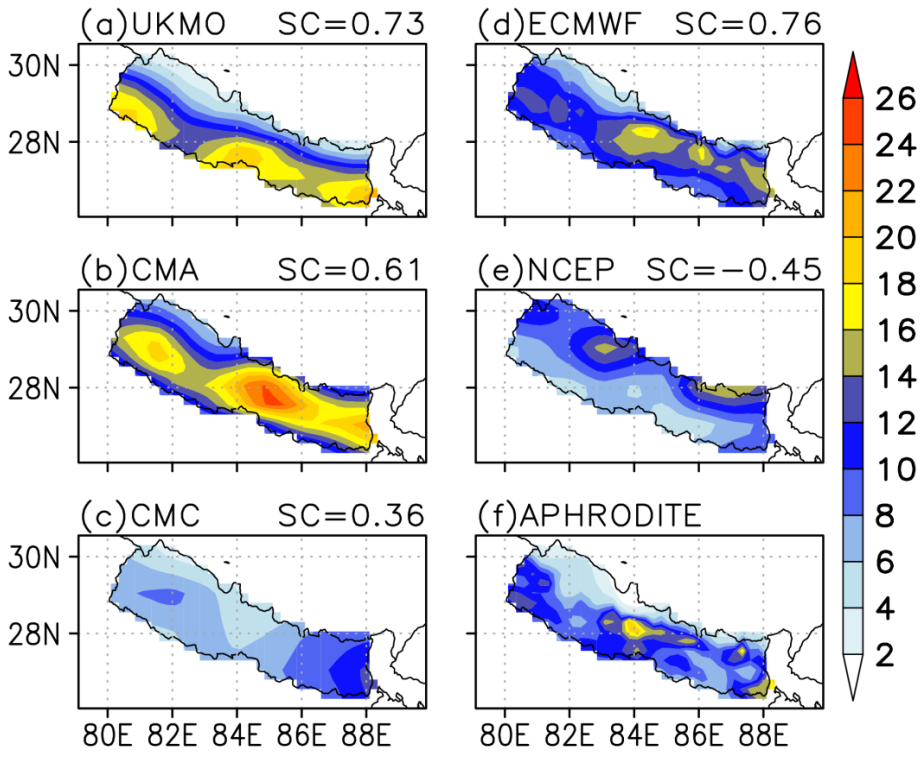


Figure 3.3. Average precipitation (mm day^{-1}) of ensemble mean forecasts for day+3 from five EPSs and APHRODITE during JJAS, 2009–2012. The pattern correlation coefficient (SC) is shown in the title.

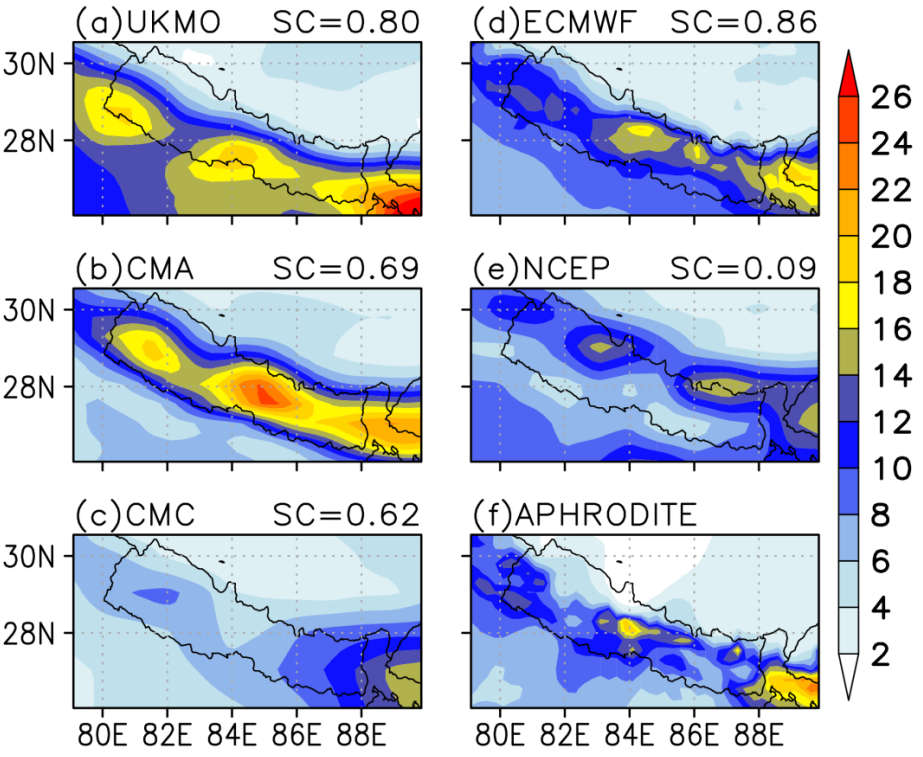


Figure 3.4. Same as figure 3.3 but showing the precipitation over Nepal and surrounding areas.

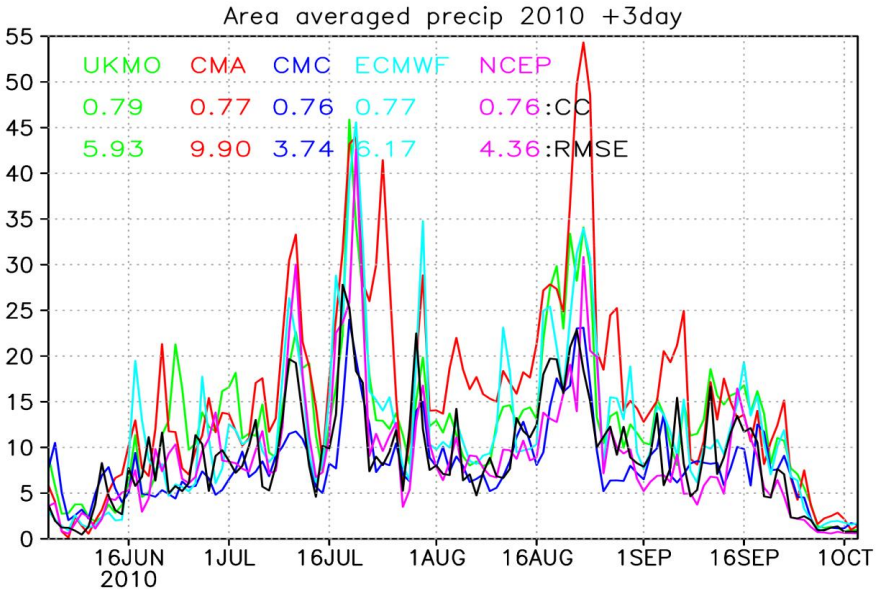


Figure 3.5. Timeseries of area averaged (Nepal) daily precipitation (mm day^{-1}) forecasts from five EPSs (colored lines) for +3day and APHRODITE (black line) during JJAS, 2010. Temporal correlation (CC) and root mean square error (RMSE) are shown inside the figure with numbers.

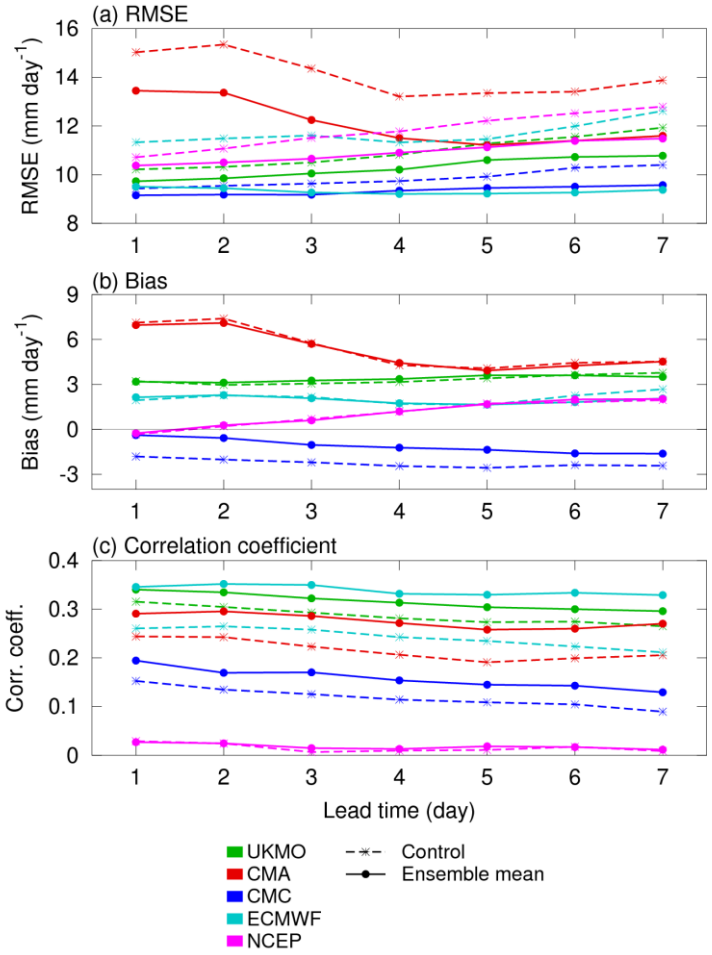


Figure 3.6. The (a) RMSE (b) Bias and (c) spatial correlation coefficient of the control forecasts (dotted) and ensemble mean forecasts excluding control forecasts (solid) during JJAS, 2009–2012.

3.3.1.2 QPFs of categorical and dichotomous events

In this section, we present the skill of EPSs to reproduce different observed precipitation categories and different precipitation thresholds. Unlike the previous investigation made for the northern hemisphere that the discrimination abilities of the models to discriminate observed precipitation categories decreases with increasing lead time *(Su et al. 2014)*, every TIGGE models shows almost consistent discrimination abilities for all lead times (day+1 to +7) (Fig. 3.7) in our case. UKMO, CMA and CMC can discriminate observed light and moderate precipitation events to some degree whereas NCEP and CMC showed poor discrimination ability (Fig. 3.7). However, UKMO, CMA and ECMWF overforecasted observed light precipitation events as moderate events with the highest value by CMA. These models also overforcasted observed heavy precipitation events as moderate events. CMA can also discriminate heavy precipitation until day+3 which could be related to very wet bias (Fig. 3.6c). However, observed heavy precipitation events are overforecasted as moderate and torrential precipitation events. NCEP and CMC poor skill could be related to their poor representation of spatial pattern and dry bias (Fig. 3.6).

Standard skill scores for dichotomous events for different thresholds (1, 10 and 25 mm day^{-1}) and lead times (day +1, +3 and +5) computed from ensemble mean QPFs during JJAs, 2009–2012 are shown in Fig. 3.8. All the skill scores except ETS are consistent among the models and lead times for the light precipitation (>1 mm day^{-1}) (Fig. 3.8), and ETS of NCEP and ECMWF are relatively higher than others (Fig. 3.8b). Although the skill scores are largely scattered at moderate and heavy precipitation thresholds (10 and 25 mm day^{-1}), skill scores of ECMWF and UKMO were generally better than others. CMA showed large frequency bias at moderate and heavy thresholds (Fig 3.8a) which contributed to increase POD and FAR (Fig 3.8c, d), and aided to reduce ETS (Fig 3.8b). Consistent with the discrimination diagram (Fig 3.7), skill scores of CMC and NCEP are generally poorer than others (Fig 3.8). However, bias score of CMC at moderate threshold and NCEP at heavy threshold are relatively better (Fig 3.8b).

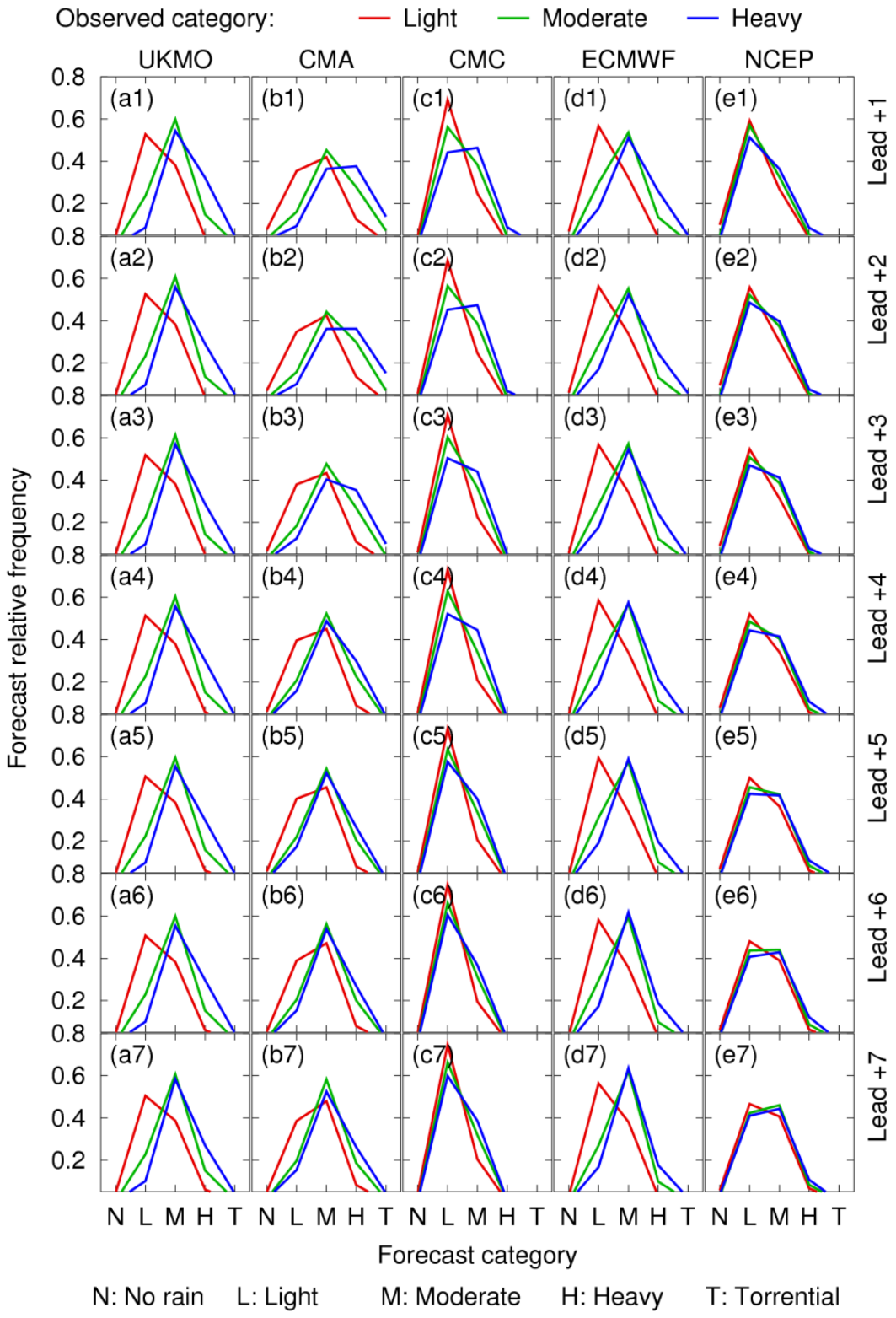


Figure 3.7. Discrimination diagram of the ensemble mean (including control experiment) QPFs in Nepal. The ordinate shows the forecast relative frequencies of observed light precipitation (1–10 mm day^{-1}, red), moderate precipitation (10–25 mm day^{-1}, green) and heavy precipitation (25–50 mm day^{-1}, blue) against five forecast categories: no precipitation (N, <1 mm day^{-1}), light precipitation (L, 1–10 mm day^{-1}), moderate precipitation (M, 10–25 mm day^{-1}) and heavy precipitation (H, 25–50 mm day^{-1}), and Torrential (T, >50 mm day^{-1}).

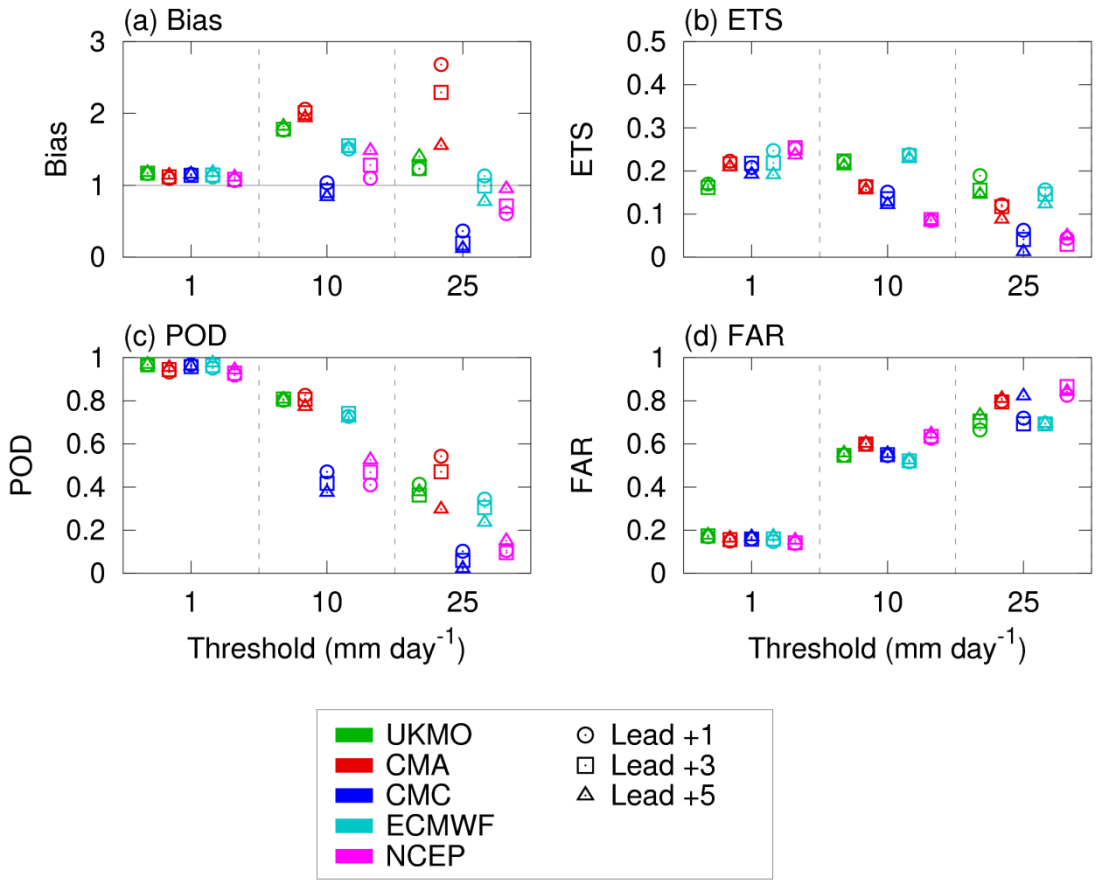


Figure 3.8. The (a) Bias, (b) ETS, (c) POD, and (d) FAR of the ensemble mean QPFs against different precipitation thresholds (1, 10, 25 mm day^{-1}) for different forecast lead times (day +1, +3, and +5) during JJAS, 2009–2012.

3.3.2 Verification of PQPFs

3.3.2.1 Spread-skill relationship

To investigate the spread-skill relationship of EPSs, we have analyzed ensemble spread and ensemble-mean error (Fig. 3.9). Theoretically, a well-constructed EPS should have the fast-growing ensemble spread consistent with the ensemble mean forecast error to capture the growth of possible forecast error *(Su et al. 2014)* since spread should reflect the true predictability of a flow. All the EPSs are underdispersive for all lead times (Fig. 3.9). NCEP and UKMO are the most severely underdispersive EPSs. However, forecast error of UKMO is slightly lower than NCEP after day+3. Spread-skill relationship of CMC (21 members) and ECMWF (51 members) are similar until day+5; afterwards RMSE and spread of ECMWF is larger than CMC. The CMC EPS uses multiphysics schemes to account model uncertainties *(Charron et al, 2010)* and has a large ensemble spread (Fig 3.9). Additionally, spread of CMA (15 members) is larger than NCEP (21 members). These results indicate that increasing ensemble sizes does not necessarily increase the ensemble spread and improve spread-skill relationship *(Su et al. 2014)*. Although the spread-skill relationship of CMC and ECMWF is similar, other statistics investigated in the above sections show that ECMWF is the best performing EPS. Therefore, it is suggested that these statistical scores should be used and interpreted with caution.

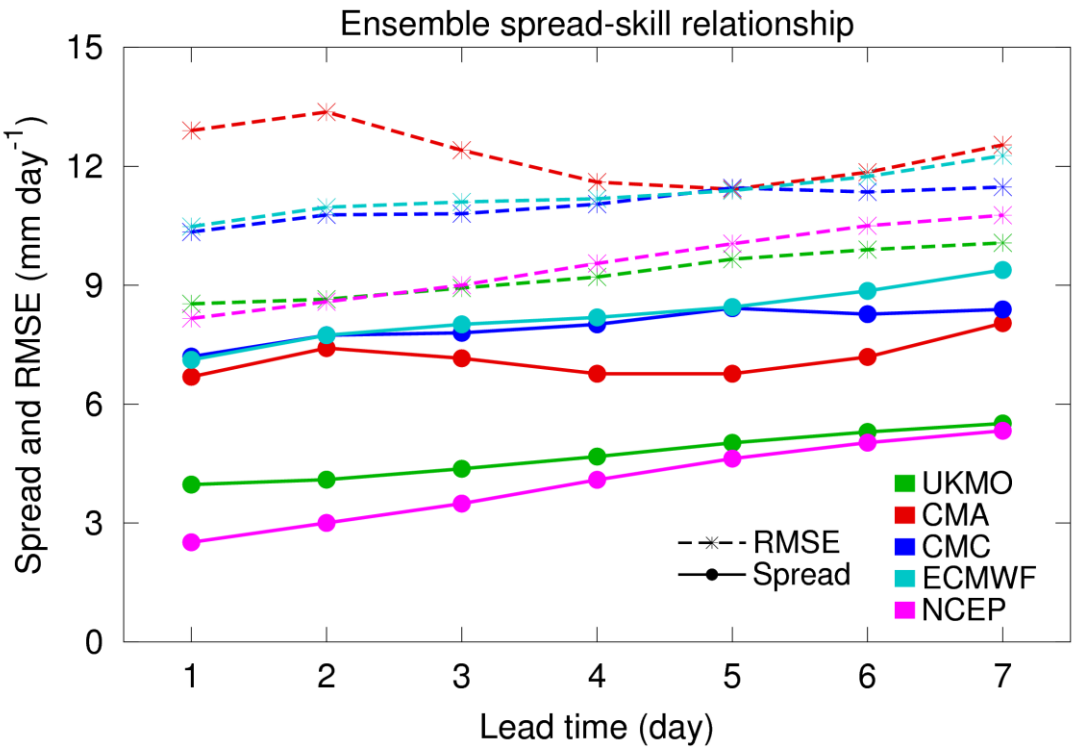


Figure 3.9. The RMSE of the ensemble mean (including control forecasts) QPFs (dotted) and ensemble spread (solid) during JJAS, 2009–2012.

3.3.2.2 PQPF of dichotomous events

The reliability diagrams of each TIGGE EPS for all lead times at three different thresholds (1, 10 and 25 mm day^{-1}) are analyzed (Fig. 3.10 and Appendix B.4.2.A1-7). Brier skill score (BSS), reliability (REL) and resolution (RES) are shown as the number at the top of each figure. Each EPS skills are different for different thresholds of precipitation but almost consistent with lead times (Fig. 3.10 and Appendix B.4.2.A1-7). All EPSs are overconfident in case of moderate (≥10 mm day^{-1}) and heavy (≥25 mm day^{-1}) precipitation events except CMC and NCEP, CMC and NCEP overestimate at lower probabilities and underestimate at higher probabilities (Fig. 10 c2 and e2) in case of moderate precipitation events. Unlike the moderate and heavy precipitation events, all EPSs show mixed skills in case of light precipitation events (≥1 mm day^{-1}), overestimate (underestimate) the high (low) risks (reliability curve is below (above) the diagonal line for higher probabilities) (Fig 10 a1, b1, c1, d1 and e1). BSS, REL and RES show that CMC and ECMWF are comparatively better than others, but are not sharp (frequencies at high probabilities categories are low or absent) for heavy events and the sharpness decreases as the lead time increases (Fig 3.10 c(1-3) and d(1-3) and Appendix B). UKMO is the third reliable models while NCEP is the poorest performing EPS followed by CMA in case of moderate and heavy events as indicated by reliability curve (flat curve parallel to climate), BSS (large negative), REL (large) and RES (low) (Fig. 3.10 a(1-3) and b(1-3)). The poor skill of CMA could be related to excessive wet bias while best skill of CMC could be related to dry bias (Fig. 3.5 and 3.6c).

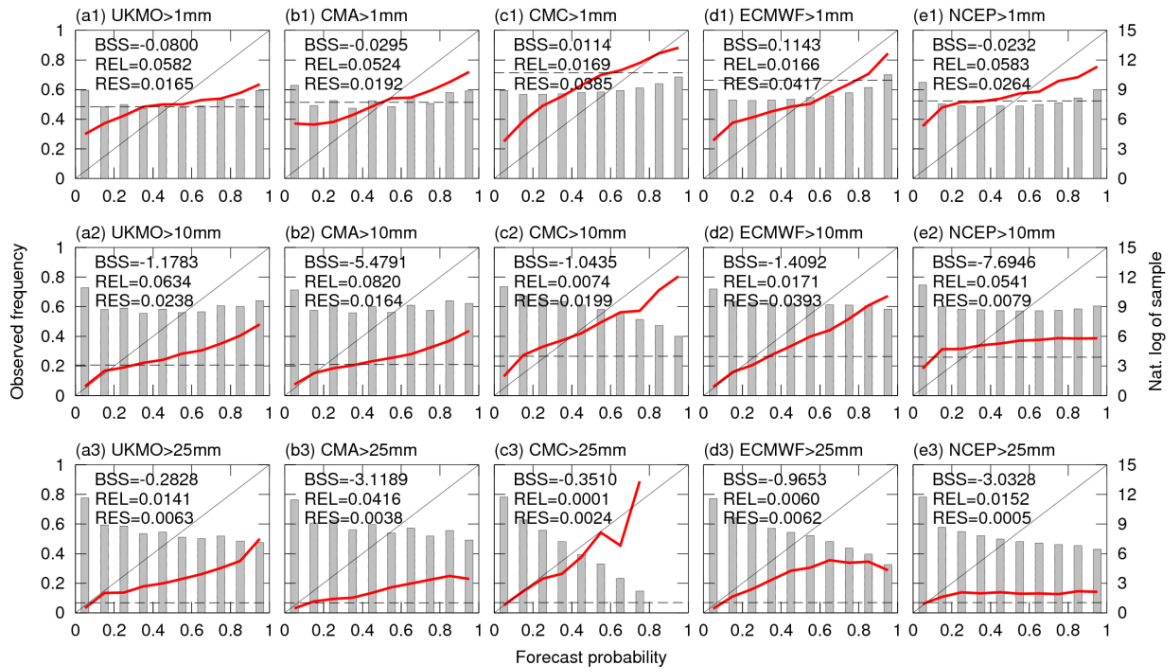


Figure 3.10. Reliability diagrams for day+3 at the different thresholds (1, 20 and 25 mm day^{-1}) (solid, red). The bar graphs show the sample frequencies in natural logarithmic scale (right side). The horizontal line (dotted, black) represents the observed sample frequency. The BSS, the reliability (REL) and resolution (RES) are shown as numbers. For brevity, probabilities of all EPSs with different member sizes are categorized into 10 bins.

3.3.3 Verification of multi-model grand ensemble means

Multi-models grand ensemble means are created with/out accounting recent past performances of the member EPSs, and comparison and verification of member EPSs and grand ensembles are performed for JJAS, 2012. The performances of EPSs for the past three consecutive seasons (JJAS, 2009-2011) have been used as training period whenever the past performances is accounted to create grand ensemble means (bias removed grand ensemble: MMM-BR, grand ensemble of weighted EPS with weight from RMSE: MMM-WR, and grand ensemble of weighted EPS with weight from CC: MMM-WC) (Appendix B.2.2).

3.3.3.1 Bias and weights during training period

The average best easy systematic bias (BES) of CMA is the highest followed by UKMO and ECMWF (Fig. 3.11) in order. Large bias of CMA in the first few days could be associated with model initialization. CMC has shown the lowest bias (Fig. 3.11) which is due to widespread dry bias over central and western parts during training period (Figs. B.4.3.1.A1-7 in Appendix B). NCEP shows the second lowest bias (Fig. 3.11) which is due to the dry bias in the southern parts (Figs. B.4.3.1.A1-7 in Appendix B).

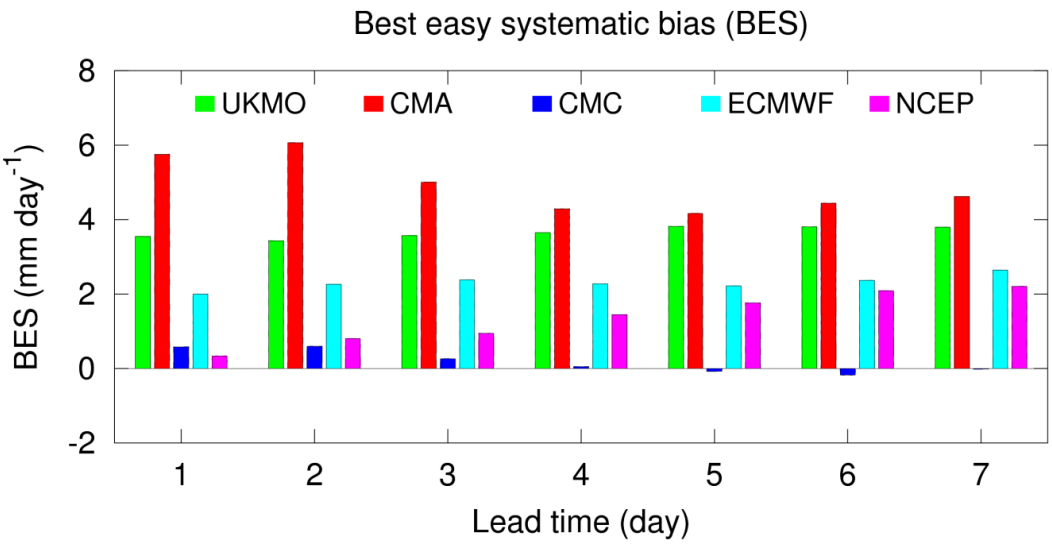


Figure 3.11. Area–averaged best easy systematic bias (BES) (mm/day) during training period (JJAS, 2009-2011).

Figure 3.12 show the area–averaged RMSE during training period for all lead time of each EPS and their respective weights. UKMO gets the highest weights until day+4 and same as CMC afterwards. CMA and NCEP, increases with lead time, get lesser weights as their respective RMSE are larger than other EPSs. ECMWF is ranked as third in terms of weight for all lead times. Broad scale features of spatial distribution of RMSE and weights show consistent pattern in all lead times (Figs. B.4.3.1.B1–7 and Figs. B.4.3.1.C1–7 in Appendix B). UKMO possesses higher weight in the northern parts whereas ECMWF and NCEP in the southern parts. CMC possesses higher weights in most parts of the country while CMA possesses lowest weight throughout the country.

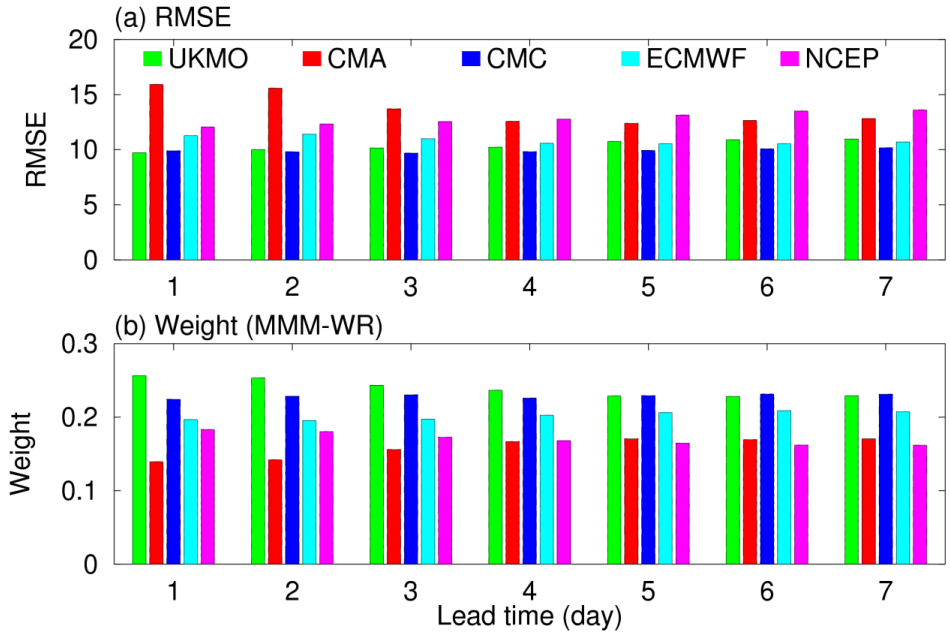


Figure 3.12. Area–averaged (a) RMSE (mm/day) during training period (JJAS, 2009-2011) and (b) weight.

The average temporal correlation coefficient (CC) of each EPS and weights based on their respective CC is shown in Fig. 3.13. In this case also, UKMO possesses the highest weight corresponding to its highest CC until day+5, and EMWWF outperforms others afterwards. Weight of CMC and CMA is almost same for all lead times whereas the NCEP is the lowest one. Unlike the Spatial patterns of weights from RMSE, spatial patterns of CC and their respective weights for all lead times are not consistent, and spatial differences do exit for different lead times (Figs. B.4.3.1.D1–7 and Figs. B.4.3.1.E1–7). Generally, UKMO and ECMWF possess the highest and NCEP possesses the lowest weights in most parts of the country.

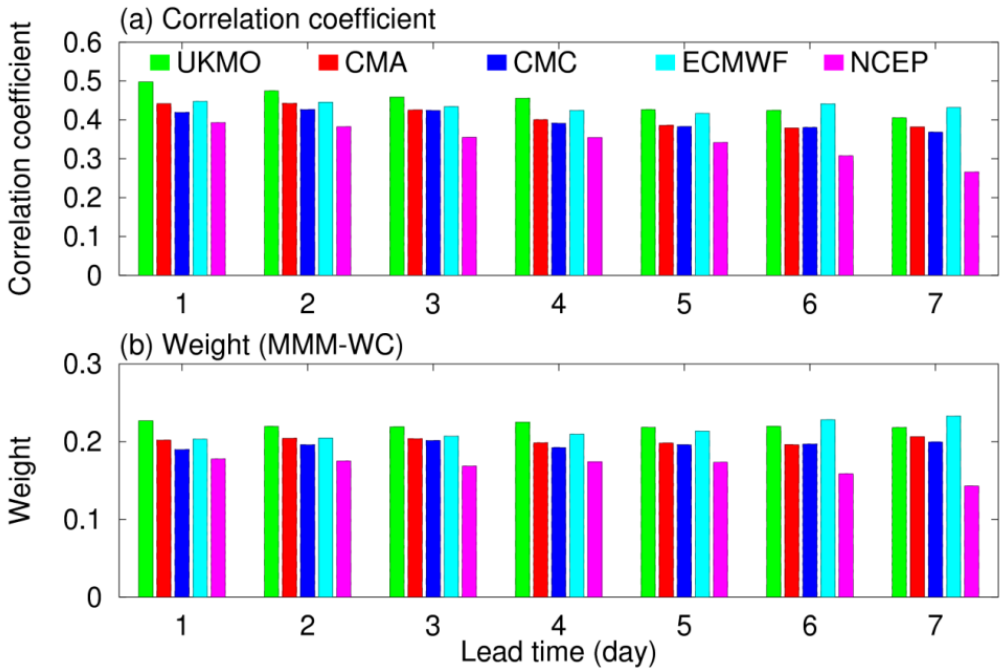


Figure 3.13. Same as Figure 3.12 but for correlation coefficient.

3.3.3.2 Verification of QPFs

Performances of all individual EPSs and multi-model grand ensemble means during JJAS, 2012 are shown in Fig. 3.14. The multi-model grand ensemble means outperform all individual models *(Matsueda and Nakazawa 2015)*, and MMM-BR outperforms all grand ensemble means (Fig 3.15). The performance of ECMWF is very close to simple multi-model grand ensemble (MMM) (Fig 3.14). As usual, CMA shows the largest mean error (Fig. 3.14a) and NCEP shows the lowest mean spatial correlation. Except CMA (CMC) whose RMSE (CC) decreases with lead time, models performance remained almost consistent for all lead times (Fig. 3.14a,b). The skill scores (RMSE and CC) of individual EPSs and multi-model grand ensemble means are compared with respect to MMM for three lead times (day+1, day+3 and day+5) (Fig. 3.15a,b). MMM-BR shows RMSE (CC) about 10% (20%) lower (higher) than MMM whereas MMM-WR shows RMSE (CC) about 5% (8%) lower (higher) than MMM. However, MMM-WC does not show any remarkable discrepancy with MMM.

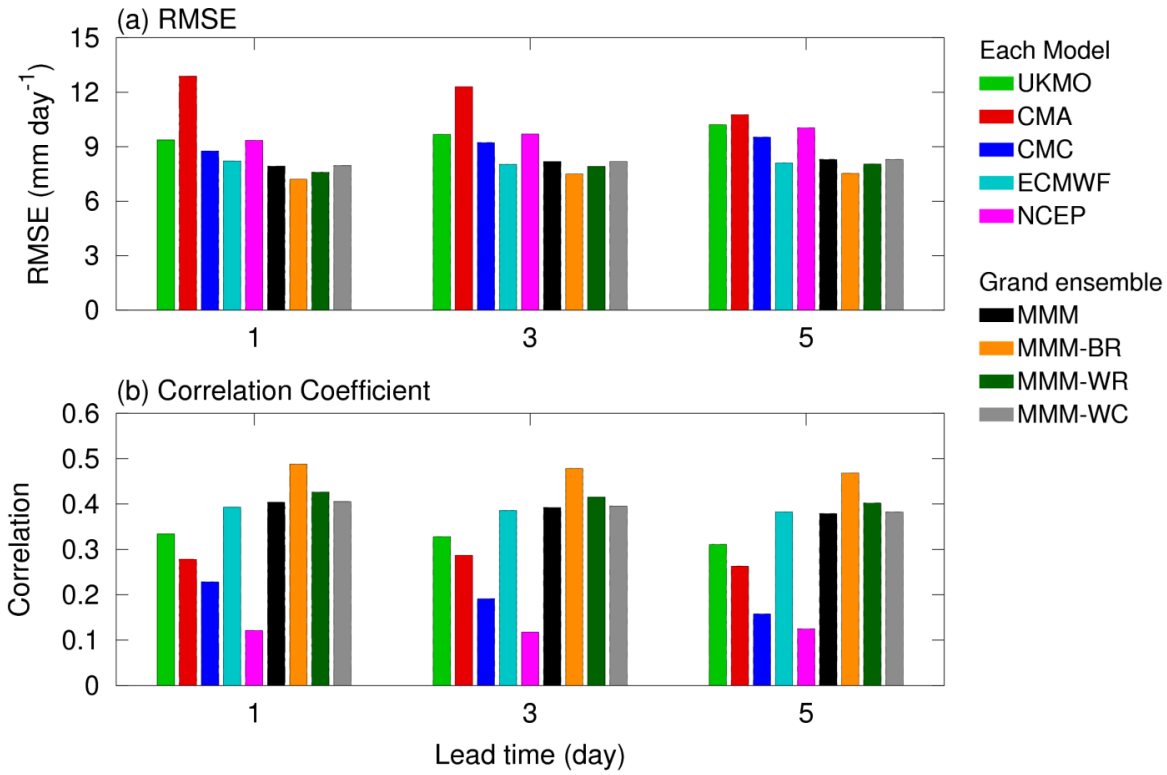


Figure 3.14. Average of daily (a) RMSE (b) correlation coefficient during monsoon 2012.

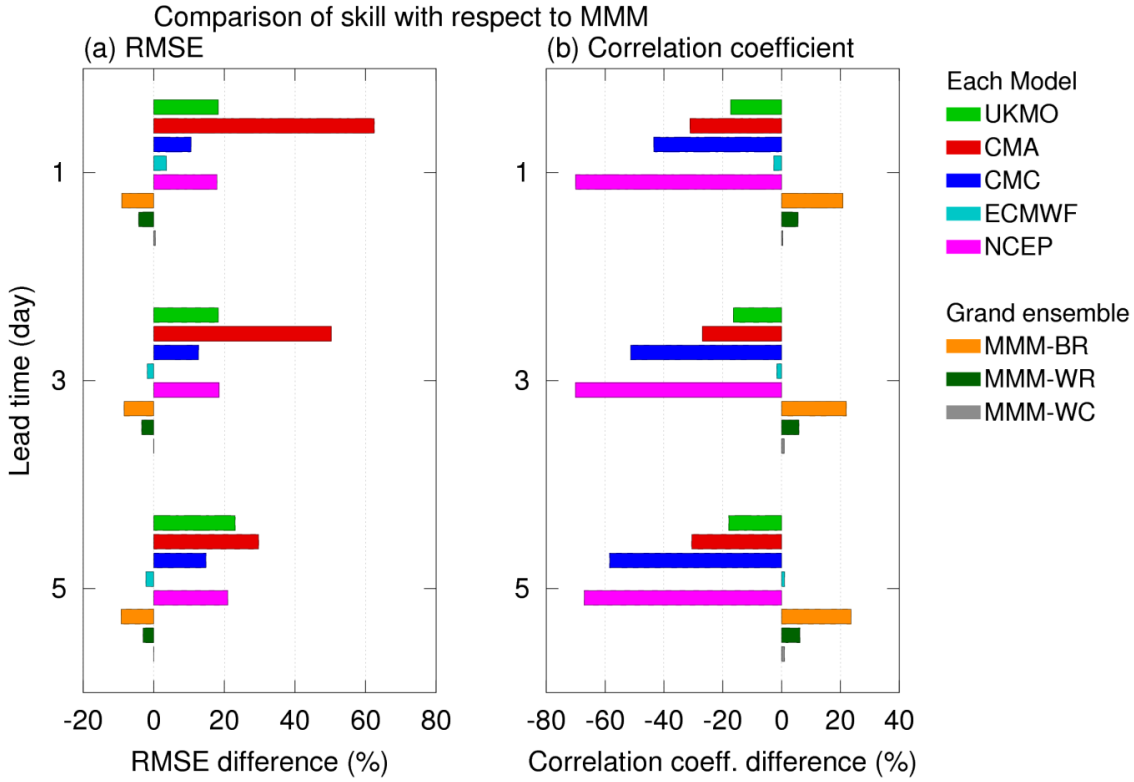


Figure 3.15. Comparison of skill difference of individual EPS and grand ensemble means with respect to MMM (a) RMSE (b) correlation coefficient during JJAS, 2012.

3.3.3.3 Verification of categorical and dichotomous events

Unlike the error statistics, multi-model grand ensemble means show diverse statistics which are used for the verification of frequency forecasts. The discrimination ability of individual EPSs and grand ensemble means to discriminate observed rainfall categories are shown in Fig. 3.16. The frequencies of observed light precipitation and moderate precipitation events are better simulated in MMM-BR (Fig. 3.16g1-7) compared to others. Most of the EPSs and grand ensemble means overestimate the observed heavy precipitation events as moderate (Fig. 3.16). In general, frequency of observed heavy events are not discriminated in all EPSs and grand ensembles and is forecasted as moderate events except in CMA in the first three lead days. The discrimination ability of MMM and MMM-WC basically follow the pattern of UKMO, and MMM-WC follows the pattern of ECMWF. However, observed relative frequencies of moderate and heavy precipitation are larger in grand ensembles than in individual EPSs at longer lead times.

The statistics computed from contingency table vary even for the same model at the same precipitation threshold (Fig. 3.17). It is therefore difficult to draw the explicit conclusions. Generally, skills of EPSs and grand ensemble means are poorer at heavy precipitation threshold ($\geq$25 mmday^{-1}) compared with light threshold ($\geq$1 mm day^{-1}) (Fig 3.17). Improvements in forecast skills using multi-model grand ensemble techniques have been seen only in certain scores. For instance, multi-model grand ensemble means show lowest FAR for day+1, day+3 and day+5 and the highest ETS for day+1 and day+3 at heavy threshold. However, their skill is either lower than or similar to other EPSs. At heavy threshold, ECMEF outperforms all EPSs and grand ensemble means for some scores (POD and Bias). UKMO is the other best performing EPS in 2012. However, FAR of ECMWF and UKMO are larger than FAR of grand ensemble means. Generally, the grand ensemble means (MMM, MMM-WR and MMM-WC) except MMM-BR do not outperform the individual EPSs at light precipitation threshold. The observed frequency at light precipitation threshold is better forecasted in MMM-BR with the lowest frequency bias and FAR, and the highest ETS (0.4) (Fig 3.17a,b,d) as reflected in the discrimination diagram (Fig. 3.16g1,g3,g5). However, POD of MMM-BR is as low as CMA (Fig 3.17c).

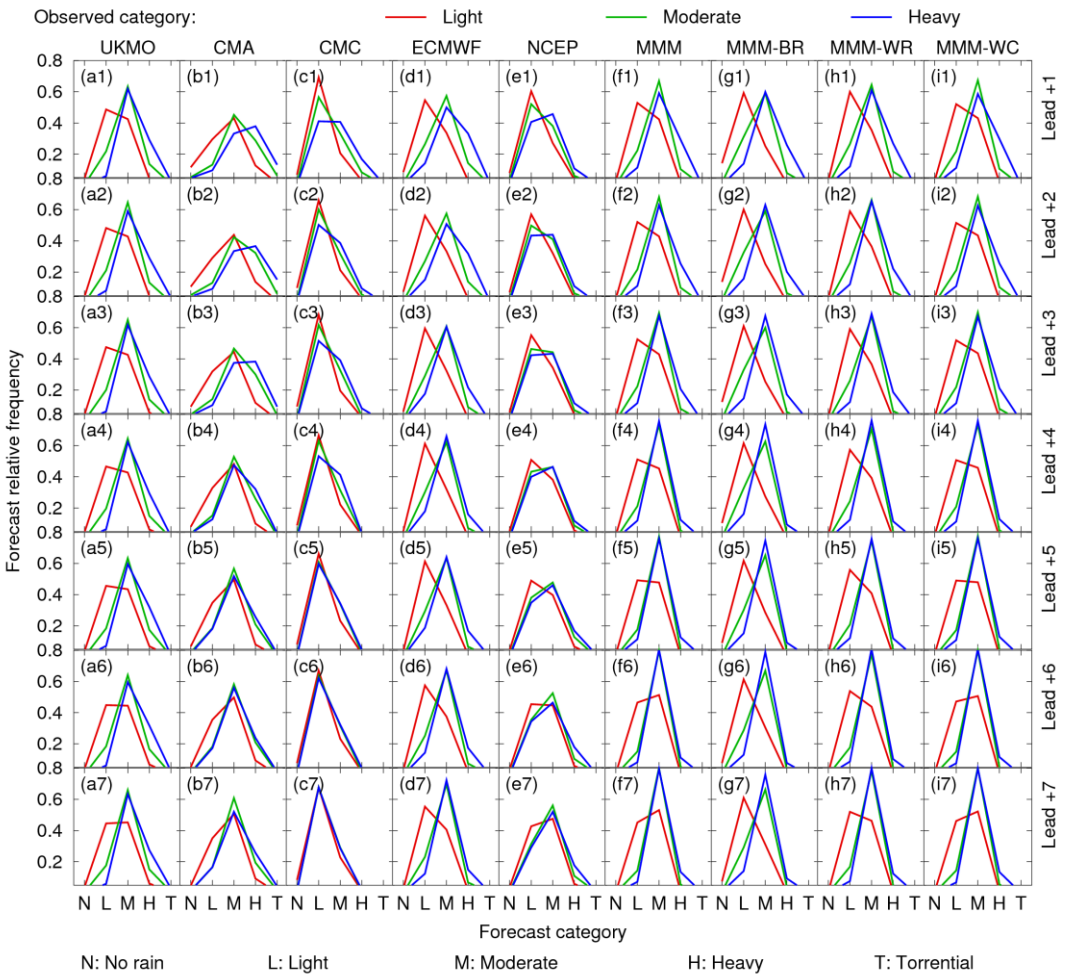


Figure 3.16. Same as Figure 3.7 but during JJAS, 2012.

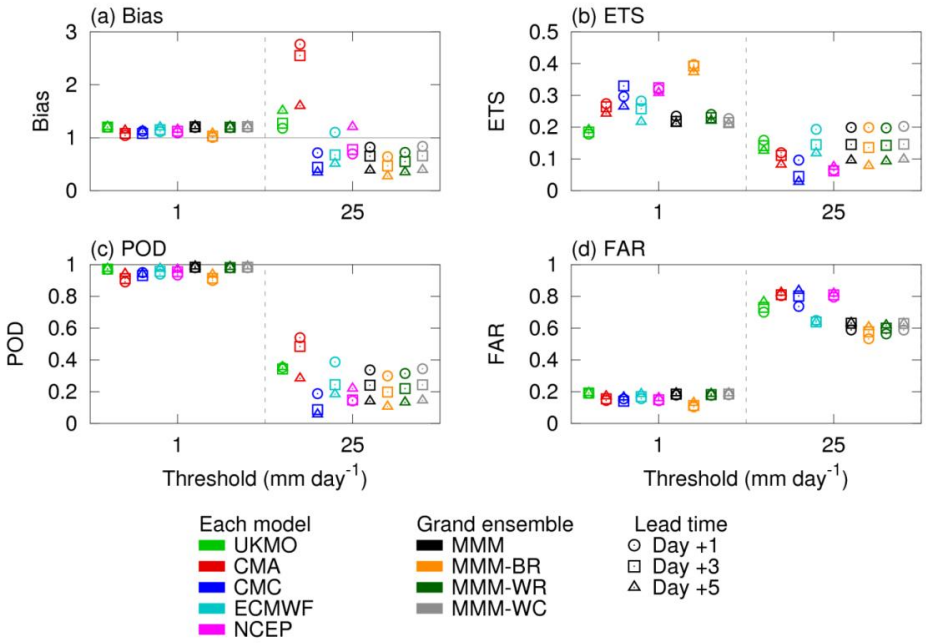


Figure 3.17. same as Figure 3.8 but for only two precipitation thresholds (1, and 25 mm day^{-1}) during JJAS, 2012.

3.4 Conclusion and discussion

In this chapter, verification and comparison of ensemble mean QPFs and PQPFs in Nepal during monsoon season of 2009–2012 from five operational global EPSs (CMA, CMC, ECMWF, NCEP and UKMO) archived in TIGGE database are performed against the APHRODITE's daily precipitation. Additionally, we have examined the performances of multi-model grand ensemble means during monsoon season of 2012. Various verification metrics have been employed to evaluate the performance at different lead times and precipitation thresholds. Skill scores highly depend on the dataset employed in verification *(Yuan et al. 2005).*

The performances of EPSs to forecast precipitation in Nepal, a complex mountainous country, is poor. Comparatively, ECMWF EPS exhibits the highest forecast skills for both QPFs and PQPFs, followed by that of UKMO. The QPFs and PQPFs skill of NCEP is the poorest among the EPSs with no skill to discriminate different precipitation categories. The poor skill of NCEP is due to the shift of precipitation area to northeastwards which could be related to topography and model resolution. Although, CMC has shown the lowest error, good spread-skill relationship and reliability diagram, spatial correlation and discrimination ability of it is very poor. Also, the CMC has lowest conditional bias which is indicated by the small value of reliability, but lacks the extremely high probabilities and is least sharp at heavy thresholds due to the large difference among the ensemble members. Consistent with previous studies in northern tropics *(Su et al. 2014)*, CMA has relatively largest bias for the first three lead days with large frequency bias and false alarm ratio. However, discrimination ability of CMA is relatively better in case of heavy precipitation for first three lead days which could be related to excessive wet bias.

Previous studies have investigated that statistical post processing can improve the forecast skill *(Roy Bhowmik and Durai 2010; Durai and Bhardwaj 2014; Matsueda and Nakazawa 2015).* In this study, multi-model grand ensemble mean created by four simple techniques are compared for monsoon season of 2012. Forecast errors are reduced in grand ensemble means than the individual EPS, with the highest improvement in bias removed grand ensemble. However, the performance of ECMWF is as good as simple multi-model grand ensemble mean. The discrimination ability of different observed precipitation categories is not improved much in grand ensemble means. However, overforecast of observed light precipitation as moderate rainfall by EPSs is significantly reduced in bias removed ensemble mean, which is also reflected by low frequency bias and FAR and high ETS.

The deficiency of ensemble spread, which is the widely noticed problem of EPS, also apparent in this study. All the TIGGE EPSs are under dispersive. The prediction systems should incorporate more sophisticated methods for simulating the model uncertainty, which ultimately will result spread more consistent with the ensemble-mean error, as expected by theory *(Swinbank et al. 2016).*

Although, this study is limited to investigate the performances of monsoon season precipitation forecast of monsoon season, the findings of this study are expected to be helpful not only to the model developers and users, but also provide the valuable information to decide the appropriate technique of post-processing to improve the forecast skill of EPSs. It is important to note that EPSs are upgraded (e.g., higher resolution, improved initialization procedures, and improved parameterizations) during the course of analysis, this study does not pay any attention towards this direction. However, previous study has investigated the influence of

system upgrade *(Su et al. 2014)*. Additionally, the TIGGE EPSs are subjected to upgrade regularly, the systematic biases in each EPS's mean forecast will decrease, and the results presented in this study will not necessarily reflect the performance of these systems at the current time.

In order to gain the comprehensive knowledge of TIGGE EPSs, in addition to the precipitation forecast skill studied here, many verification studies should be carried out for more systems (e.g., JMA, BoM, and KMA), for different variables at different levels (e.g., wind at 800 hPa and 200 hPa, temperature at 2 m) in the future. Such studies can be helpful to immensely improve the EPS performance and operational forecasters to understand the uncertainty of the concerned models.

Chapter 4

Reproducibility of an intense precipitation event in Nepal in JMA-NHM

4.1 Introduction

Extreme precipitation events can severely impact the society causing severe floods and landslides, damaging agricultures and infrastructures, and consequently obstructing the overall socio-economic development. The disaster risk is higher in developing country which has poor infrastructures and economy such as Nepal. Situated in the Central Himalaya, Nepal is highly prone to the developments of deep and severe convective events *(Houze et al. 2007; Shrestha 2016)* leading to heavy rainfall that trigger the natural disasters such as flash floods and landslides, cause huge loss of lives and properties and deterioration of ecosystems. In Nepal, precipitation-induced disasters are the major life threating and property damaging disasters. Moreover, observed heavy precipitation events are increasing in Nepal particularly in western parts *(Baidya et al 2008; Karki et al 2017a)* and are projected to increase in future under warming scenarios which we have investigated in the chapter 2.

Given the high exposure and vulnerability of the region to the extreme precipitation events, it is vital for Nepal to have accurate predictions of such high-impact weather events to devise proactive plans in order to reduce and mitigate related disasters. However, as we have investigated in chapter 3, the state-of-the-art global operational models sophisticated with ensemble prediction systems are still inadequate to forecast precipitation in Nepal accurately, in particular heavy precipitation events. The poor performance of the model could be related to the complex topography of Nepal which is not sufficiently represented in the coarse resolution models, and also could be related to the imperfect model configurations such as physical parameterizations and model dynamics. *Wehner et al. (2010)* examined the extreme precipitation simulation over the continental United States using atmospheric general circulation model (AGCM), and investigated that high resolutions models are more reasonable to reproduce the extreme precipitation amount than coarse resolution, and coarse resolution models underestimate the precipitation amount. However, running high resolution AGCM is computationally very expensive. It is generally anticipated that higher resolution regional climate models (RCMs) would be able to add value in simulation owing to their ability to better simulate fine temporal and spatial scale features and variability *(Rummukainen 2010; Feser et al. 2011; Karmacharya et al. 2016)*, are widely used to study local atmospheric features in recent decades.

RCM uses large-scale meteorological fields from GCM runs or global reanalysis dataset as initial and time-dependent lateral boundary conditions (LBC) for high resolution simulations. The performance of the RCMs depends on various factors such as initial and boundary conditions *(Wu et al. 2005)*, model domain size and location *(Jones et al. 1995; Leduc and Laprise 2009)*, model resolution *(Giorgi and Marinucci 1996)*, physics parameterization *(Giorgi and Marinucci 1996)* or a combination of several of those factors. The detail issues of RCMs are described in *Giorgi and Mearns (1999)* and *Rummukainen (2010)*.

A number of studies conducted on numerical simulation over the Himalaya and south Asia investigated that high resolution RCMs can realistically simulate precipitation in comparison to observations *(Rupa Kumar et al. 2006; Kulkarni et al. 2013; Menegoz et al. 2013; Collier and Immerzeel 2015; Karmacharya et al. 2016; Norris et al. 2016; Karki et al. 2017b)*. However, most of them reported the capability of RCMs to simulate

climatic features. *Rupa Kumar et al. (2006)* employed RCM to simulate precipitation by covering most of South Asia. They investigated that RCM is able to capture the spatial pattern of summer monsoon rainfall in India. Another numerical experiment by *Ji and Kang, (2012)* examined that RCM simulated precipitation does not represent as good as temperature over the Tibetan Plateau. *Kulkarni et al. (2013)* concluded that RCM is well capable in simulating the spatial pattern of precipitation over the central Himalaya. However, they noticed wet bias over the Himalayan region in monsoon season. Significant precipitation bias over complex terrain mountainous regions where topography influences orographic precipitation and monsoon circulation is also reported in *Solman et al. (2013)*. In spite of common findings such as over estimation of precipitation at mountainous regions, *Menegoz et al. (2013)* investigated the dry bias over the central Himalaya running RCM in a small domain. However, they claimed that their model is able to simulate snow cover reasonably.

In a study by *Collier and Immerzeel (2015)* investigated that models with very high spatial resolution (5 and 1 km) capture the main features of observed meteorological variability (spatial, seasonal and diurnal) in the catchment located in the high mountainous regions of Nepal. They further revealed that improvement in precipitation simulation is comparatively lower than temperature when model is run down to 1 km resolution from 5 km. *Karmacharya et al. (2016)* performed two experiments with RCMs at medium (0.44°) and high (0.11°) spatial resolutions to simulate south Asian monsoon precipitation, and revealed that compared to the medium resolution RCM, the high resolution RCM is able to better resolve the interaction of the low level monsoon flow with the Himalayan orography leading to added value in simulating many aspects of precipitation such as the seasonal mean, relative frequency distribution of daily precipitation, and various metrics of precipitation extremes. However, they noted that the added value is more over the Indo-Gangetic plain than over the complex Himalayas. There are other very recent reports that show the improvements of spatial, seasonal and diurnal variability of precipitations in very high resolution RCMs (<10 km) examined in the mountainous catchments of Nepal *(Norris et al 2016; Karki et al, 2017b)*.

Study on simulation of extreme precipitation events in RCMs over Nepal and Himalaya is still rare *(Das et al. 2006; Maharjan and Regmi, 2015; Shrestha et al. 2015)*. *Das et al. (2006)* examined an extreme rainfall event occurred for a short time, called cloudbrust, in the Himalaya on 16 July 2003 using very high resolution model (3 km) and found that model is capable to predict rainfall amount with location bias of cloudburst tens of kilometers. *Maharjan and Regmi (2015)* examined an extreme precipitation event that occurred in western Nepal on 18 August, 2001 using a very high resolution model (1 km) and found that although observed precipitation is significantly underestimated, the model is capable to reproduce the essential features of the precipitation. *Shrestha et al. (2015)* ran very high resolution (2.8 km) RCMs to examine the mechanism of an extreme precipitation event that occurred over the Uttarakhand region of the central Himalayas on 13 September 2012, and found that the spatial pattern of daily-accumulated precipitation and atmospheric state profiles simulated by the model compared well with the observations and suggested underlying mechanisms. *Ngo-Duc et al. (2017)* examined 12 extreme indices of both rainfall and temperature (7 rainfall and 5 temperature) using the 21 years (1989-2008) simulations from a RCMs with 18 different configurations at spatial resolutions of 36 km, investigated that extreme rainfall events are not simulated as good as temperature events. In this regard, several studies are essential to fully describe the capability of RCMs to simulate extreme events that occurred in Nepal and the Himalayas.

In this chapter, we propose to evaluate the performance of a RCM developed by the Japanese research communities in simulating an intense precipitation event occurred on 13-15 August, 2014 in Nepal. With the introduction in section 4.1, this chapter is arranged as follows. The synoptic overview of the event is given in section 4.2 followed by model configuration and analysis data in section 4.3. The results are presented in section 4.4 and conclusions and discussions are presented in section 4.5.

4.2 Extreme precipitation event and synoptic condition

An extreme precipitation event occurred on 13-15 August, 2014 (Fig. 4.1) in Nepal caused widespread floods in the southern flat land, flash floods in several rivers and landslides in several places, claimed many lives and left many people injured (http://lib.icimod.org/record/31711/files/icimodNDR015-Ch4.pdf) (Photos in Appendix C). The unprecedented event concentrated over the southwestern parts on 14 August released more than 400 mm rainfall in 24-hour over the region (Fig 4.1b, left panel). A rain-gauge located in the extreme event area broke its historical record of the highest daily rainfall recording (423.1 mm in 24 hr from 03UTC, 14 August to 03UTC 15 August) and monthly total rainfall of the August (Fig. 4.2a,b). Heavy precipitation over the region is largely underestimated by TRMM (> 160 mm) on 14 August, 2014 (Fig.4.1b right panel and Fig.4.2a).

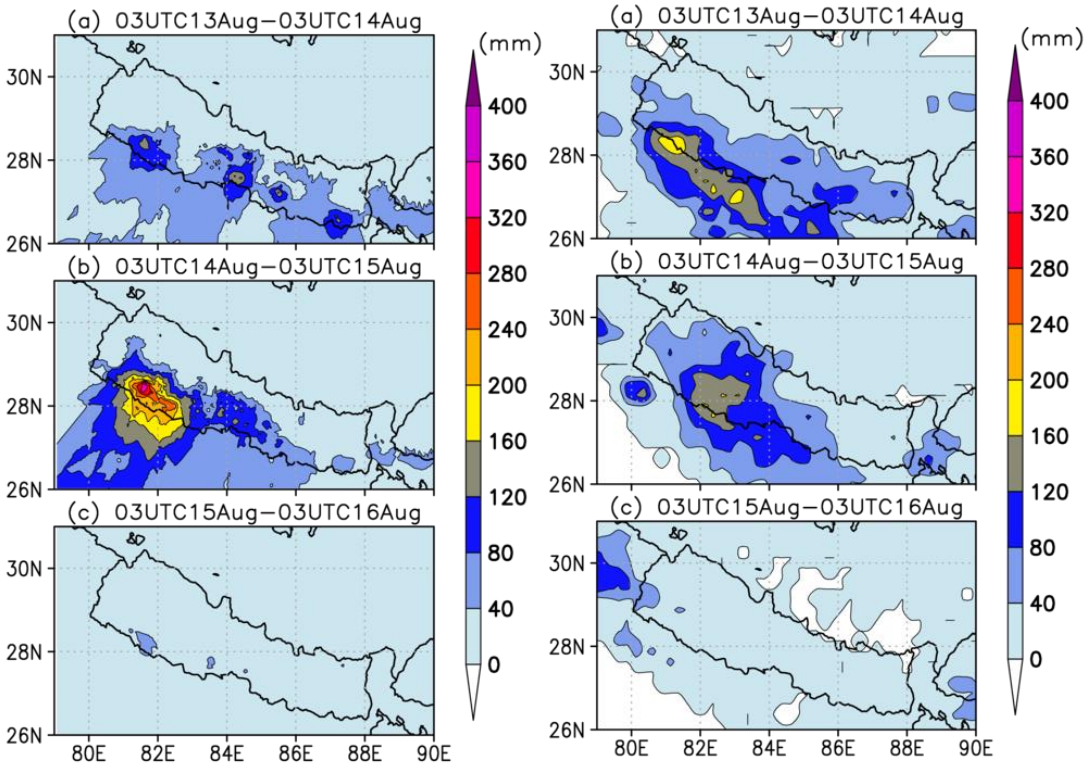


Figure 4.1. 24-hour accumulation of precipitation during 13-16 August 2014 measured at synoptic time in Nepal (a) 03 UTC 13 August–03 UTC 14 August 2014 (b) 03 UTC 14 August–03 UTC 15 August 2014 (c) 03 UTC 15 August–03 UTC 16 August 2014. Left panel APHRODITE and right panel TRMM.

During the event, an upper-level extra-tropical westerly trough over the northwest of the Nepal (although weak) and a shift of the westerly and easterly jets towards the Himalayas (Fig. 4.14c) are observed which are the common synoptic conditions observed during the extreme event *(Shreshtha. 2016)*. Additionally, a northward shift of the monsoon trough towards the Himalayas and the extension of westerly surface wind flow directly from the Arabian Sea to the Himalayas (Fig 4.12c) are the other common conditions during the extreme event *(Shreshtha. 2016)*. Moreover, in such events, upper-level positive potential vorticity anomaly, associated with a steep tropopause fold around the westerly jet over the Himalayas north of the extreme event regions, is also reported *(Shreshtha. 2016)*.

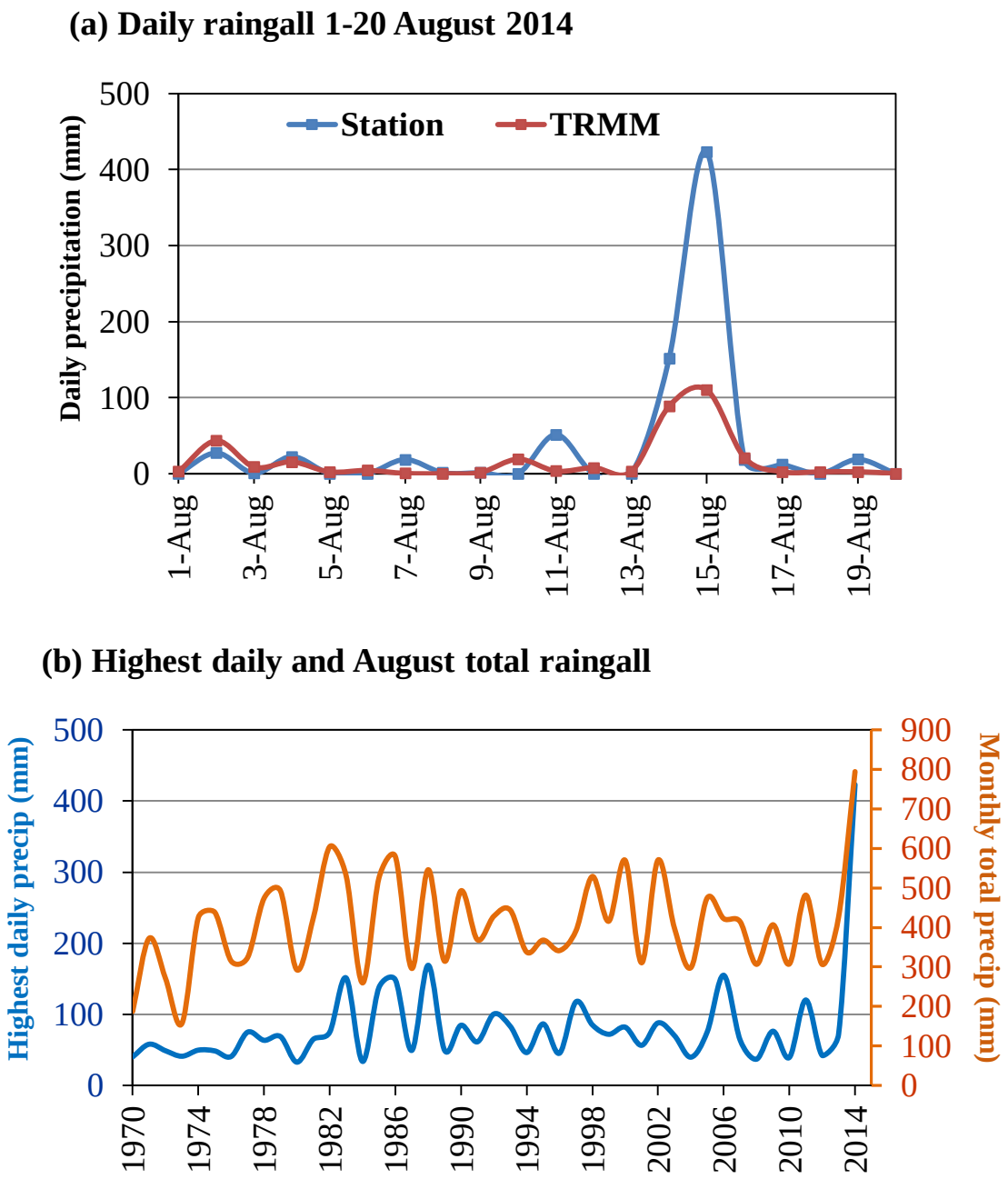


Figure 4.2. a) Daily precipitation 1–20 August 2014, b) monthly total rainfall in August and the annual highest daily rainfall at Surkhet, Nepal (28.6 °N and 81.6 °E).

4.3. Data and model configuration

4.3.1 Verification data

As mentioned in precious chapter that the recent APHRODITE's precipitation dataset is available only for Nepal, we have used Tropical Rainfall Measuring Mission (TRMM) verion3B42v7 satellite precipitation product *(Huffman et al. 2007)* as verification datasets. TRMM 3B42V7 is available in gridded format with 3 hour temporal resolution and 0.25° spatial resolution extending quasi-globally (50°S –50°N and 0–360°) since 1998. 3B42V7 estimates are produced using calibrated microwave precipitation estimates which are further combined with infrared estimates.

Additionally, we have used JRA55 *(Kobayashi et al. 2015)* for analysis of synoptic conditions and verification of simulated atmospheric fields.

4.3.2 Model configuration

The RCM employed in this study is the Japan Meteorological Agency non-hydrostatic model (JMA-NHM; hereinafter NHM) *(Saito et al. 2007)*. NHM is fully compressible non-hydrostatic model. The model is set up for two domains configured at 25 and 5 km using one-way nesting approach for two sets of experiments namely big domain and small domain. Big domain experiment at 25 km (128×105, hereinafter BD25) covers 16.26–40.78° N and 64.06–102.22° E (Fig. 4.3a), and the corresponding nested 5km domain (341×241, hereinafter BD5) (Fig 4.4a) spans 22.36–33.40° N and 73.76–92.22° E. Similarly small domain experiment at 25 km (84×68, hereinafter SD25) covers 20.16–35.58° N and 72.37–95.62° E (Fig. 4.3b), and the corresponding nested 5km domain (120×100, hereinafter SD5) spans 25.76–30.21° N and 78.90–85.10° E (Fig. 4.4b). All the experiments are run in Lambert Conformal map projection with 50 vertical levels (lowest grid: 40m and highest: 1240m) with model top located at about 30km.

In Himalaya, convection plays important role in geographical distribution and accumulation of precipitation *(Houze et al. 2007)* enhanced by intertwined effect of topography and southeasterly moisture flow from Bay of Bengal and Arabian sea during summer time. In RCM, the topographical effect can be modulated by the parameterization schemes used for cumulus convection *(Giorgi and Marinucci 1996)*. In this study, we have used the Kain–Fritsch (KF) cumulus parameterization scheme *(Kain and Fritsch 1990; Kain 2004)* only in BD25 and SD25 experiments, and turned off in BD5 and SD5 as precipitation output may worsen by using cumulus parameterization schemes compared to convection permitting experiment on that scale *(Collier and Immerzeel 2015)*. Furthermore, we have chosen KF-scheme as this comparatively produces reasonable precipitation over the Himalaya *(Kadel 2014, Collier and Immerzeel 2015, Karki et al. 2017b)*.

The bulk method including the ice phase was chosen for cloud microphysics *(Ikawa and Saito, 1991)* and for turbulence, the Mellor–Yamada–Nakanishi–Niino Level 3 scheme *(Nakanishi and Niino 2004)* is employed. Spectral boundary coupling (SBC) method *(Yasunaga et al. 2005)* is employed for the adjustment of inconsistencies between the external and internal model states. In the SBC method, the large-scale wave components of internal model values are replaced by those of external model values with an interval. A more complete description of the NHM is provided in *Saito et al. (2007)*.

Experiments at 25km (BD25 and SD25) are initialized and forced at boundary by the European Centre for Medium-Range Weather Forecasts (ECMWF) latest global atmospheric 6 hourly reanalysis dataset: ERA-Interim *(Dee et al. 2011)* which has spatial resolution of 0.75°×0.75°. National Oceanic and Atmospheric Administration (NOAA) daily optimum interpolation sea surface temperature (OISST) observed by the Advanced Very High Resolution Radiometer (AVHRR) version 2 *(Reynolds et al. 2007)* has been used. Both OISST and boundary conditions are updated at every 6 hours. OISST-AVHRR is provided with spatial resolution of 0.25°. Topography dataset used in this study is GTOPO30 developed by U.S. Geological Survey. The model is run for 6 days simulation starting from 00UTC 10 August in case of 25 km experiments, and 5 days simulation starting from 00UTC 11 August in case of 5km experiments. Simulation outputs for outputs are analyzed only for 13–15 August, 2014.

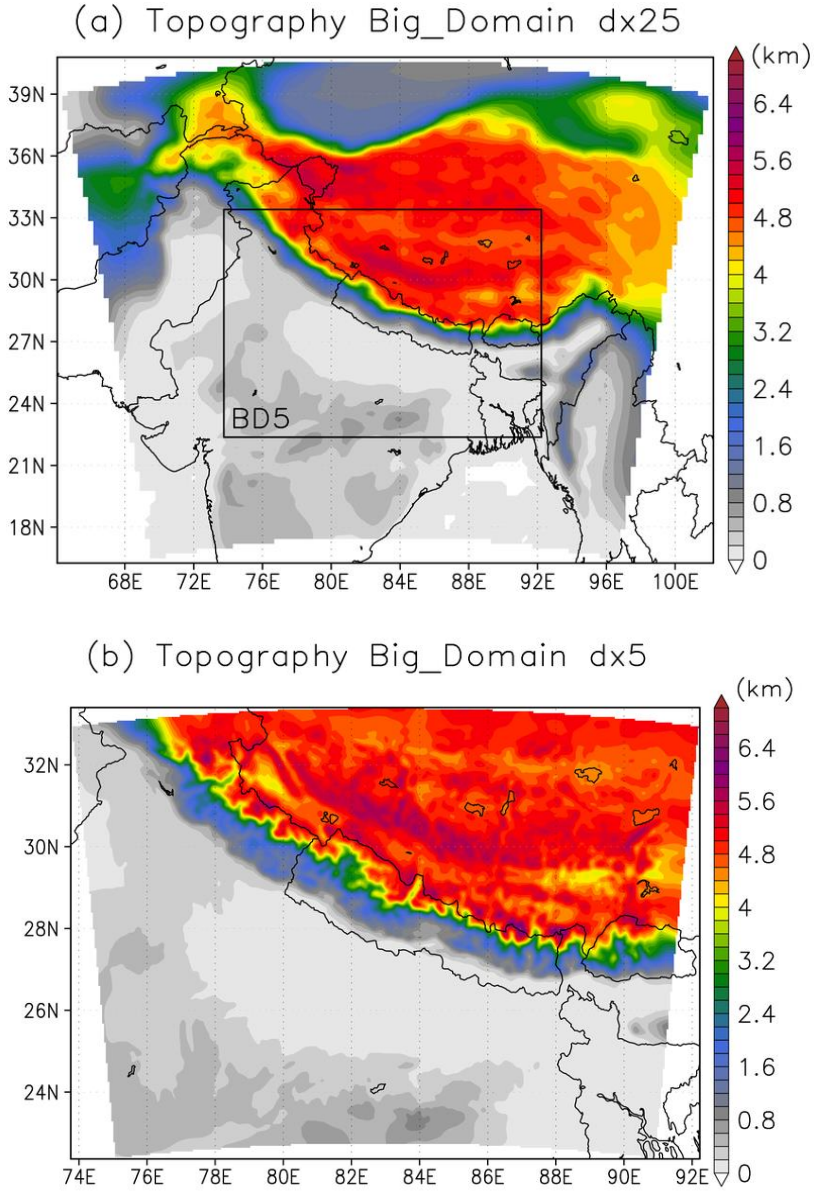


Figure 4.3. Model domain and topography of big domain experiment. a) 25km domain. Box indicates nested domain of 5km resolution. b) 5 km horizontal resolution nested domain indicated in a). Topography in b) is smoothed confirm stable model integration.

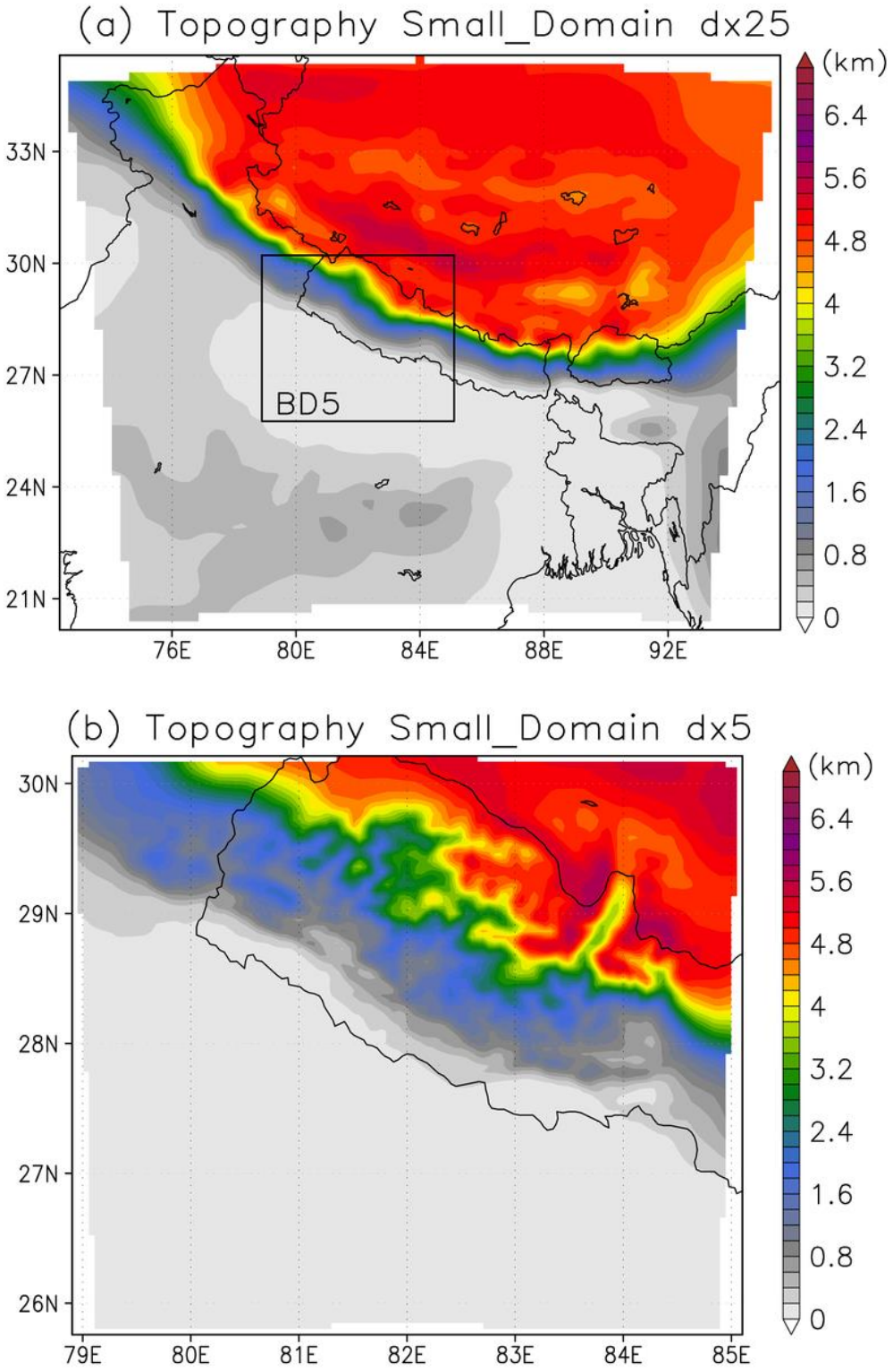


Figure 4.4. Same as Fig 4.3 but for small domain experiment. In this case, topography in b) has not been smoothed.

4.4 Results

The simulation of an extreme event by JMA–NHM is analyzed and presented in this section. Firstly, we validate the simulated precipitation and then we present the simulated atmospheric variables such as wind and geopotential height.

4.4.1 Verification of precipitation

The results show that BD25 and SD25 are capable to indicate the heavy precipitation (simulated precipitation >60 mm) in the target area, show identical broad features of the simulated precipitation in two experiments which indicates the low impact of domain size. However, the spatial coverage of the targeted event is largely reduced and shifted slightly towards north (Figs. 4.5–4.7). The most important thing to be noticed is the consistent extreme wet bias over the eastern Nepal during the event period (Figs. 4.5–4.7). Unlike in BD25 and SD25, simulated precipitation patterns in BD5 and SD5 vary each other and from day to day during the event period (Figs. 4.8–4.10). The improvement relative to the parent domain is apparent in the case of BD5 with reduction of excessive wet bias that is noticed in BD25 (Figs. 4.8–4.10a) however degradation is clear in the case of SD5 on 14 and 15 August (Figs. 4.9 and 4.10b). The SD5 is capable to simulate extreme precipitation on 13 August (Fig. 4.8b) while it failed to reproduce on 14 August (Fig. 4.9b) and show excessive wet bias on 15 August (Fig. 4.10b). The results indicate that domain size has large impact on the simulated precipitation in the case of 5 km experiments showing bigger domain is better than smaller domain.

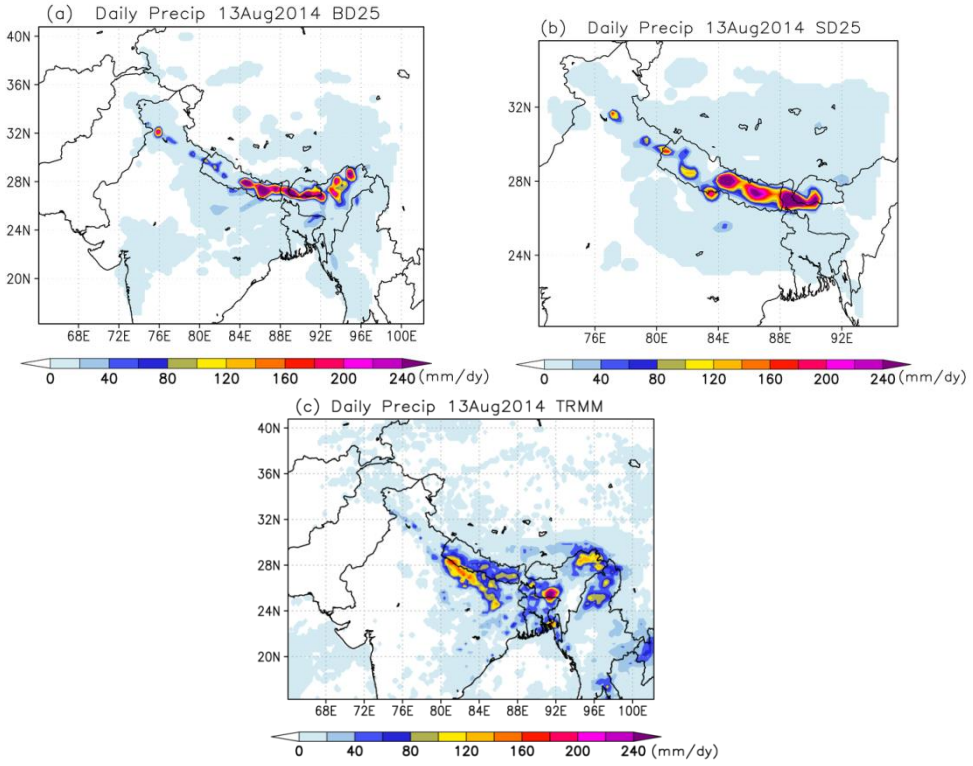


Figure 4.5. Daily precipitation (00 UTC – 00 UTC) on 13 August, 2014 a) simulated by NHM at 25km using big domain, b) simulated by NHM at 25km using small domain, c) TRMM

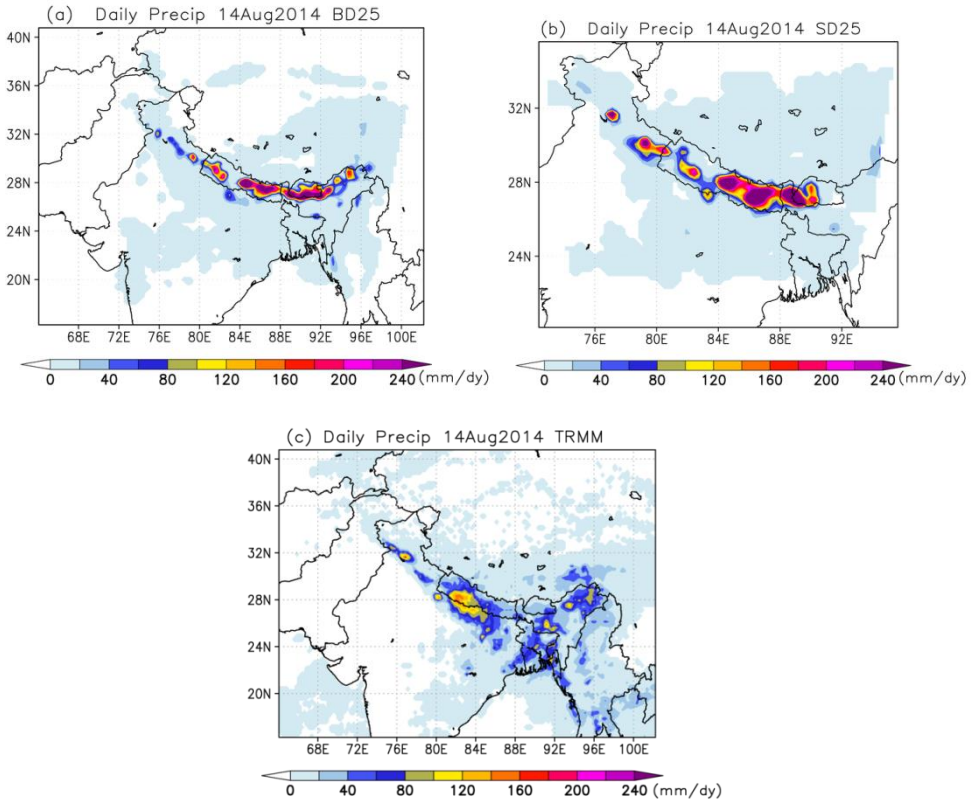


Figure 4.6. Same as figure 4.5 but on 14 August, 2014.

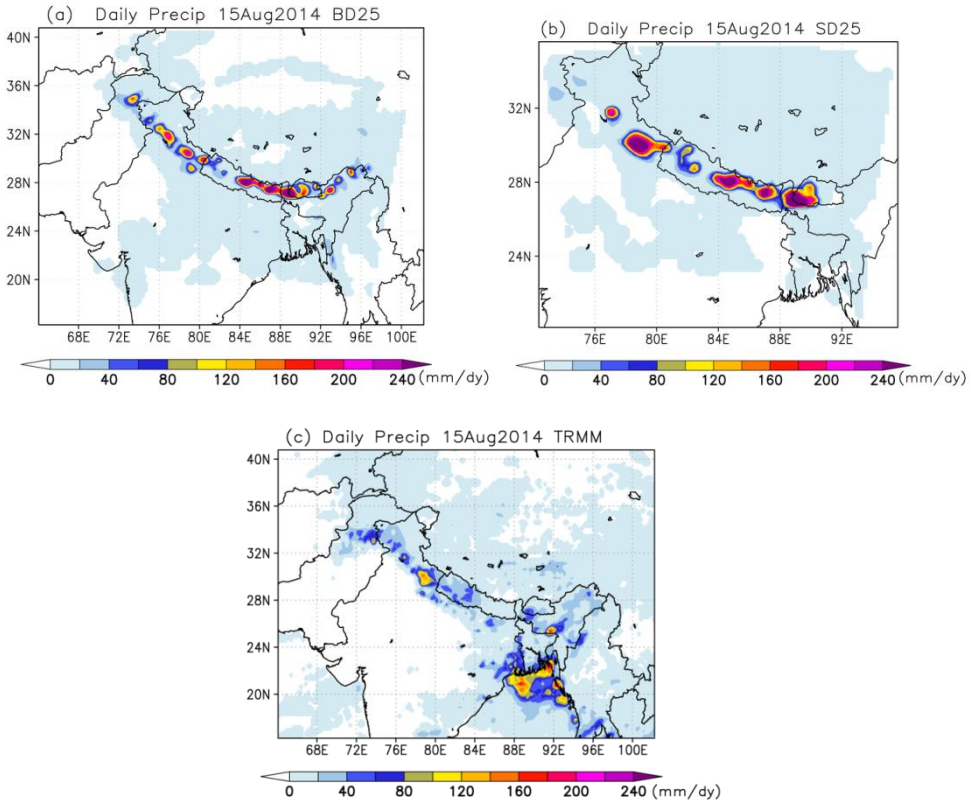


Figure 4.7. Same as figure 4.5 but on 15 August, 2014.

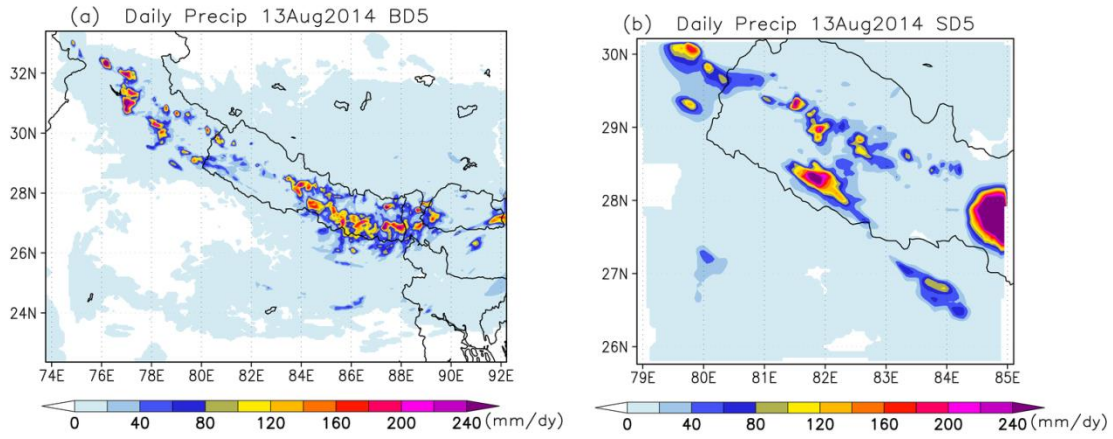


Figure 4.8. Daily precipitation simulated by 5 km models on 13 August, 2014 a) big domain b) Small domain on 15 August, 2014.

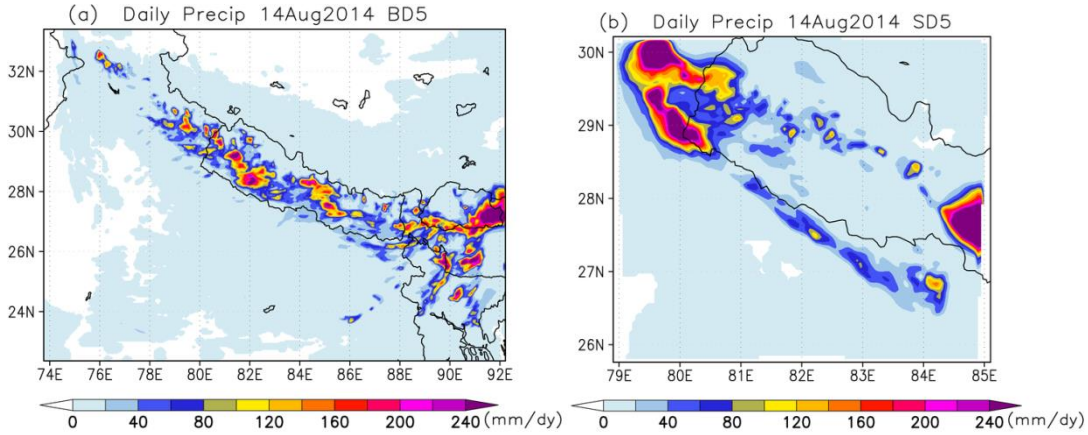


Figure 4.9. Same as figure 4.8 but on 14 August 2014.

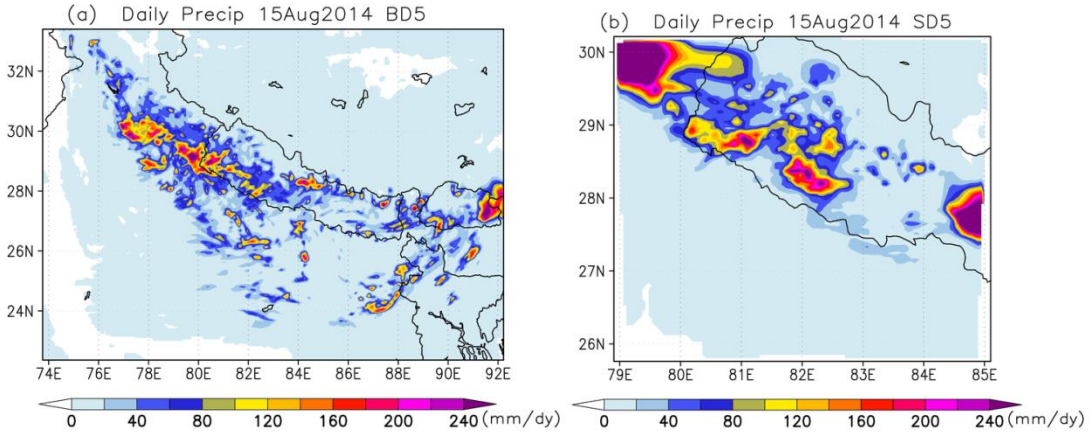


Figure 4.10. Same as figure 4.8 but on 15 August 2014.

4.4.2 Analysis of simulated atmospheric variables

In the case of BD25 and SD25, simulated mean sea level pressure (MSLP) over the Tibetan plateau is very low (Figs. 4.11a,b) compared to the JRA55 (Fig. 4.11c), causes strong wind flow from the Arabian sea and the Bay of Bengal which converges in the foothills of eastern Nepal, contributed heavy precipitation over the region. Similar wind pattern is also reflected at lower level (850 hPa level). The lower level flow is stronger in BD25 than SD25 and JRA55 (Figs 4.12a-c). Similarly, penetration of westerly jet stream reaches more south in BD25 compared to SD25. Although, in general, the broad pattern of circulation is captured by the models, the pattern of geopotential height is not well simulated at upper (200 hPa) level (Figs. 4.13 and Appendix C).

In the case of BD5 and SD5, simulated features (MSLP and surface wind; lower and upper level wind and geopotental height) are better captured in BD5 compared to SD5 (Figs. 4.14-4.16, Appendix C). The low level flow is northwesterly in the case of SD5 whereas it is south westerly in the event area supplying necessary moisture for favorable condition of heavy rainfall (Figs. 4.14 and 4.15). However, on 13 August, when SD5 is reasonable to capture rainfall, low level circulation is towards the event area (Figs. C.7 and C.8 in Appendix C).

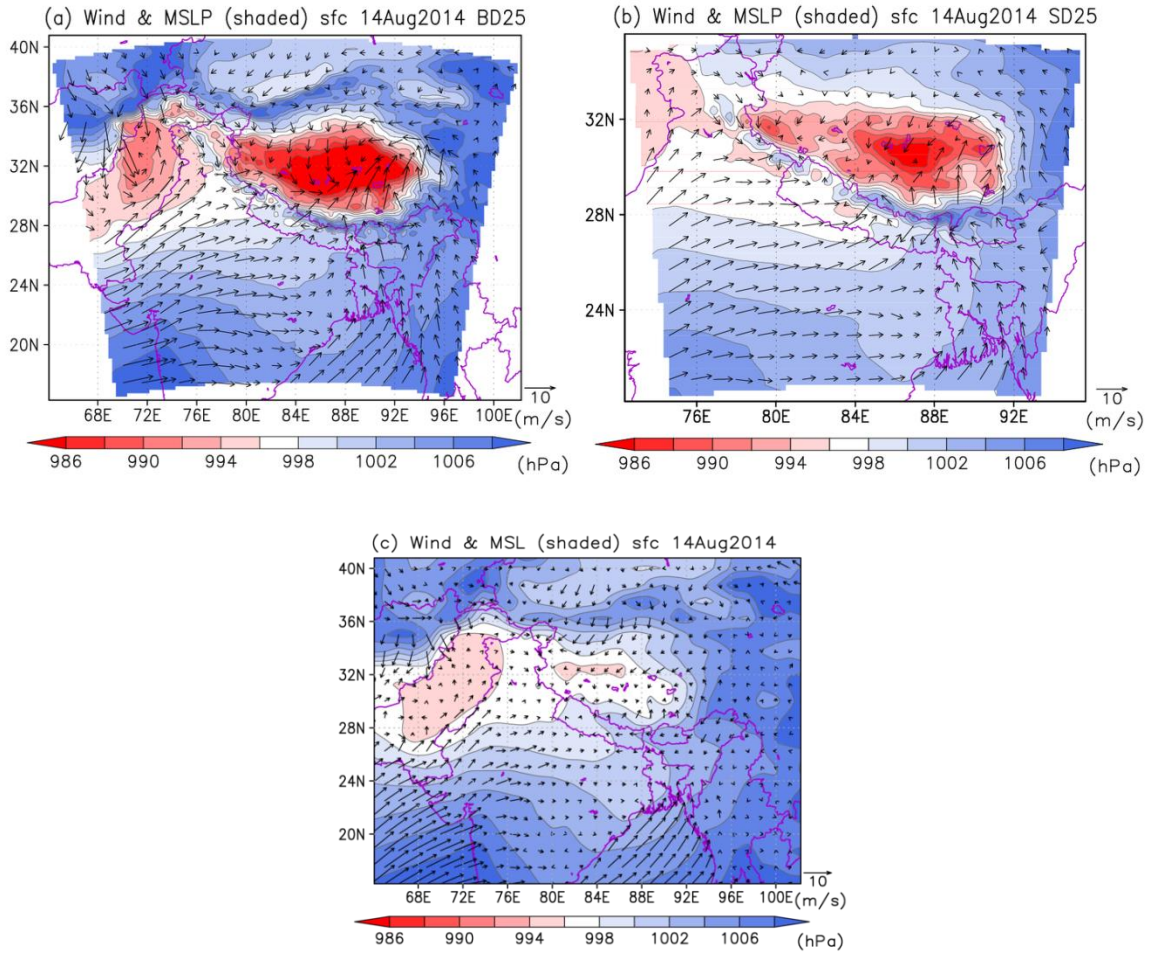


Figure 4.11. Mean sea level pressure (shaded, hPa) and wind at 10 m (vector, m/s) on 14 August 2014 a) BD25, b) SD25, and c) JRA55.

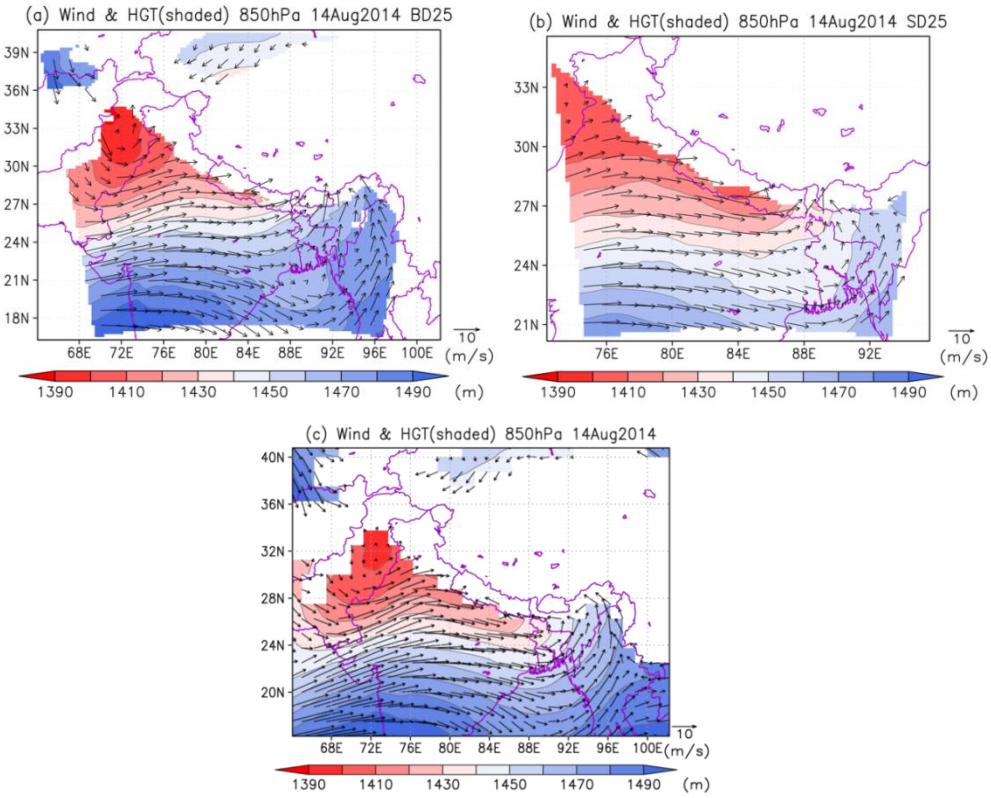


Figure 4.12. Geopotential height (shaded, m) and wind at 850 hPa (vector, m/s) on 14 August 2014 a) BD25, b) SD25, and c) JRA55.

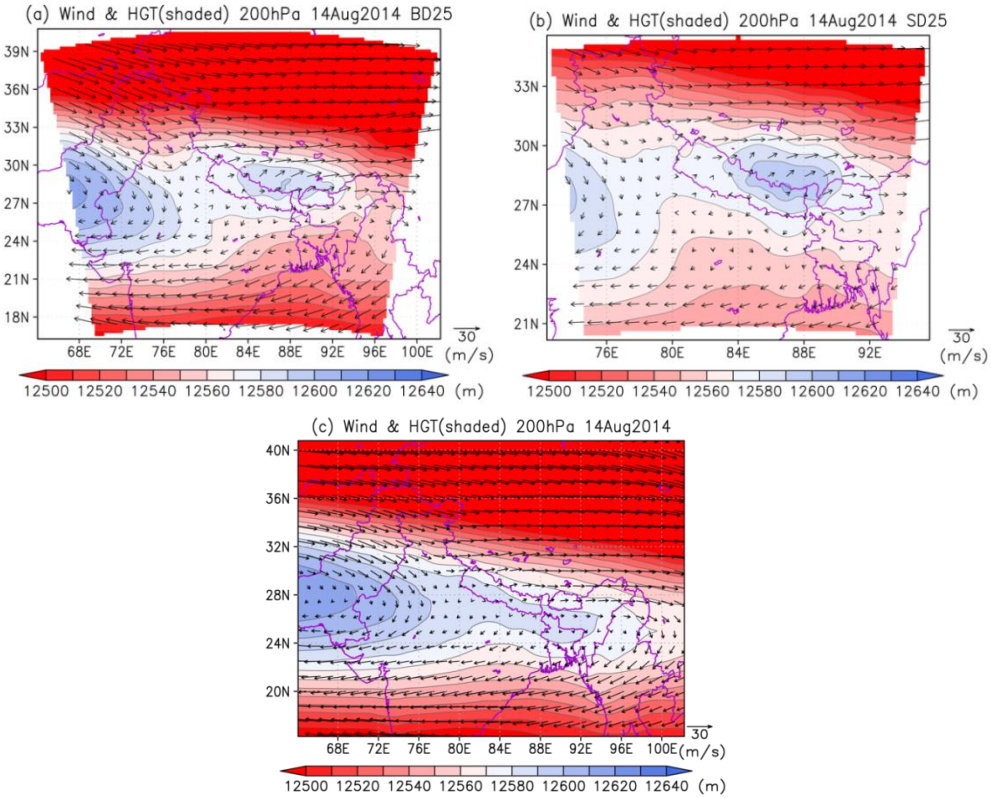


Figure 4.13. Geopotential height (shaded, m) and wind at 200 hPa (vector, m/s) on 14 August 2014 a) BD25, b) SD25, and c) JRA55.

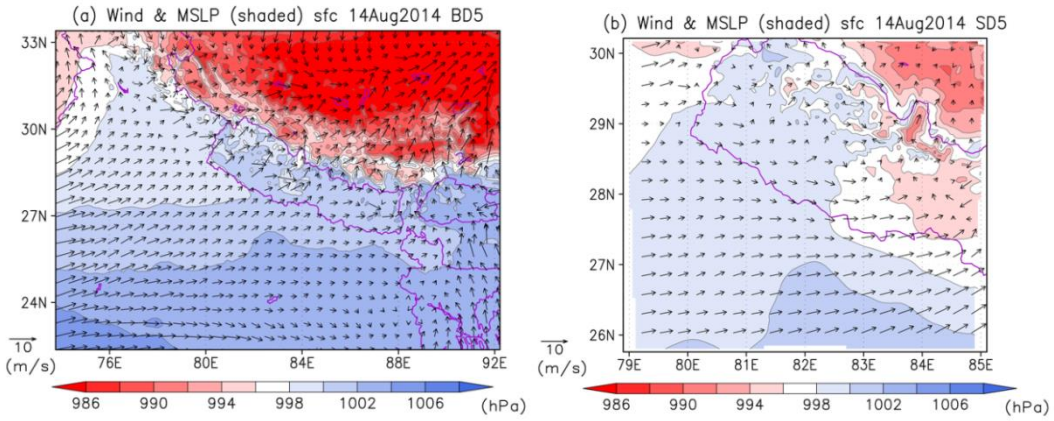


Figure 4.14. Mean sea level pressure (shaded, hPa) and wind at 10 m (vector, m/s) on 14 August 2014 a) BD5, and b) SD5.

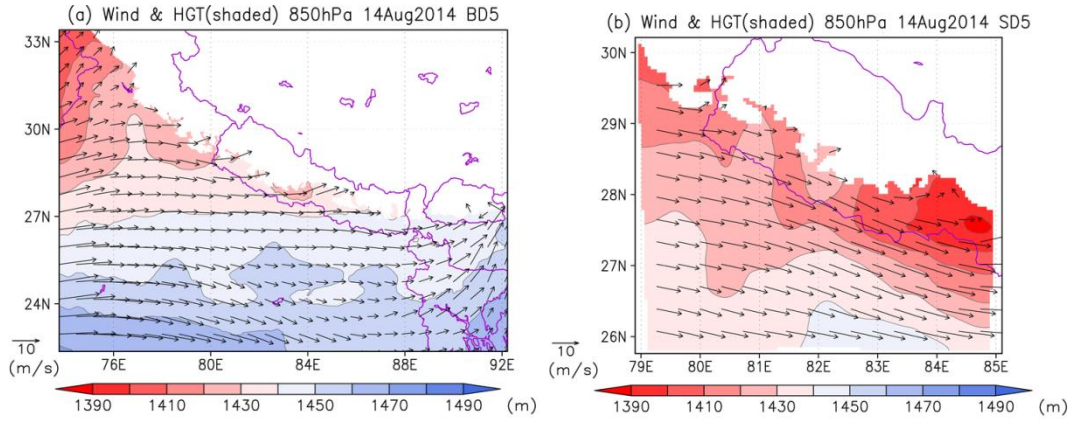


Figure 4.15. Geopotential height (shaded, m) and wind at 850 hPa (vector, m/s) on 14 August 2014 a) BD5, and b) SD5.

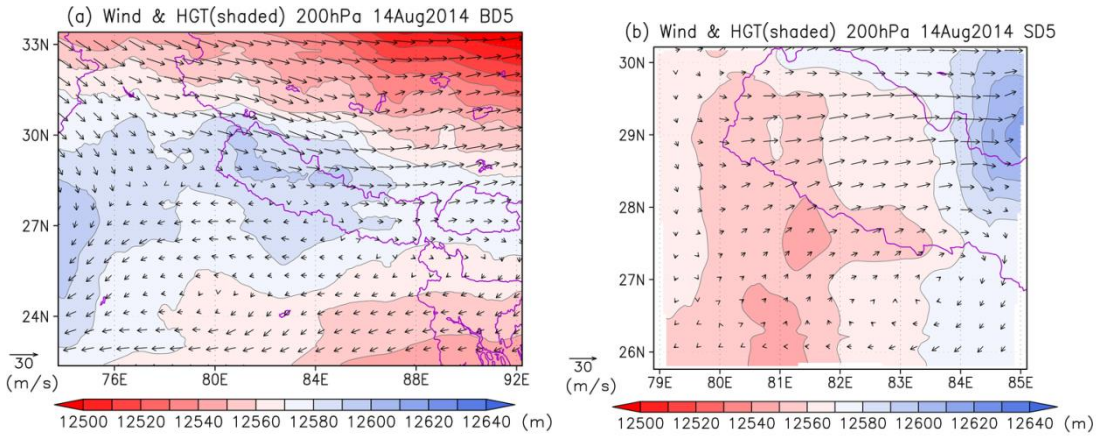


Figure 4.16. Geopotential height (shaded, m) and wind at 200 hPa (vector, m/s) on 14 August 2014 a) BD5, and b) SD5.

4.5 Conclusion and discussion

In this chapter, we have investigated the skill of JMA−NHM to reproduce an intense precipitation event that occurred on 13-15 August, 2014 in Nepal using four experiments: two at 25 km horizontal resolution with big and small domain sizes (BD25 and SD25) and two at 5 km with nested domains (BD5 and SD5). The BD25 and SD25 are initialized and forced at boundary by JRA55 reanalysis data and NOAA OISST, and integrated for 6 days initializing at 00UTC 10August 2014. The nested experiments are initialized at 00UTC of 11 August 2014 and are integrated for 5 days. All the experiments are run with the same model settings except domain sizes, and the cumulus parameterization scheme is turned off in the case of BD5 and SD5.

The results show that BD25 and SD25 are generally capable to simulate the circulation to some extent however the simulated precipitation is underestimated in the targeted area of the extreme events and extremely over estimated in the eastern parts. The unrealistic precipitation could be related to topography, model resolution and cumulus parameterization schemes partly; and also the initial and boundary conditions because low level flow which is expected to converge or directed towards the event area is not happen in JRA55 as well (Fig. 4.12c). Overestimation of daily rainfall intensity by JMA−NHM is also reported in *Cruz and Sasaki (2017)*.

The poor performance 5km runs basically SD5 on 14 and 15 August could be caused by unsuitable domain size. However, achieving added value from high resolution RCM simulation is an important issue because simply increasing the model resolution does not automatically always result in improved model performance *(Rauscher et al. 2010; Karmacharya et al. 2016)*. Many previous studies have reported that high resolution simulations may lack any added value or even show degradation owing to problems in the lateral forcing *(Warner et al. 1997)*, domain size *(Vannitsem and Chome 2005; Leduc and Laprise 2009)*, parameterization schemes *(Jiao and Caya 2006)*, or a combination of several of those factors *(Castro et al. 2005)*. For instance, overall degradation in precipitation climatology at higher resolution is reported by *Pope and Stratton (2002)*. Lack of improvement with higher resolution is also documented in the recent studies *(e.g. Dobler and Ahrens 2010; Haslinger et al. 2013)*. Moreover, seasonality and geographical location also possess large impacts on model performance at the higher resolution as these factors control the climate, and hence the precipitation type *(Iorio et al. 2004; Rauscher et al. 2010; Chan et al. 2013)*. A study reported the failure of RCMs to simulate summertime over East Asia is partly caused by the models' failure to capture the features of tropical cyclones *(Zhong, 2006)* which indicates that models are still lacking to address the extreme weather events.

To generalize, the performance of NHM to simulate the extreme precipitation, more sensitivity experiments (e.g., changing domain size and boundary schemes) and many cases are necessary be considered.

Chapter 5

General conclusion

Prediction of the precipitation over the Central Himalaya in and around Nepal either at short to medium range or at climate scale, is crucial from multiple perspectives such as for managing weather and climate induced hazards, planning agriculture and water resource, and knowing its thermodynamic effect in modulating the Asian monsoon system. Recently revealed evidences such as increasing extreme precipitation events causing floods and droughts change of precipitation trend and variability, delay monsoon onset and shift of ecosystem boundary to upper altitude fueled the emergency on the study of precipitation. In addition, dependency of Himalayan people directly and indirectly on precipitation for their daily wellbeing and their poor technical as well as economic capability to handle the adverse conditions exacerbated by climate change make them more vulnerable to weather/climate hazards. Furthermore, precipitation over the Himalaya not only threatens the people over the Himalaya but also the millions of people living in the downstream in India and Bangladesh through the availability of the discharge in the river originating from the Himalaya. In this regard, evaluations of performances of the state–of–the–art numerical models are very essential and demanding. In this study, we have investigated the performances of the state-of-the-art global climate models (GCMs) and global operational ensemble prediction models (EPSs) to reproduce the precipitation in summer monsoon season (SMS) over the central Himalaya. Additionally, we have examined the skill of a regional climate model to reproduce an extreme precipitation event.

First, the capabilities of 38 GCMs participated in Coupled Model Intercomparison Project phase 5 (CMIP5) to reproduce precipitation at present climate are evaluated systematically. Results revealed that majority of the models are capable to simulate annual cycles of mean precipitation over the central Himalaya to some degree. However, simulation of the spatial pattern of SMS mean climatology and interannual variability (IAV) is remained a major challenge to the models. About 66 % of the models show wet bias. Out of 38 models, only six GCMs viz., ACCESS1.0, CNRM-CM5, EC-ERATH, HadGEM2-AO, HadGEM2-CC, and HadGEM2-ES are found relatively better than others. To maintain the model spread, only HadGEM2-ES is selected from HadGEM2 family as this model showed skills either similar or better than others. Using the selected best four models and their multi–model mean, the assessment of the future change under two warming scenarios namely representative concentration pathway (RCP) 4.5 and RCP 8.5 are performed for three future tri-decadal periods (near: 2011–2040, mid: 2041–2070 and far: 2070–2099). Best models and multi-model mean project the increase of seasonal mean precipitation linearly from near to far future periods with increasing amplitude and confidence. However, the projected change in variability is very uncertain. Furthermore, moderate and heavy precipitation days are projected to increase whereas dry days are projected to decrease. However, regional uncertainty does exit. Additionally, the frequency and length of active monsoon spells and frequency of dry spells are projected to increase consistently in future periods. The increase in the seasonal mean and the extremes could be related to the increase in low level flow from the Arabian Sea, increase in humidity and convective activities. Our results suggest that the adverse impact of extreme events might be increased in future making mountain communities more vulnerable if appropriate adaptation measure is not implemented.

Secondly, we have documented the daily precipitation forecast skills of the five major global operational EPSs (CMA, CMC, ECMWF, NCEP and UKMO) which contributed in The Observing System Research and Predictability Experiment (THORPEX) Interactive Grand Global Ensemble (TIGGE). The results show that the performances of EPSs to forecast precipitation in Nepal, a mountainous country, are poor. Comparatively, the ECMWF exhibits the highest forecast skills for both quantitative precipitation forecast (QPFs) and probabilistic QPFs (PQPFs) followed by UKMO. Although, CMC has shown comparatively lower error, better spread-skill relationship and reliability diagram, forecasts are least sharp and the spatial correlation and discrimination ability is very poor. The skill scores to verify QPFs and PQPFs show that NCEP is the least reliable. Similarly, CMA shows large wet bias particularly for the first three lead days, and over estimates medium and heavy rain events. Furthermore, we investigate the performance of multi–model grand ensemble means by applying various methods in 2012. The results show that forecast errors are reduced in grand ensemble means than the individual EPS, with the highest improvement in bias removed grand ensemble. However, the performance of ECMWF is as good as the simple multi–model grand ensemble mean.

Finally, we have reported the reproducibility of an extreme event occurred on 13–15 August, 2014 by Japan Meteorological Agency Non–hydrostatic Model (JMA–NHM). Four experiments are performed with two sets of domain sizes (big and small) at 25km and 5km spatial resolutions. Results show that JMA–NHM is able to simulate extreme precipitation at 25km to some extent. However, the spatial coverage of the event is greatly diminished, and excessive wet bias is consistent in the eastern parts of Nepal. The large wet bias could be rated to the topography, cumulus parameterization scheme and bias inherited from initial and boundary conditions. With the use of high resolution (5km) model, the simulation of precipitation is not improved, and even deteriorated in the case of small domain on 14 and 15 August. Additionally, the impact of domain size is insignificant in the case of 25km experiments compared to the systematic wet bias particularly in the eastern parts of Nepal. However, domain size has shown large difference in simulation in the case of 5km experiments.

In summary, precipitation over the central Himalaya is highly variable both in space and time. The state–of–the–art numerical models are not sufficient to accurately reproduce the precipitation characteristics. Therefore, detail precipitation mechanisms should be understood, and numerical models need to be improved accordingly. Additionally, in this study, we have evaluated the global models (climate and operational) based on their skills to simulate precipitation only over the central Himalaya and Nepal. To describe the models' performances fully, thorough analysis of other simulated variables such as wind, temperature and moisture is also necessary.

Appendix

Appendix A

A.1 Spectral analysis to determine a dominant period

As mentioned in section 2.2.4.2 of chapter 2, rainfall index (RI) used to define active and break spells is computed by applying 6–90–day band pass filter (BPF) to daily rainfall data. The choice of 6–90–day period is made based on the spectrum analysis (Fig. A1) of detrended daily anomaly of area averaged precipitation over the central Himalaya (26–30° N and 80–89° E) for 30 years (1971–2000), and variance explained by the respective period (Fig. A2) *(Fujinami et al., 2011)*. The detrended anomaly is created by removing the climatological mean of first three harmonics (periods of 365.25, 182.625, and 121.75 days) of the annual cycle for 1971–2000. The spectrum analysis shows that the intraseasonal variability of precipitation over the central Himalaya is modulated by multiple dominant periods, viz. 3–5–day, 6–20–day and 21–90–day (Fig. A1). The variance explained by each dominant period is shown in Fig. A2. Comparatively, 6–20–day period explains the highest variance followed by 21–90–day and 3-5–day periods in order. However, variance of 21–90–day spectrum in some years is large or equal to variance contributed by 6–20–day spectrum. Therefore, we have selected 6–90–day band period which is significant at 95% confidence level. To calculate the significance levels of spectra, we refer to *Wilks (2006)* method that utilizes a theoretical spectral density function $S_0(f_k)$ of lag-1 autocorrelation properties AR(1). Adopting Equation 8.77 in *Wilks (2006)*, $S_0(f_k)$ is defined as

$$S_0(f_k) = \frac{4\sigma_\varepsilon^2/n}{1 + \phi^2 - 2\phi\cos(2\pi f_k)}, \quad 1 \le k \le n/2$$

where ϕ is the is the lag-1 autocorrelation value; f_k is the frequency at wavenumber k; σ_ε^2 and n are the data variance and number of data samples, respectively.

Then, the spectrum limit for the significance level at frequency f_k can be formulated as (Equation 8.81 in *Wilks (2006)*)

$$Sig_k \equiv \frac{S_0(f_k)}{\nu}\chi_\nu^2(1-\alpha),$$

where, $\chi_\nu^2(1-\alpha)$ is the critical value denoting right-tail quantiles of the appropriate Chi-square distribution. In this study, we have used the 95% and 99% confidence levels which correspond to the critical values of 5.991 and 9.211, respectively; ν is the number of degrees of freedom and is equal to 2. BPF is applied to each monsoon season using Lanczos filtering method with 180 daily weights at each end.

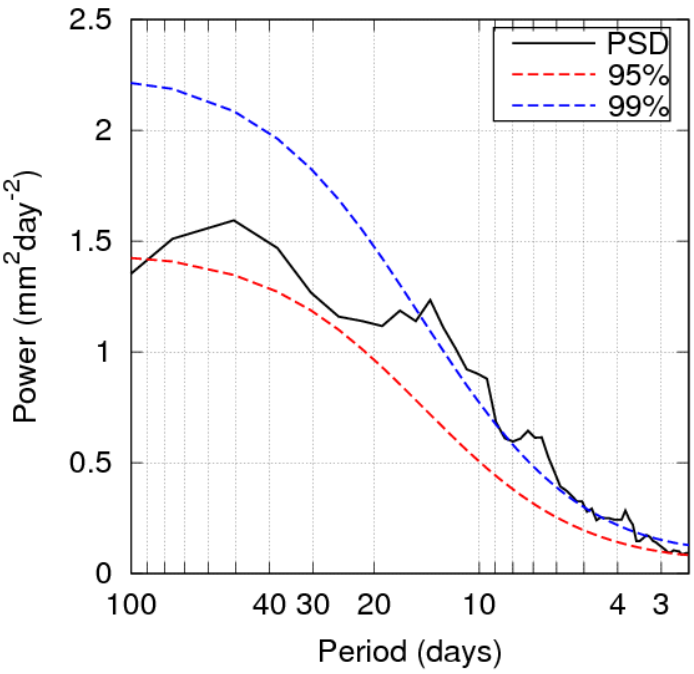


Figure A1. Ensemble mean of power spectra of daily rainfall index (JJAS) from 1971 to 2000 calculated using Discrete Fourier Transform (DFT) method (black line). Seasonality is removed before the calculation by removing the climatological mean of first three harmonics. Dashed lines denote confidence levels of 95% (red) and 99% (blue.

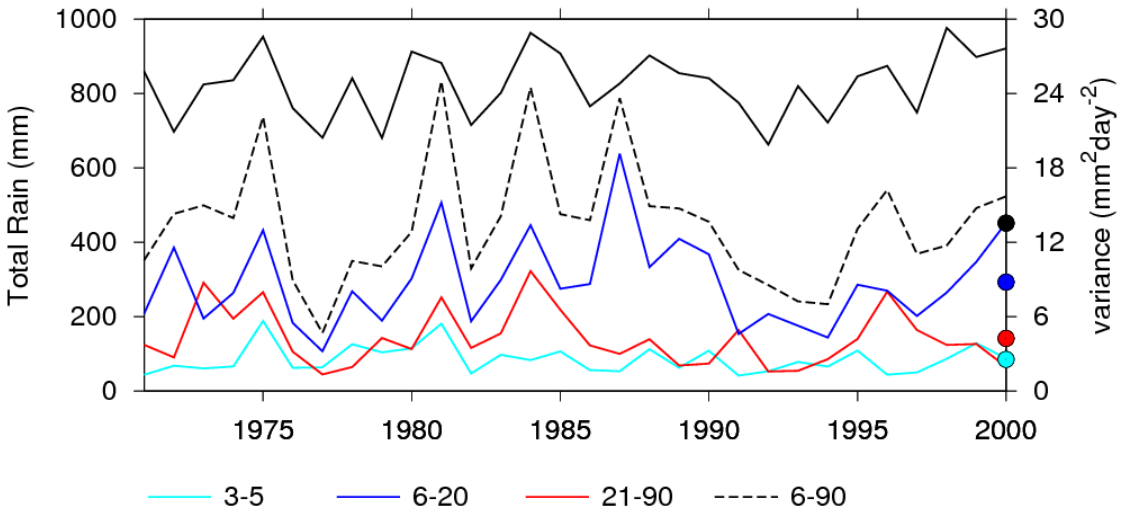


Figure A2. Interannual variation of total summer rainfall (solid black line, left axis) and daily variance (right axis) of 3-5 day (light blue line), 6-20 day (blue line), 21-90 day (red line), and 6-90 day (dashed line) band periods from 1971 to 2000. Mean variances are shown as circles in the right axis.

Appendix B

B.1 Summary of TIGGE models

THORPEX Interactive Grand Global Ensemble (TIGGE) was established in 2005 as a major component of The Observing System Research and Predictability Experiment (THORPEX), a World Meteorological Organization's (WMO) program with the aim to accelerate the improvements in the accuracy of 1–day to 2–week high–impact weather forecasts (*Bougeault et al. 2010).* TIGGE integrates the forecasts from ten major global operational forecast centers throughout the world namely, the Bureau of Meteorology of Australia (BoM), China Meteorological Administration (CMA), Canadian Meteorological Centre (CMC), the Centro de Previsão de Tempo e Estudios Climáticos of Brazil (CPTEC), European Centre for Medium-Range Weather Forecasts (ECMWF), the Japan Meteorological Agency (JMA), the Korea Meteorological Administration (KMA), the National Meteorological Service of France (Meteo-France), U.S. National Centers for Environmental Prediction (NCEP), and United Kingdom Meteorological Office (UKMO). TIGGE data are available quasi-real time (about a delay of 48 h) through the ECMWF data portals. In his study, we have selected five models: CMA, CMC, ECMWF, NCEP and UKMO, which differ from each other in many respects such as model resolution, initial perturbation methods, forecast lengths and ensemble members. These centers are selected mainly based on their data availability during our analysis period. However, CPTEC is not included in this study because of the poor performance investigated in our preliminary analysis. TIGGE models are upgraded regularly. The detail information about the models is available online (https://software.ecmwf.int/wiki/display/TIGGE).

The CMA uses bred vectors (BVs) for the T213 global model (~0.5625°) *(Wang et al. 2008)* as the initial perturbations to construct the EPS, and model uncertainties have not been taken into account. CMA EPS has not been upgraded during our study period *(Su et al. 2014).*

The CMC EPS uses ensemble Kalman filter (EnKF) to generate initial perturbations *(Charron et al. 2010).* Multiphysics schemes (such as different deep convections, surface schemes, mixing lengths, vertical diffusions, and gravity wave drags) and two stochastic parameterization schemes (perturbations of physics tendencies (PTP) and stochastic kinetic energy backscatter (SKEB)) are used to account model uncertainties *(Charron et al. 2010; Su et al. 2014).* On 17 August 2011, the CMC EPS has been upgraded with the finer model horizontal grid spacing of 66 km changing from about 100 km with the change of the model grid from a latitude-longitude grid to a Gaussian grid *(Su et al. 2014).*

The ECMWF EPS has been upgraded many times during the study periods. The initial-time singular vectors (EVO-SVINI) method was adopted as its initial perturbations until 24 June 2010, and after this, it has been upgraded to the ensemble of data assimilation and the initial-time singular vectors (EDA-SVINI) *(Buizza et al. 2010).* The stochastic perturbation of physics tendency (SPPT) has been applied to account for model uncertainties *(Su et al. 2014).* More detail information can be accessed online (https://software.ecmwf.int/wiki/display/TIGGE).

The NCEP EPS uses the bred vector-ensemble transform with rescaling (BV-ETR) to generate initial perturbations. Since 23 February 2010, the stochastic total tendency perturbation (STTP) scheme has been introduced into the NCEP EPS to account for model uncertainties *(Cui et al. 2012)* with the increase in model horizontal resolution from T126 (~110 km) to T190 (~70 km) *(Su et al. 2014).* It has changed model version to

9.0 with the adoption of the improved BV-ETR initialization and STTP schemes, and the upgraded horizontal resolution of T254 (~55 km) for 1–8 day forecasts (9–16 day forecasts remain T190) on 14 February 2012 *(Su et al. 2014).*

The UKMO EPS uses the ensemble transform Kalman filter (ETKF) *(Bowler et al. 2008)* as the initial perturbation strategy. Random parameters and stochastic kinetic energy backscatter schemes have been adopted to represent model uncertainties. The version of the UKMO EPS has been changed several times during our analysis periods. It has been upgraded to version 8.0 with the improvement in horizontal resolution from 1.25° × 0.83° to 0.83° × 0.56° on 9 March 2010.

B.2 Computation of single model ensemble mean and multi-model grand ensembles

B.2.1 Ensemble mean of ensemble members of each models

Ensemble mean (EnM) is the average of all members at the given time, and is defined by

$$\mathrm{EnM} = \frac{1}{\mathrm{M}} \sum_{\mathrm{im}=1}^{\mathrm{M}} \mathrm{F}_{\mathrm{im}} \quad \text{……………………………………………….. (B1)}$$

where, M is the number of ensemble members, F_{im} represents the forecast of the im^{th} member at a time.

B.2.2 Multi-model grand ensembles

a) Simple grand ensemble of member models

Simple multi-model grand ensemble mean (MMM) of all five model is the simple mean of all five member models ensemble mean forecasts, and is defined as

$$\mathrm{MMM} = \frac{1}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{EnM_i} \quad \text{………………………………...………………. (B2)}$$

where, N is the total number of member models (in this study, N=5), EnM_i represents the ensemble mean of the i^{th} member model at a time (i = 1,2,3,…,N).

b) Simple grand ensemble of bias removed member models (MMM-BR)

Best easy systematic bias (BES) of ensemble mean forecast of each member model at each grid is first computed for past three consecutive seasons (June–September, 2009-2011) using the following formula *(Durai and Bhardwaj, 2014).*

$$\mathrm{BES}\ (\mathrm{b_i}) = \frac{1}{4}\ (\mathrm{Q_1} + 2\mathrm{Q_2} + \mathrm{Q_3}) \quad \text{…………….………………………… (B3)}$$

where, Q_1, Q_2 and Q_3 are the 1st, 2nd and 3rd quartile of forecast error (EnM_i – O) respectively and O is observation.

Then, multi-model grand ensemble of bias removed forecasts (MMM–BR) is computed as

$$\mathrm{MMM} - \mathrm{BR} = \frac{1}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} (\mathrm{EnM_i} - \mathrm{b_i}) \quad \text{……………………………….…. (B4)}$$

c) Grand ensemble with weights from RMSE (MMM-WR)

Root mean square error (RMSE) of ensemble mean forecast of each member model is computed for the past three consecutive seasons (June–September, 2009-2011) at each grid point using following formula

$$\mathrm{RMSE_i} = \sqrt{\frac{1}{\mathrm{T}}\sum_{\mathrm{it}=1}^{\mathrm{T}}(\mathrm{EnM_{it}} - \mathrm{O_{it}})^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(B5)}$$

where, it is forecast day and T is total number of forecast days (366).

Then, weight of each model is computed as

$$\mathrm{weight\ (we_i)} = \frac{\mathrm{a_i}}{\sum_{\mathrm{i}=1}^{\mathrm{N}}\mathrm{a_i}},\ \mathrm{a_i} = \frac{1}{\mathrm{RMSE_i}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(B6)}$$

Finally, multi-model grand ensemble forecast with weight (MMM-WR) is created as follows.

$$\mathrm{MMM-WR} = \sum_{\mathrm{i}=1}^{\mathrm{N}}\mathrm{we_i} \times \mathrm{EnM_i} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(B7)}$$

d) Grand ensemble with weights from correlation coefficient (MMM-WC)

Multi-model grand ensemble with weights from correlation coeffificent (MMM-WC) (Roy Bhowmik and Durai, 2010) is calculated as follows.

$$\mathrm{MMM-WC} = \sum_{\mathrm{i}=1}^{\mathrm{N}}\mathrm{wc_i} \times \mathrm{EnM_i} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(B8)}$$

where, $\mathrm{wc_i}$, weight from correlation coefficient (CC) of ith member model, is define as

$$\mathrm{weight\ (wc_i)} = \frac{\mathrm{CC_i}}{\sum_{\mathrm{i}=1}^{\mathrm{N}}\mathrm{CC_i}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(B9)}$$

and $\mathrm{CC_i}$, correlation coefficient between forecasted and observed precipitation, is computed at each grid as

$$\mathrm{CC_i} = \frac{\sum_{\mathrm{it}=1}^{\mathrm{T}}(\mathrm{EnM_{it}} - \overline{\mathrm{EnM}})(\mathrm{O_{it}} - \bar{\mathrm{O}})}{\sqrt{\sum_{\mathrm{it}=1}^{\mathrm{T}}\mathrm{EnM_{it}} - \overline{\mathrm{EnM}}^2}\sqrt{\sum_{\mathrm{it}=1}^{\mathrm{T}}(\mathrm{O_{it}} - \bar{\mathrm{O}})^2}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(B10)}$$

For the computational consistency, $\mathrm{CC_i}$ is taken as 0.0001 in case $\mathrm{CC_i} \leq 0$.

B.3 Verification Indices

B.3.1 Deterministic verification

a) Root mean square error

Root Mean Square Error (RMSE) is one of the most widely used measures in the verification of QPFs which represents the accuracy of forecasts, and is defined by

$$RMSE = \frac{1}{T}\sum_{it=1}^{T}\sqrt{\frac{1}{G}\sum_{ig=1}^{G}(EnM_{ig,it} - O_{ig,it})^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(B11)}$$

where, G is the total number of grids, T is total analysis days, O is observation.

b) Pattern correlation coefficient (CC)

Pattern correlation coefficient during analysis period is computed as

$$CC = \frac{1}{T}\sum_{it=1}^{T} CC_{it} \quad \text{..... (B12)}$$

where,

$$CC_{it} = \frac{\sum_{ig=1}^{G}(EnM_{ig} - \overline{EnM})(O_{ig} - \bar{O})}{\sqrt{\sum_{ig=1}^{G} EnM_{ig} - \overline{EnM}^2}\sqrt{\sum_{ig=1}^{G}(O_{ig} - \bar{O})^2}} \quad \text{..... (B13)}$$

c) Ensemble spread

Ensemble Spread (STD) is a widely used to measure the degree of the forecast uncertainty in the ensemble forecast. It is the standard deviation of the ensembles, and is defined by

$$\mathrm{STD} = \sqrt{\frac{1}{\mathrm{M}-1}\sum_{\mathrm{m}}^{\mathrm{M}}(\mathrm{F_m} - \mathrm{EnM})^2} \quad \text{..... (B14)}$$

Average ensemble spread is calculated over the study area for the evaluation period using following equation

$$\text{average ensemble spread} = \frac{1}{\mathrm{T}}\sum_{\mathrm{it}=1}^{\mathrm{T}}\frac{1}{\mathrm{G}}\sum_{\mathrm{ig}=1}^{\mathrm{G}}\mathrm{STD}_{\mathrm{ig,it}} \quad \text{..... (B15)}$$

where, G is number of grids (293) and T (488) is number of forecast days used in the study.

d) Verification indices for categorical forecasts

The contingency table is very useful way to see what types of errors are being made. In the contingency table, the categorical forecasts of a binary event are divided into four possible outcomes, namely, hits, false alarms, misses and correct rejections (or correct negatives) (Table B1). A perfect forecast system would produce only hits and correct negatives, and no misses or false alarms. A forecast (>= a threshold) is considered as a hit if is occurred, and false alarm if is not occurred. Similarly, a forecast is considered as a miss if not forecasted (>= a threshold) event is occurred, and correct rejection if there is no forecast and no observation for the threshold.

Table B1. Contingency table

		Observations		Total
		Yes	No	
Forecasts	Yes	Hits (a)	False alarms (b)	a+b
	No	Misses (c)	Correct rejections (d)	c+d
	Total	a+c	b+d	N = a+b+c+d

A large variety of categorical skill score can be computed from the elements in the contingency table to describe particular aspects of forecast performance. In this study, only four scores namely frequency bias (bias), probability of detection (POD), false alarm ratio (FAR) and equitable threat score (ETS) are computed.

i) Bias

Frequency bias (BIAS) is the ratio of forecasted events to the frequency of observed events. It answers whether the forecast system has a tendency to underforecast (BIAS<1) or overforecast (BIAS>1) events, and is defined as

$$BIAS = \frac{a+b}{a+c} \quad \text{(B16)}$$

The value of BIAS varies 0 to ∞. The value of unity indicates that the forecast system is perfect.

ii) POD

POD indicates what fraction of the observed "yes" events is correctly forecasted, and is defined as

$$POD = \frac{a}{a+c} \quad \text{(B17)}$$

POD can be artificially improved by issuing more "yes" forecasts if the forecasts system has tendency of wet bias. The value of POD ranges between 0 and 1. Higher the value of POD better the forecast system is.

iii) FAR

FAR indicates what fraction of the predicted "yes" events actually did not occur, and is defined as

$$FAR = \frac{b}{a+b} \quad \text{(B18)}$$

The value of FAR ranges between 0 and 1. The perfect value is zero.

iv) ETS

ETS indicates how well does the forecast "yes" events correspond to the observed "yes" events accounting for hits due to random chance, and is defined as

$$ETS = \frac{a-a_r}{a+b+c-a_r} \quad \text{(B19)}$$

where,

$$a_r = \frac{(a+c)(a+b)}{N} \quad \text{(B20)}$$

The value of ETS ranges between -1/3 to 1. Zero value indicates that forecast system shows no skill, and value of unity indicates the forecast system is perfect.

B.3.2 Probabilistic verification

a) Brier score (BS)

A BS is a way to verify the accuracy of a probability forecast, and is defined by

$$BS = \frac{1}{T}\sum_{it=1}^{T}(p_{it} - O_{it})^2 \quad \text{(B21)}$$

where, T is the evaluation period in days. O_{it} is the actual outcome of the observed event. The value of O_{it} is equal to 1 (0) for the occurrence (nonoccurrence) of the event. p_{it} is the forecast probability and is defined by

$$p_{it} = \frac{1}{M}\sum_{im=1}^{M} F_{im} \quad \text{(B22)}$$

where, M is the total number of ensemble members, F_{im} represents the forecast of the im^{th} member and is equal to 1(0) for the occurrence (nonoccurrence) of the event of the given thresholds.

The value of BS ranges between 0 and 1. BS=0 means that the forecasts are perfect, and BS=1 means that forecasts are not accurate**.**

The BS can be decomposed into three components: reliability, resolution and uncertainty (i.e. BS = Reliability - Resolution + Uncertainty) *(Murphy 1973)*. The reliability measures the closeness of the forecast probability with the observed frequency whereas the resolution measures the difference of the forecast probability with the climatological probability of the event. Consider the probability of the forecast is classified into K intervals. Let the sample number in the ikth interval be n_{ik}, and the number of observed events in n_{ik} be m_{ik}. Then,

$$Relability = \sum_{ik=1}^{K}\frac{n_{ik}}{N}\left(\frac{m_{ik}}{n_{ik}} - P_{ik}\right)^2 \quad \text{(B23)}$$

$$Resolution = \sum_{ik=1}^{K}\frac{n_{ik}}{N}\left(\frac{m_{ik}}{n_{ik}} - \frac{M}{N}\right)^2 \quad \text{(B24)}$$

where, P_{ik} is the representative value in the ik^{th} interval of the predicted probability, $N = \sum_{ik}^{K} n_{ik}$ and $M = \sum_{ik}^{K} m_{ik}$.

b) Brier skill score (BSS)

BSS tells the relative skill of the probabilistic forecast over that of climatology, in terms of predicting whether or not an event occurred, and is defined as

$$BSS = 1 - \frac{BS}{BS_{ref}} \quad \text{(B25)}$$

where, BS_{ref}, reference brier score, is defines as

$$BS_{ref} = \frac{1}{T}\sum_{it=1}^{T}(p_c - O_{it})^2 \quad \text{(B26)}$$

where, p_c is probability of climate and is defined by

$$p_c = \frac{1}{T}\sum_{it=1}^{T} O_{it} \quad \text{(B27)}$$

A BSS has a range of -∞ to 1. Negative values mean that the forecast is less accurate than a standard forecast, and 0 values indicates no skill and one indicates perfect skill compared to the reference forecast.

B.4 Figures

B.4.1 Verification of QPFs

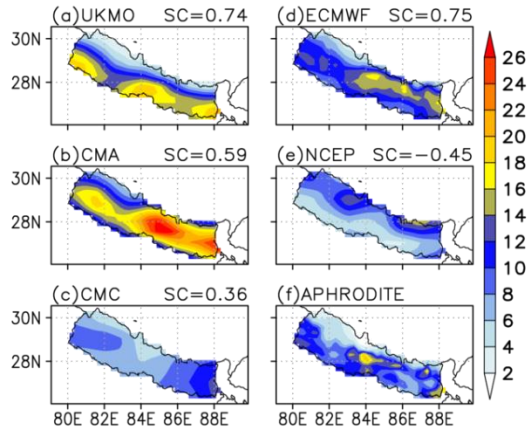


Figure B.4.1.A1. Average precipitation (mm day^{-1}) of ensemble mean forecasts for day+1 from five EPSs and APHRODITE during JJAS, 2009–2012. The pattern correlation coefficient (SC) is shown in the title.

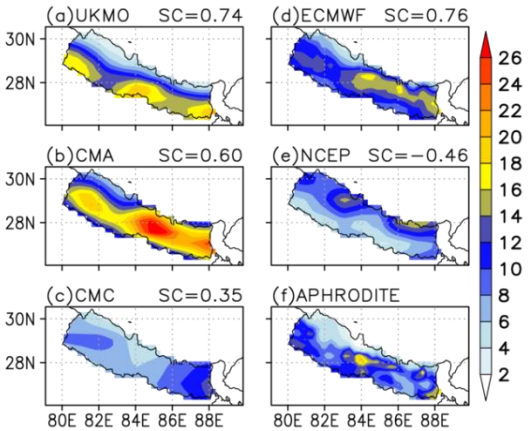


Figure B.4.1.A2. Same as Figure B.4.1.A1 but for day+2 .

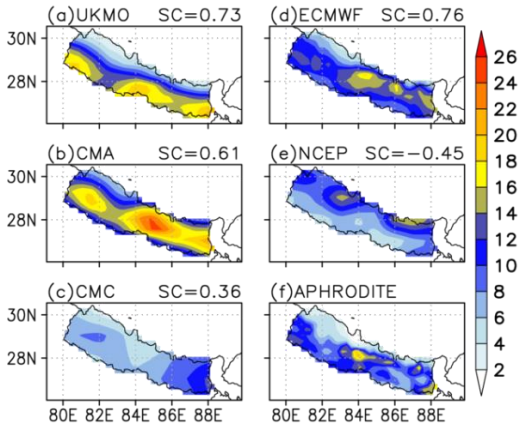


Figure B.4.1.A3. Same as Figure B.4.1.A1 but for day+3.

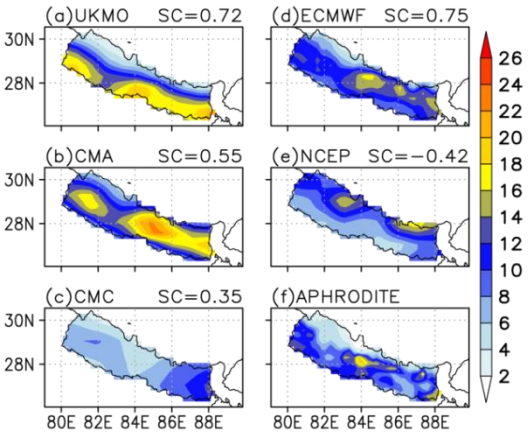


Figure B.4.1.A4. Same as Figure B.4.1.A1 but for day+4.

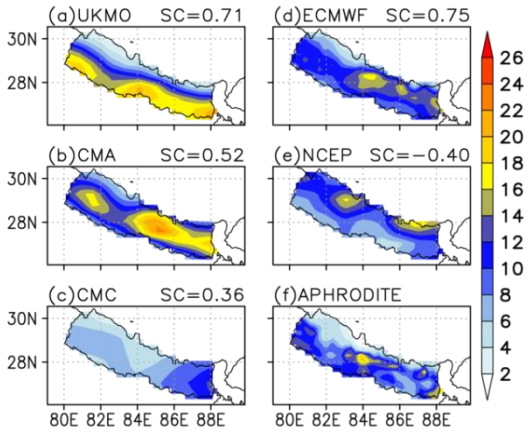


Figure B.4.1.A5. Same as Figure B.4.1.A1 but for day+5.

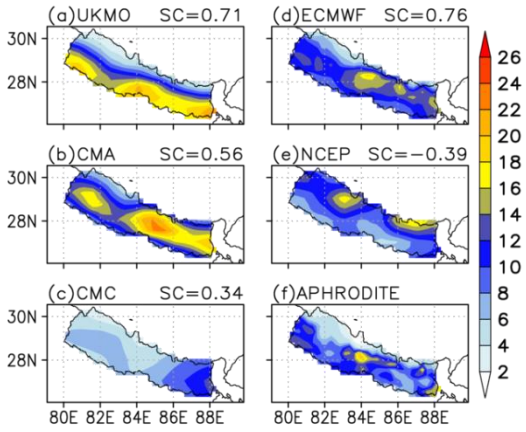


Figure B.4.1.A6. Same as Figure B.4.1.A1 but for day+6.

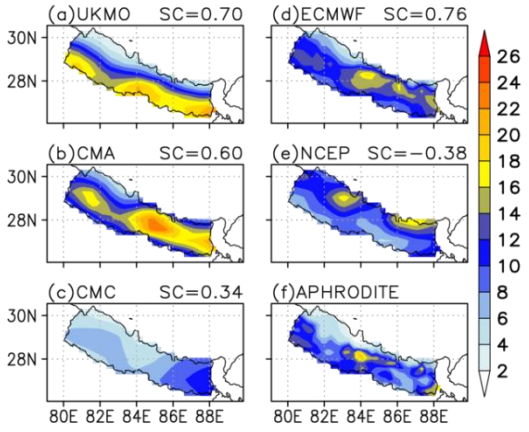
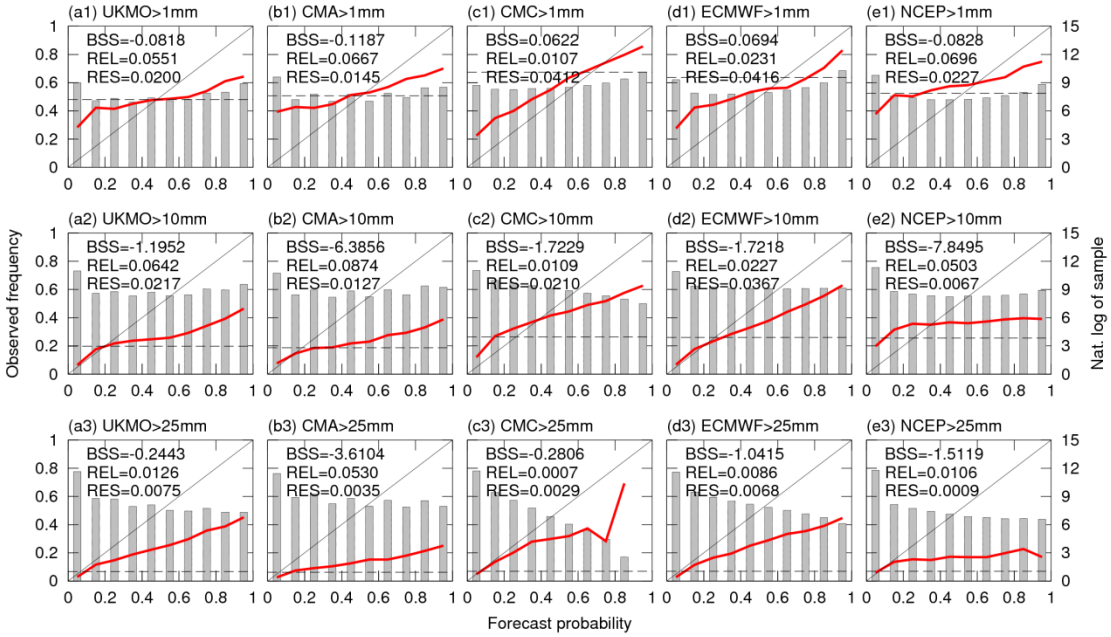
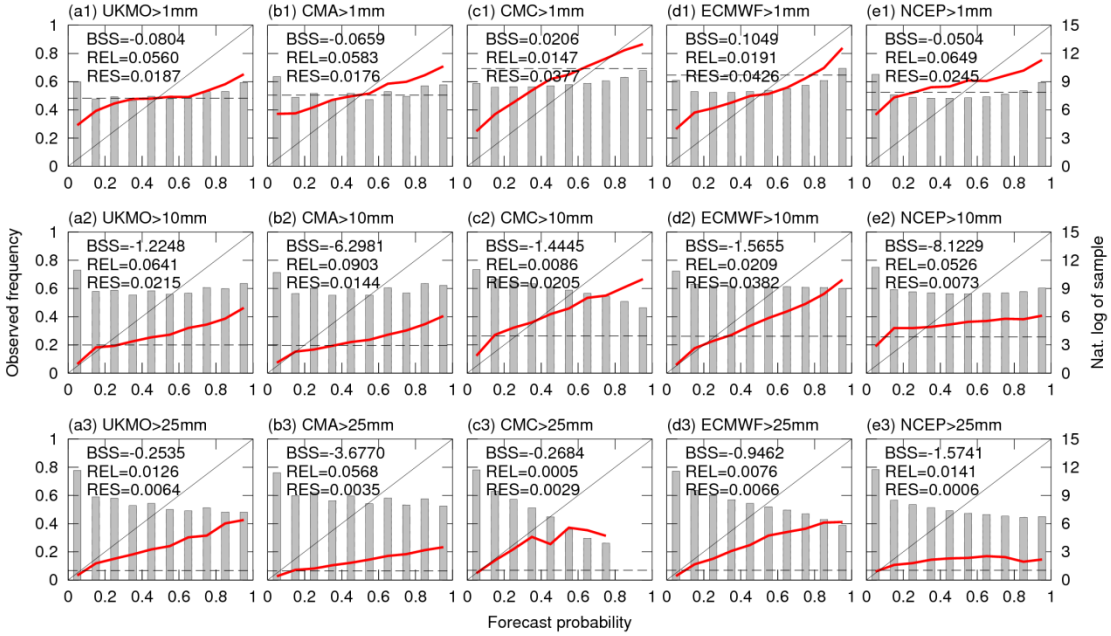


Figure B.4.1.A7. Same as Figure B.4.1.A1 but for day+7.

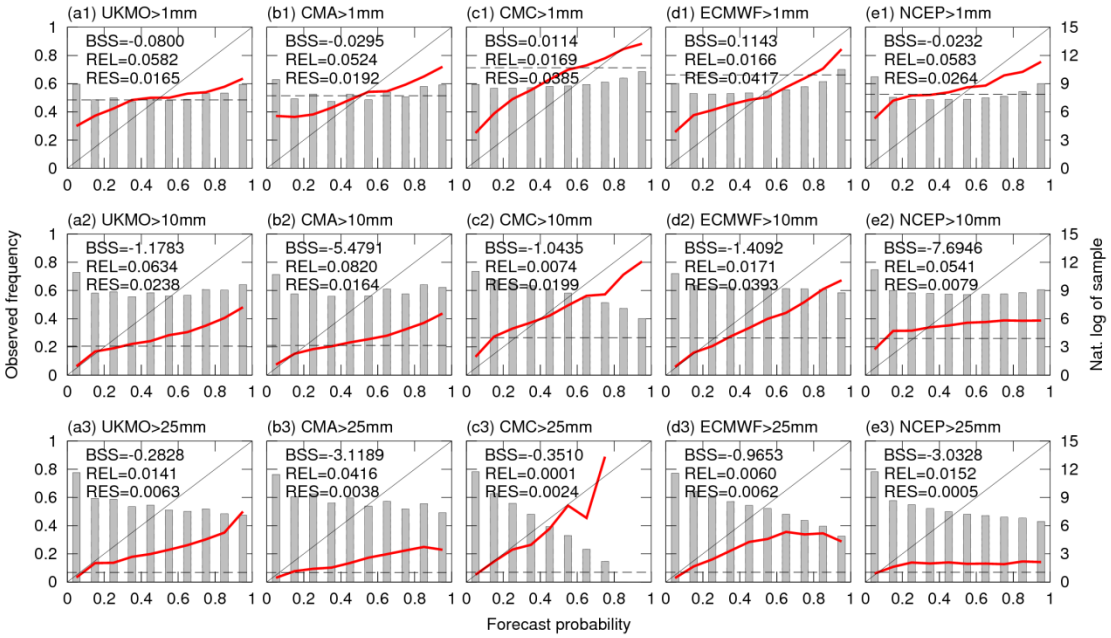
B.4.2 Verification of PQPFs



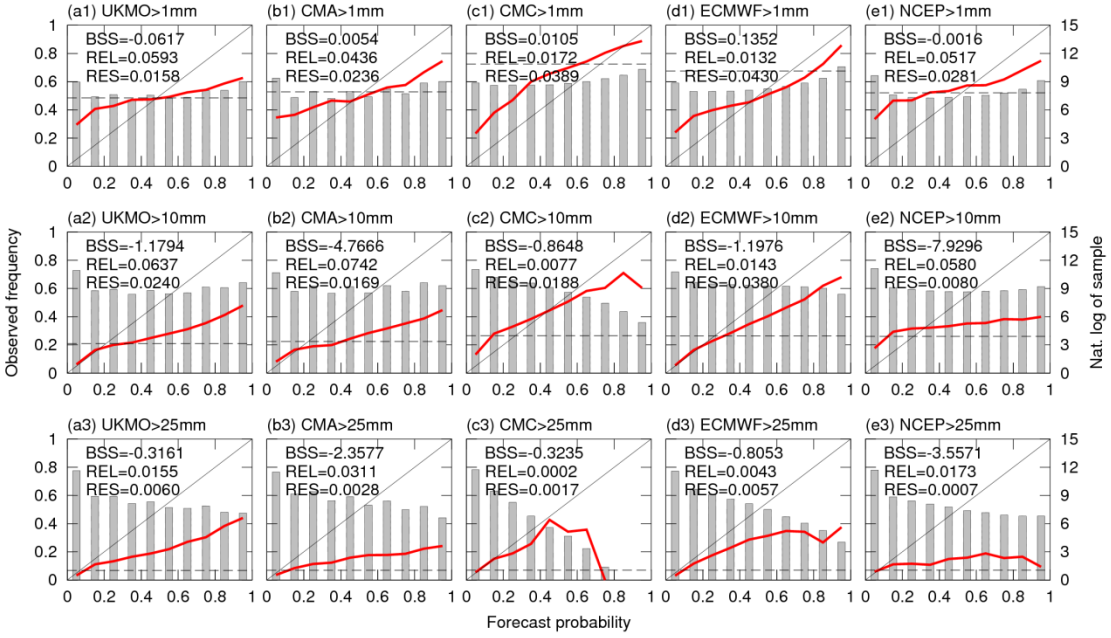
B.4.2.A1. Reliability diagrams for day+1 at the different thresholds (1, 20 and 25 mm day^{-1}). The bar graphs show the sample frequencies at the natural logarithmic scale (right side). The horizontal line (dotted, black) represents the observed sample frequency. The BSS, the reliability (REL) and resolution (RES) are shown as numbers. For brevity, probabilities of all EPSs with different member sizes are categorized into 10 bins.



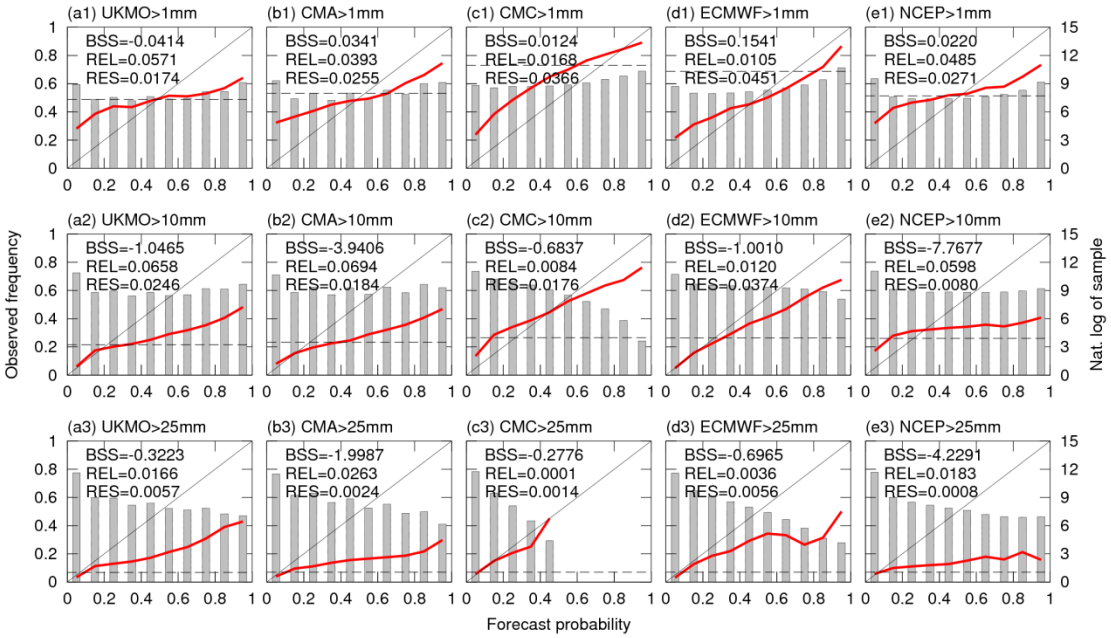
B.4.2.A2. Same as figure B.4.2.A1 but for day+2.



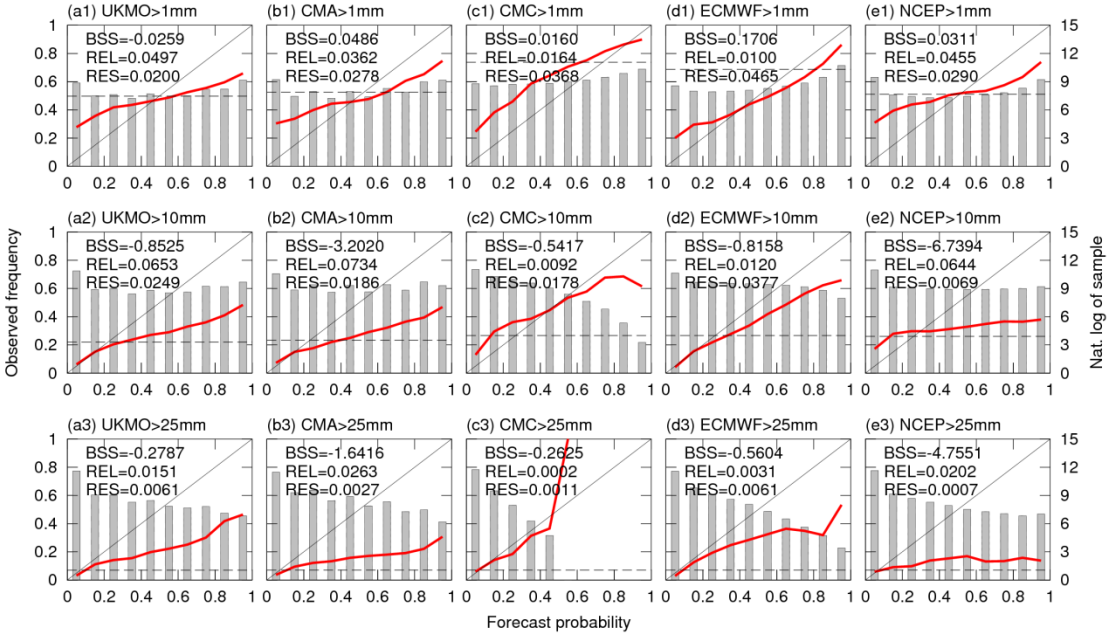
B.4.2.A3. Same as figure B.4.2.A1 but for day+3.



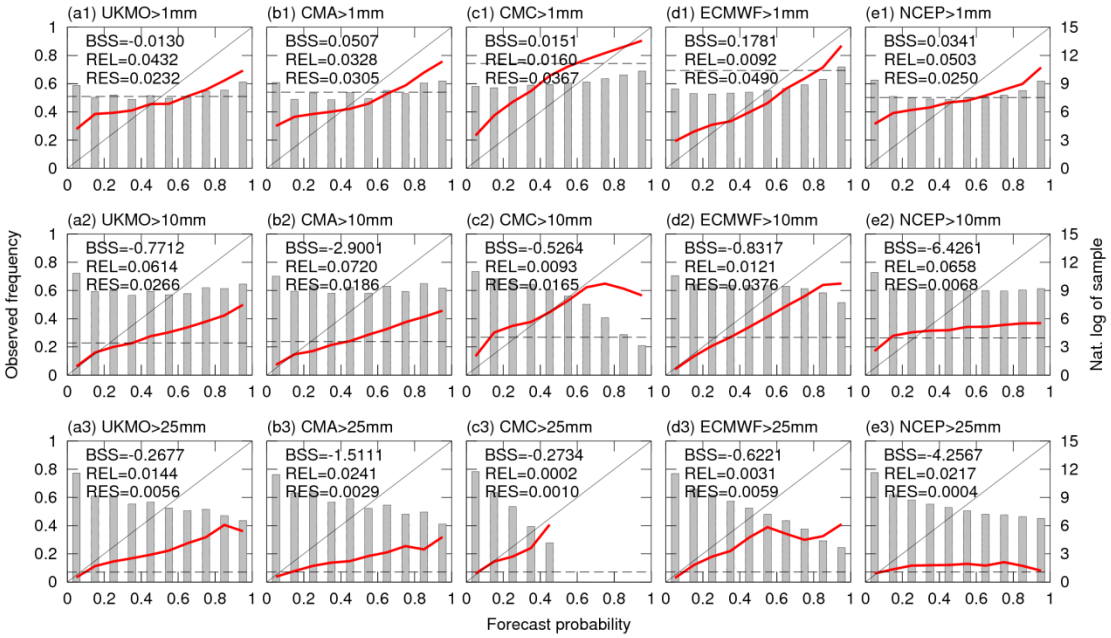
B.4.2.A4. Same as figure B.4.2.A1 but for day+4.



B.4.2.A5. Same as figure B.4.2.A1 but for day+5.



B.4.2.A6. Same as figure B.4.2.A1 but for day+6.



B.4.2.A7. Same as figure B.4.2.A1 but for day+7.

B.4.3 Verification of multi-model grand ensembles

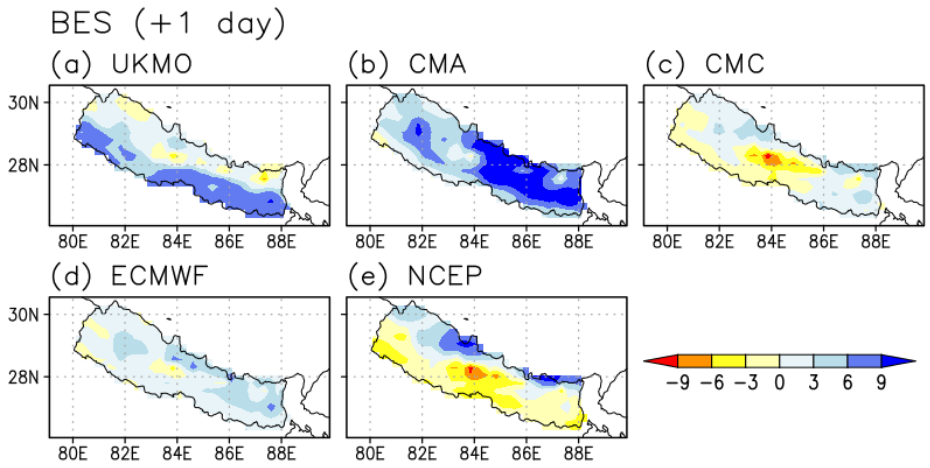


Figure B.4.3.1.A1. Best easy systematic bias (BES) during training period (JJAS, 2009-2011) for day+1.

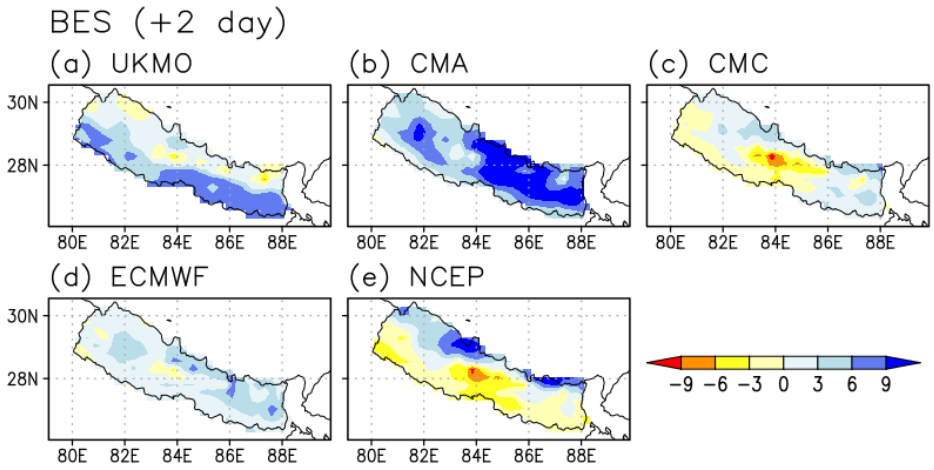


Figure B.4.3.1.A2. Same as figure B.4.3.1.A1 but for day+2.

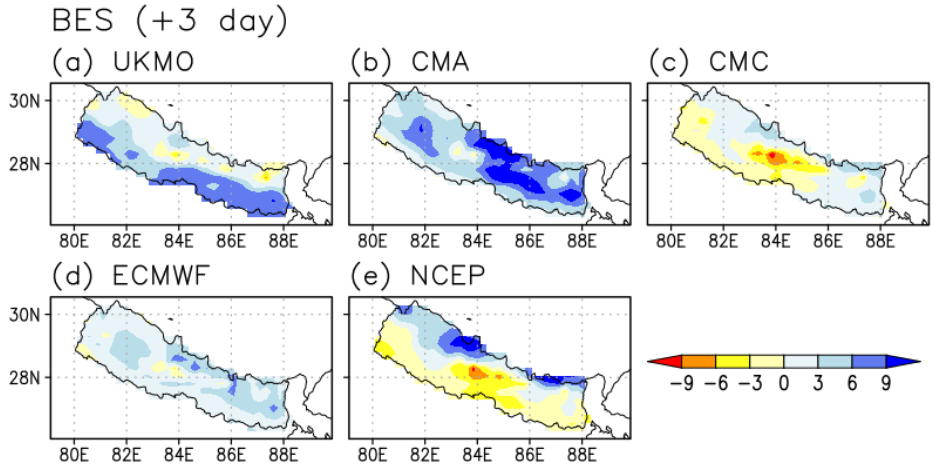


Figure B.4.3.1.A3. Same as figure B.4.3.1.A1 but for day+3.

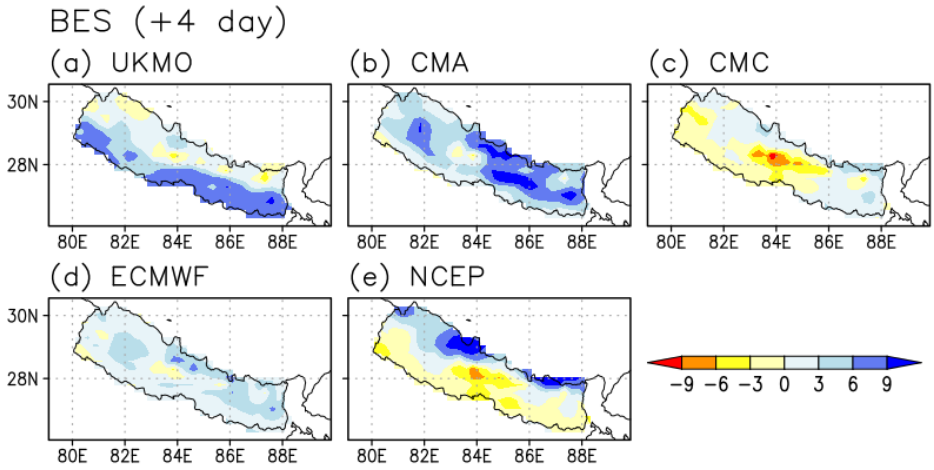


Figure B.4.3.1.A4. Same as figure B.4.3.1.A1 but for day+4.

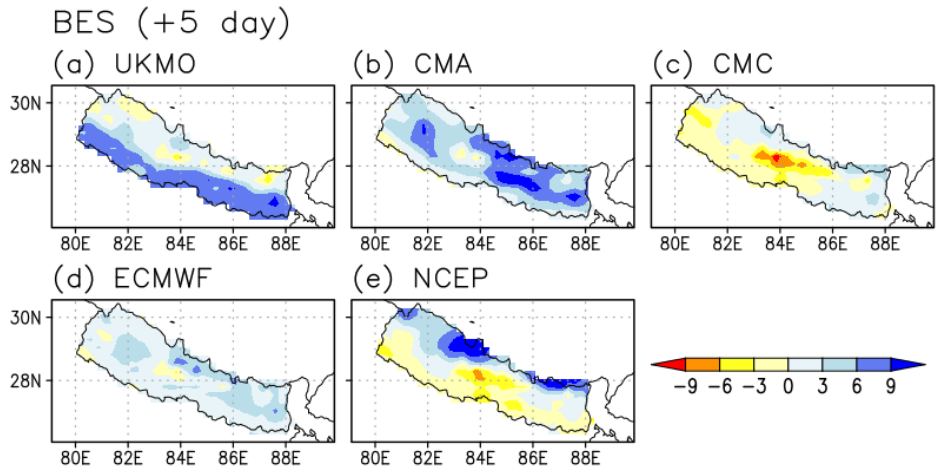


Figure B.4.3.1.A5. Same as figure B.4.3.1.A1 but for day+5.

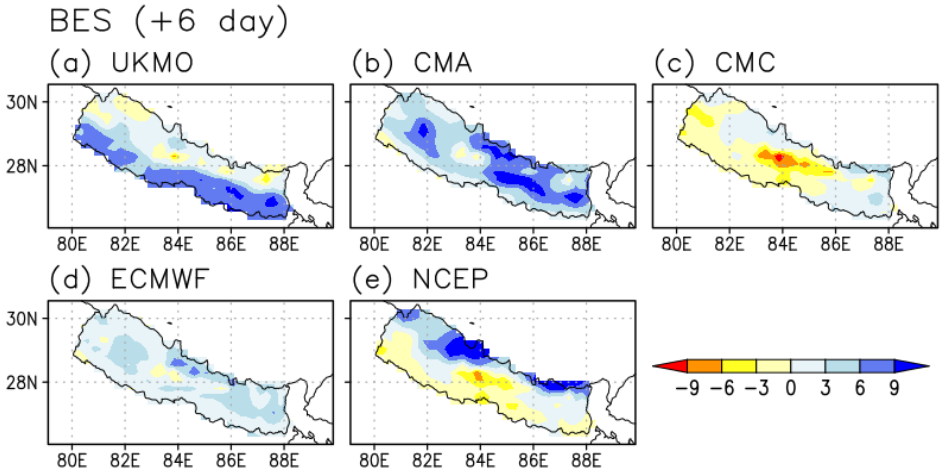


Figure B.4.3.1.A6. Same as figure B.4.3.1.A1 but for day+6.

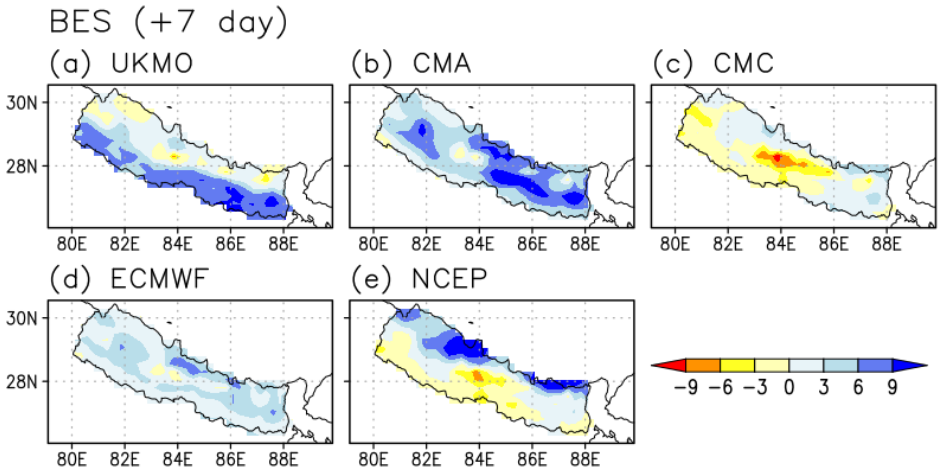


Figure B.4.3.1.A7. Same as figure B.4.3.1.A1 but for day+7.

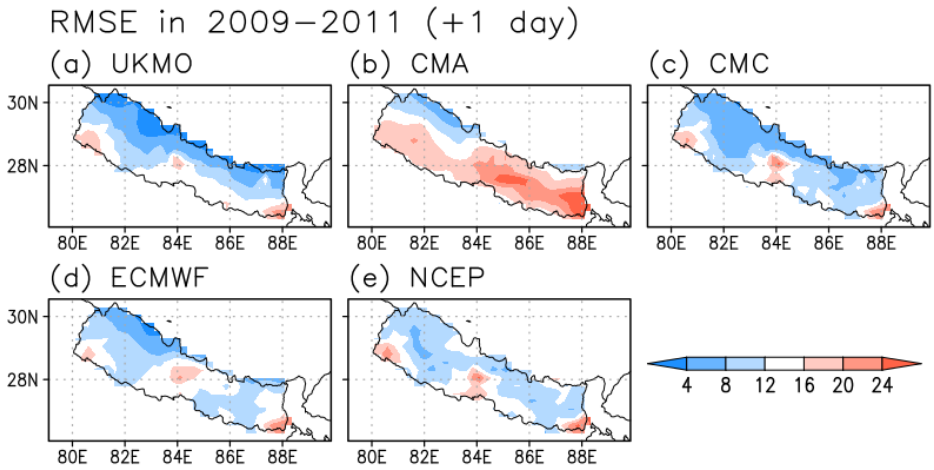


Figure B.4.3.1.B1. Root mean square error (RMSE) during training period (JJAS, 2009-2011) for day+1.

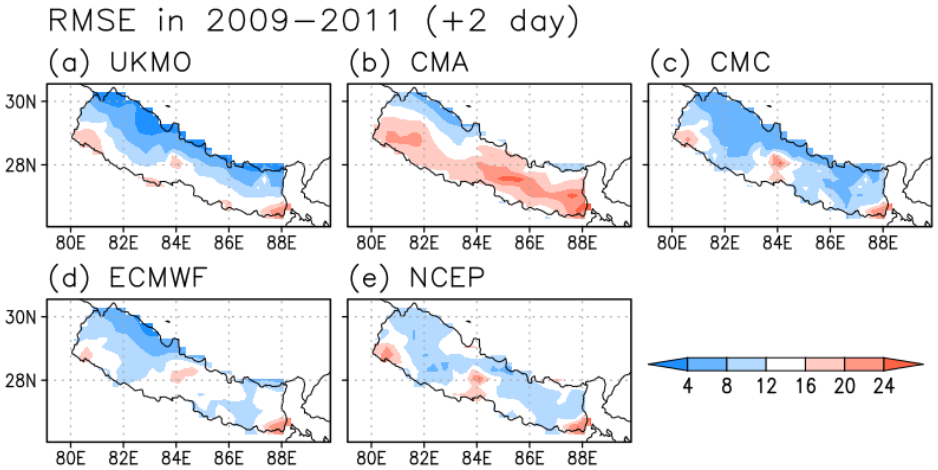


Figure B.4.3.1.B2. Same as figure B.4.3.1.B1 but for day+2.

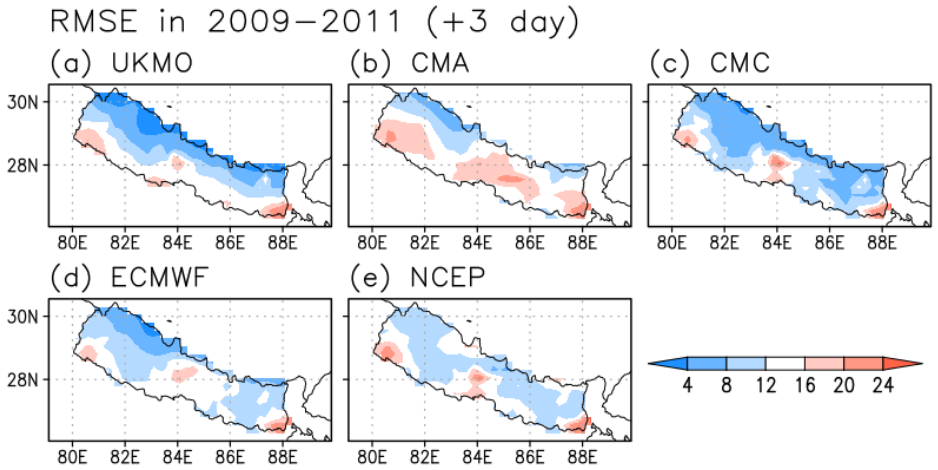


Figure B.4.3.1.B3. Same as figure B.4.3.1.B1 but for day+3.

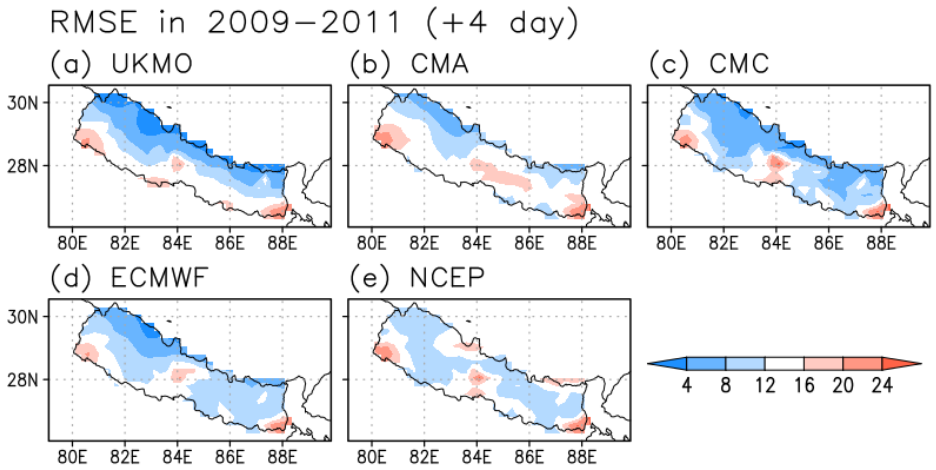


Figure B.4.3.1.B4. Same as figure B.4.3.1.B1 but for day+4.

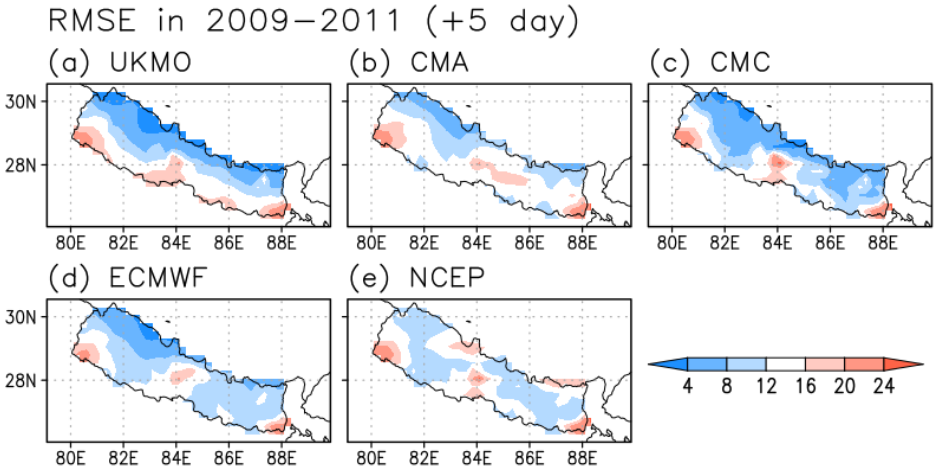


Figure B.4.3.1.B5. Same as figure B.4.3.1.B1 but for day+5.

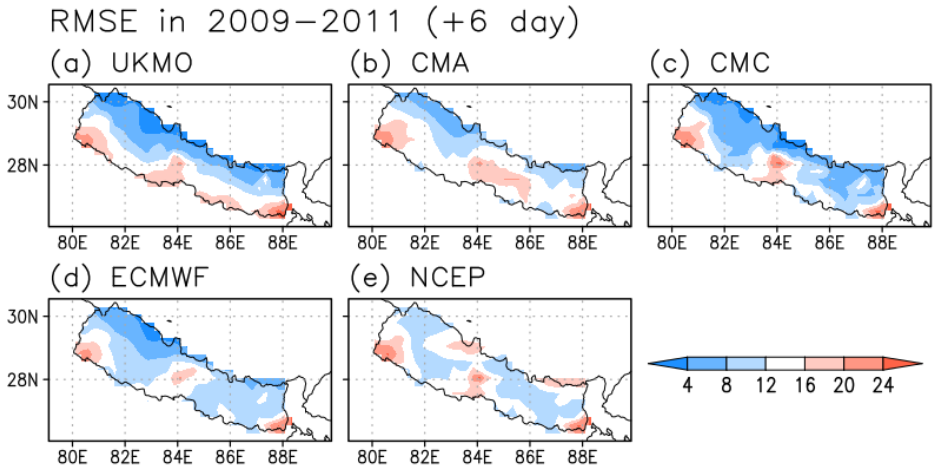


Figure B.4.3.1.B6. Same as figure B.4.3.1.B1 but for day+6.

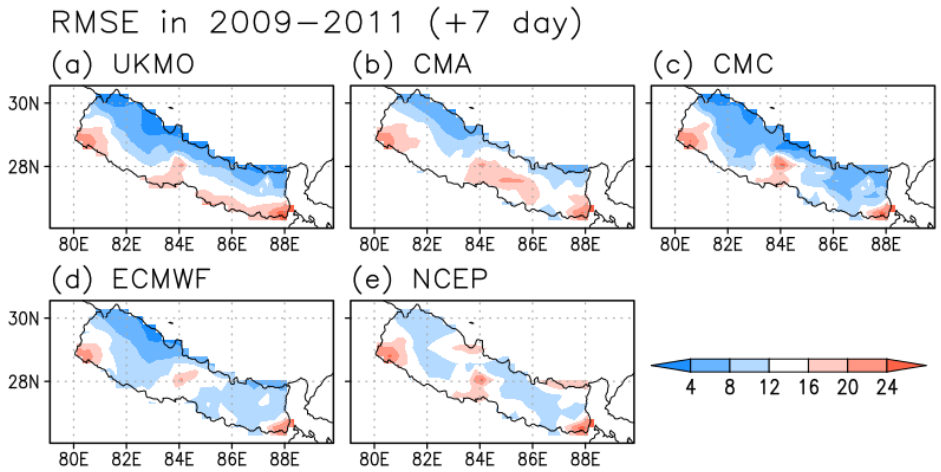


Figure B.4.3.1.B7. Same as figure B.4.3.1.B1 but for day+7.

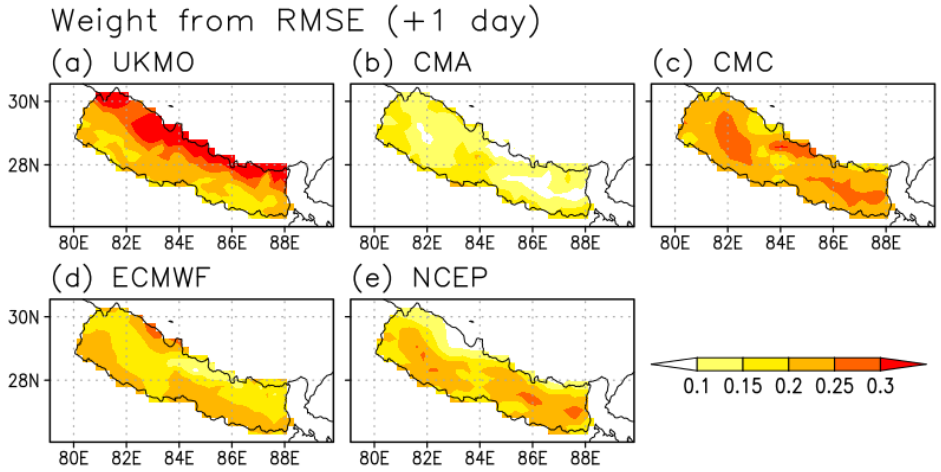


Figure B.4.3.1.C1. Weight based on RMSE during training period (JJAS, 2009-2011) for day+1.

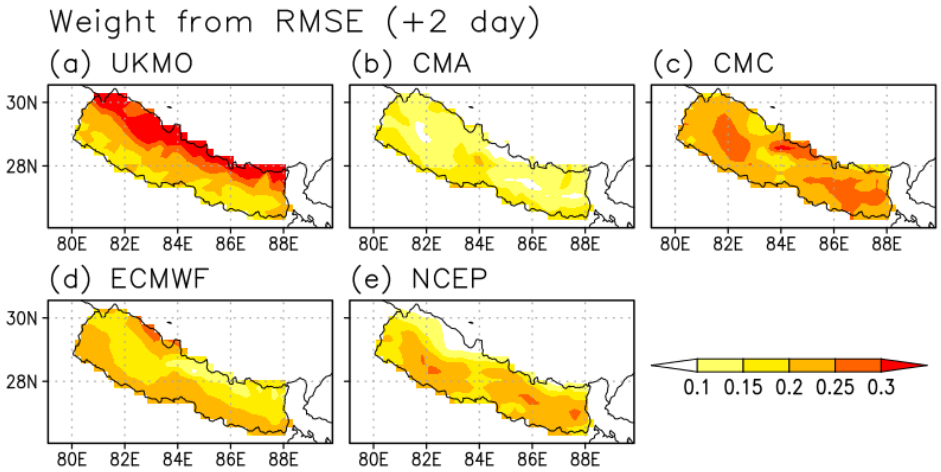


Figure B.4.3.1.C2. Same as Figure B.4.3.1.C1 but for day+2.

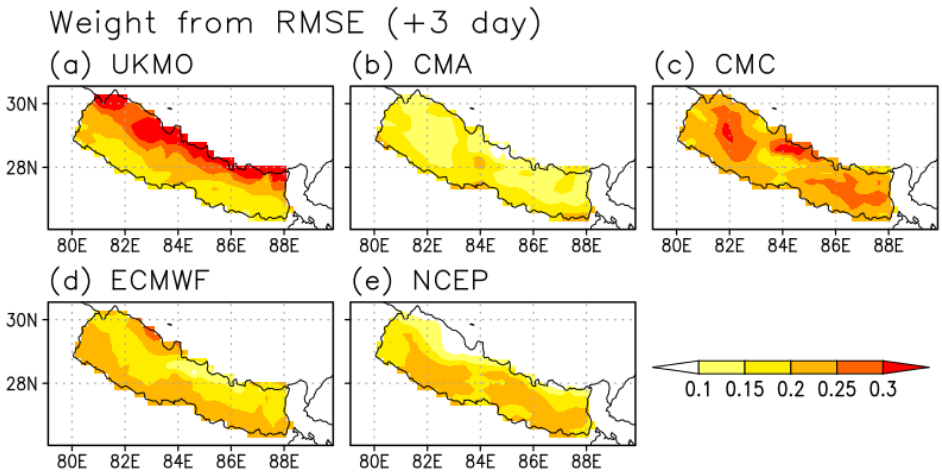


Figure B.4.3.1.C3. Same as Figure B.4.3.1.C1 but for day+3.

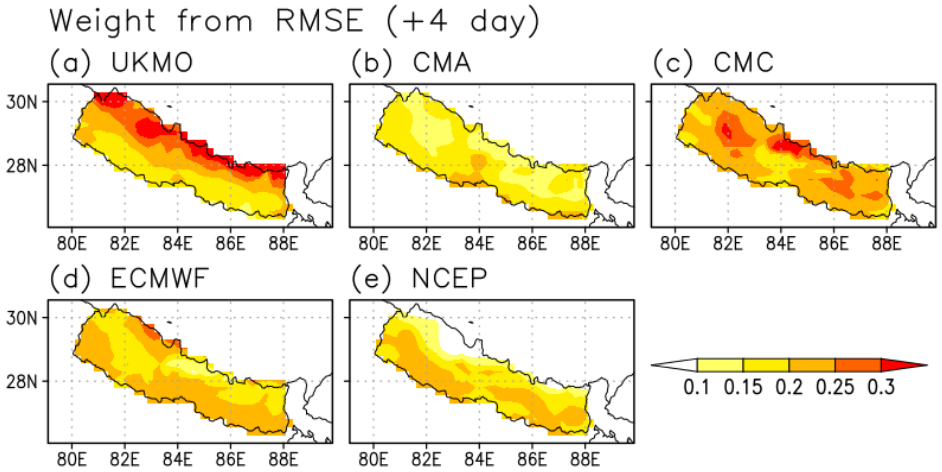


Figure B.4.3.1.C4. Same as Figure B.4.3.1.C1 but for day+4.

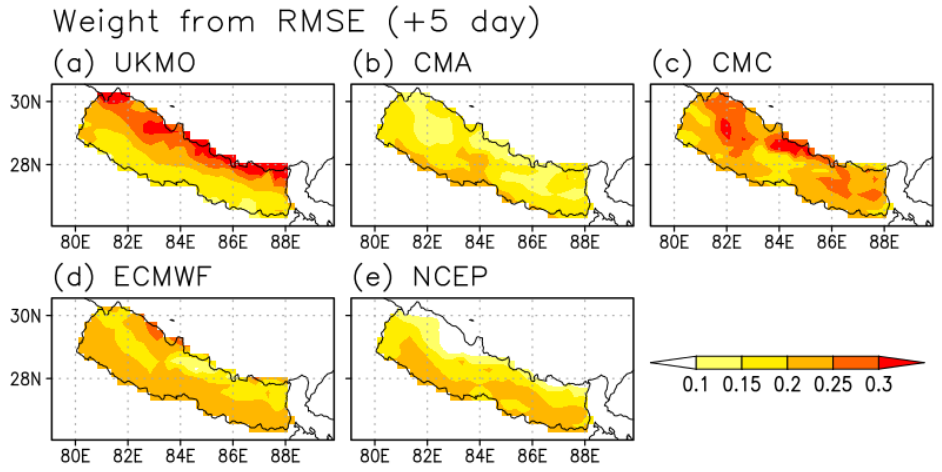


Figure B.4.3.1.C5. Same as Figure B.4.3.1.C1 but for day+5.

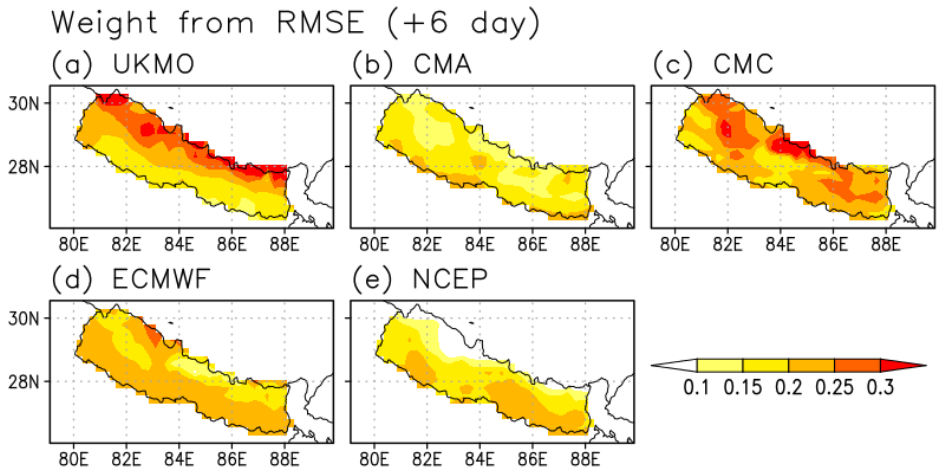


Figure B.4.3.1.C6. Same as Figure B.4.3.1.C1 but for day+6.

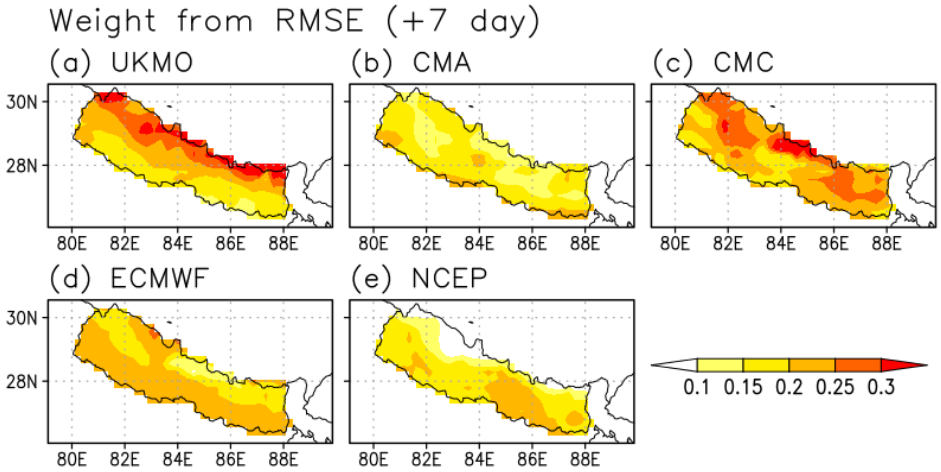


Figure B.4.3.1.C7. Same as Figure B.4.3.1.C1 but for day+7.

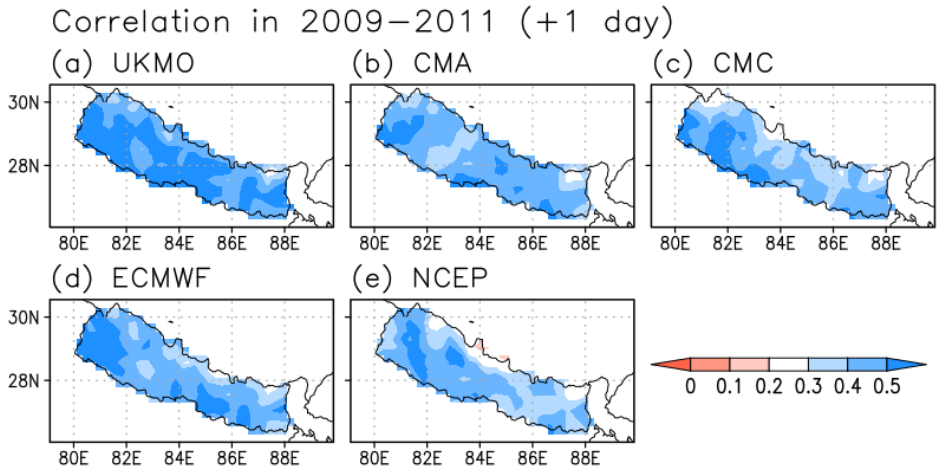


Figure B.4.3.1.D1. Correlation coefficient during training period (JJAS, 2009-2011) for day+1.

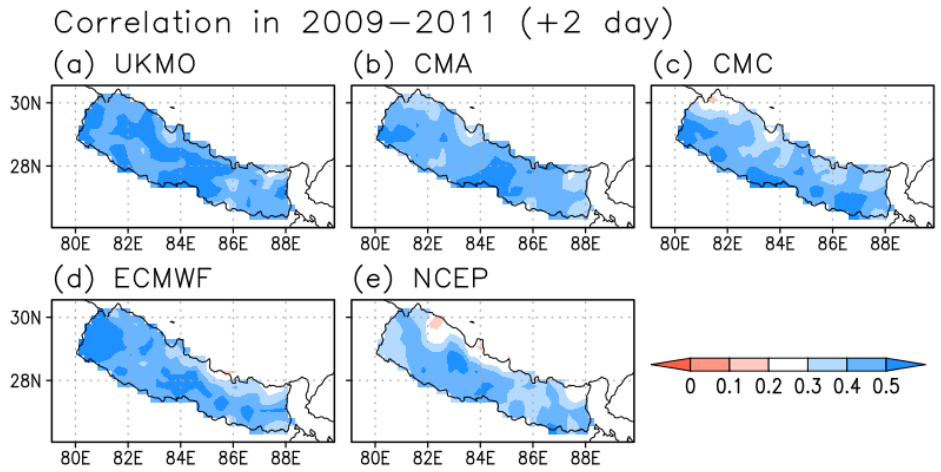


Figure B.4.3.1.D2. Same as Figure B.4.3.1.D1 but for day+2.

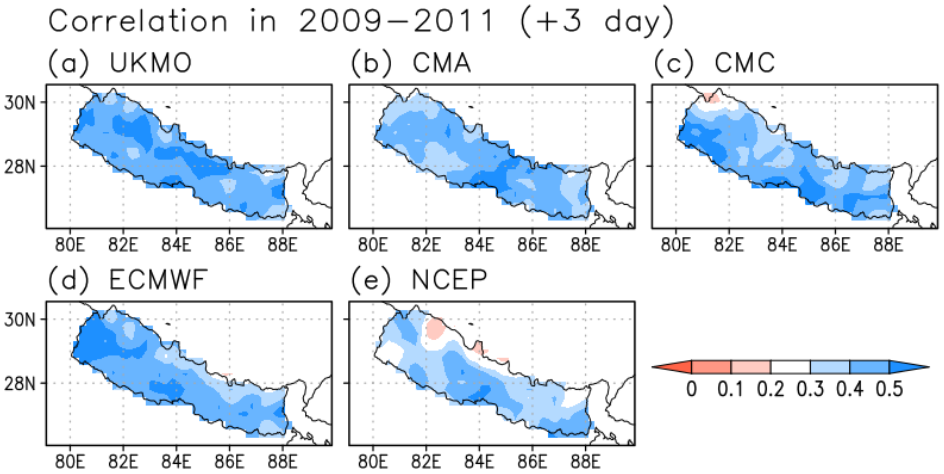


Figure B.4.3.1.D3. Same as Figure B.4.3.1.D1 but for day+3.

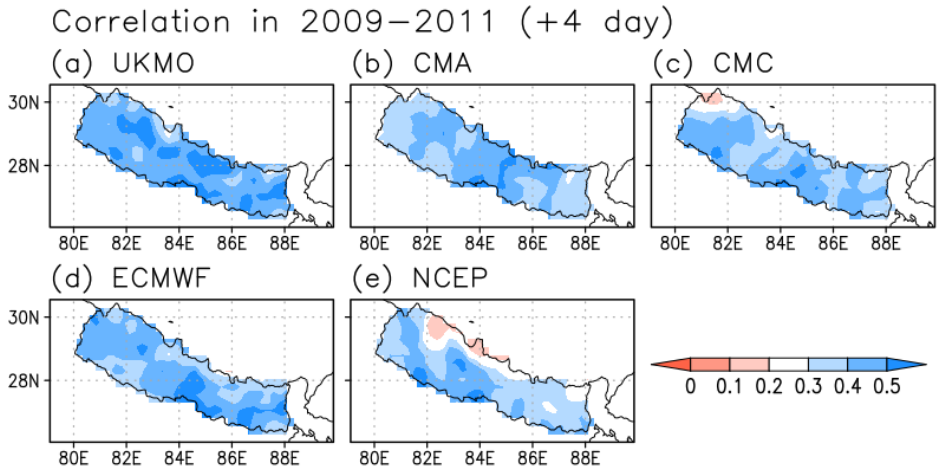


Figure B.4.3.1.D4. Same as Figure B.4.3.1.D1 but for day+4.

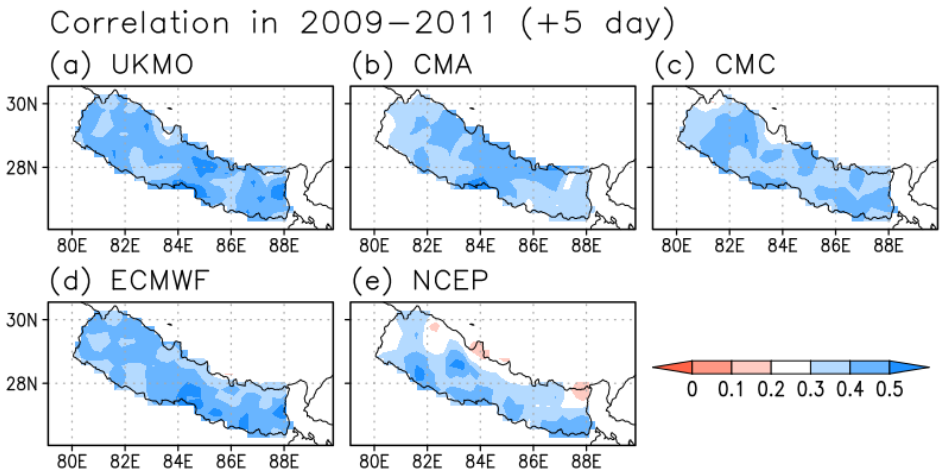


Figure B.4.3.1.D5. Same as Figure B.4.3.1.D1 but for day+5.

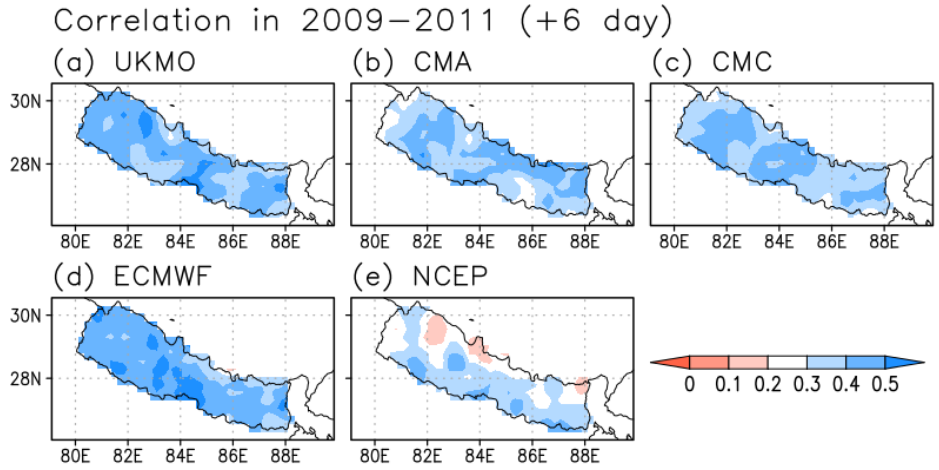


Figure B.4.3.1.D6. Same as Figure B.4.3.1.D1 but for day+6.

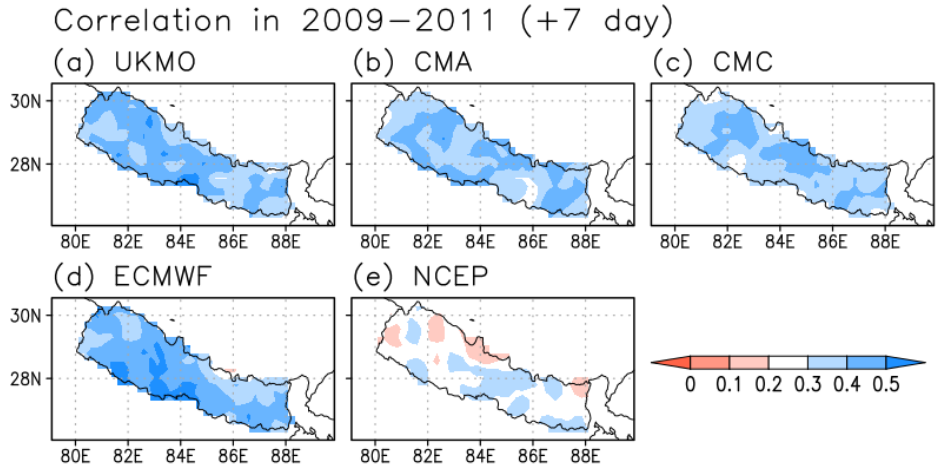


Figure B.4.3.1.D7. Same as Figure B.4.3.1.D1 but for day+7.

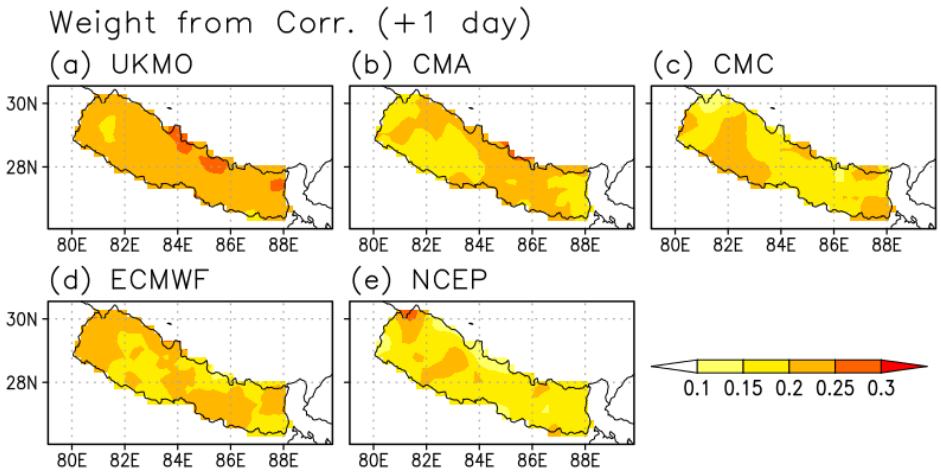


Figure B.4.3.1.E1. Weight based on Correlation coefficient during training period (JJAS, 2009-2011) for day+1.

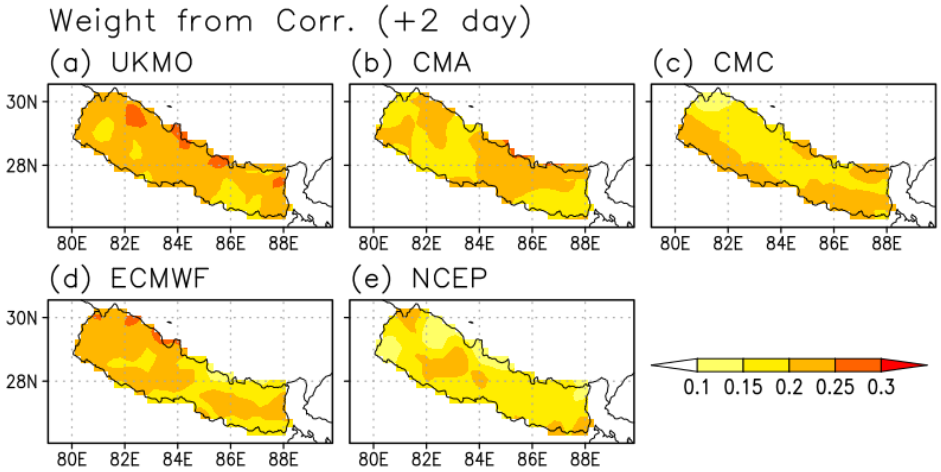


Figure B.4.3.1.E2. Same as Figure B.4.3.1.E1 but for day+2.

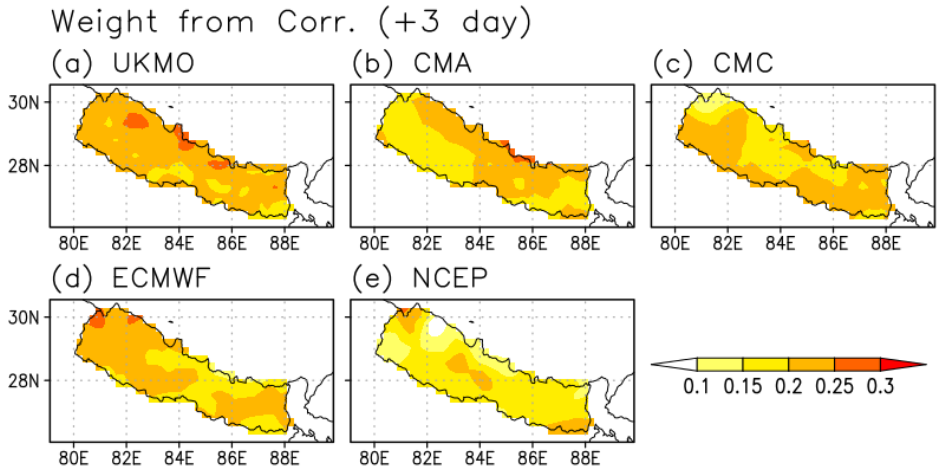


Figure B.4.3.1.E3. Same as Figure B.4.3.1.E1 but for day+3.

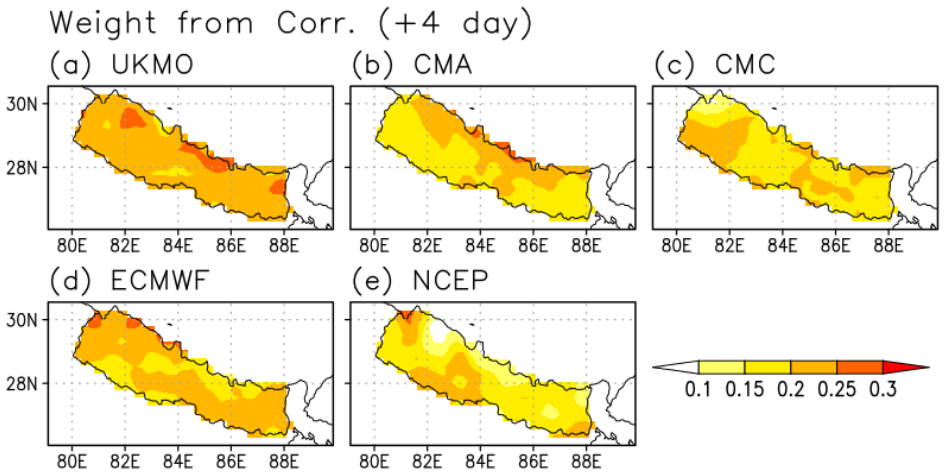


Figure B.4.3.1.E4. Same as Figure B.4.3.1.E1 but for day+4.

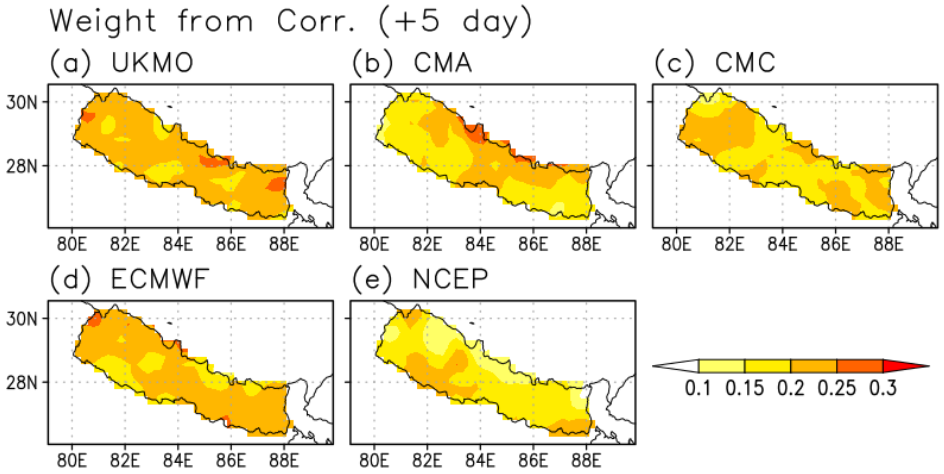


Figure B.4.3.1.E5. Same as Figure B.4.3.1.E1 but for day+5.

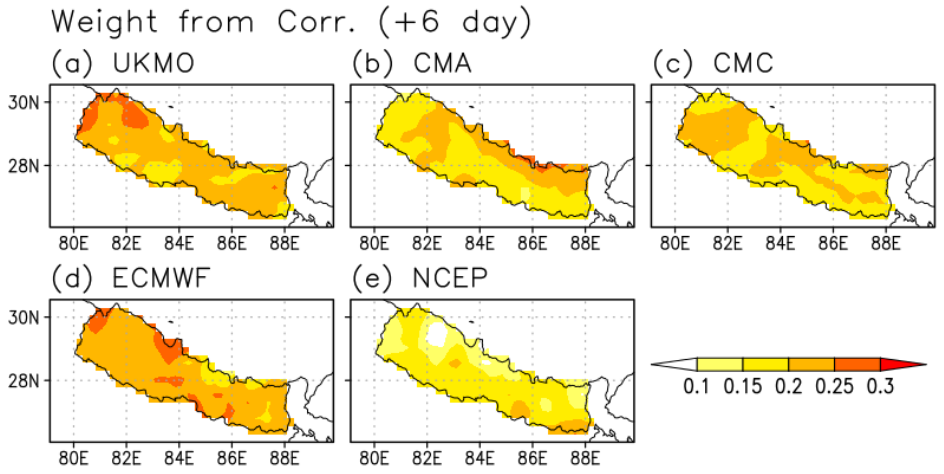


Figure B.4.3.1.E6. Same as Figure B.4.3.1.E1 but for day+6.

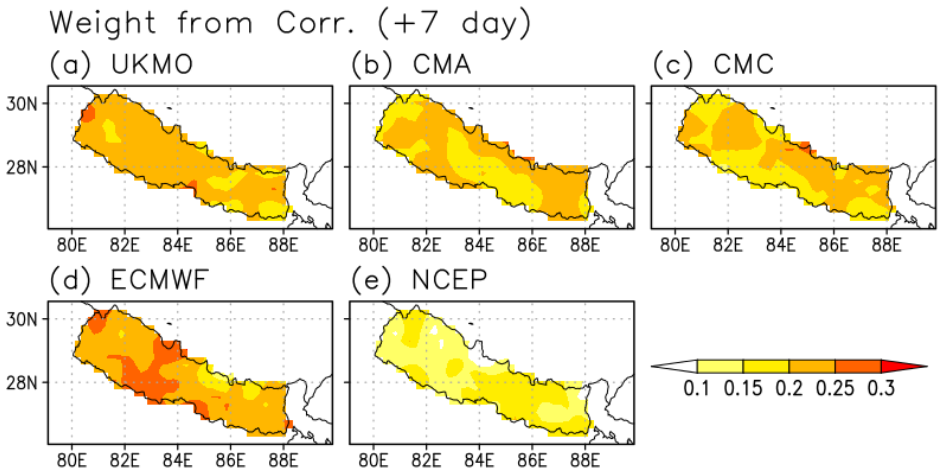


Figure B.4.3.1.E7. Same as Figure B.4.3.1.E1 but for day+7.

Appendix C

Source:http://www.npr.org/sections/goatsandsoda/2014/09/10/347398623/nepal-struggles-to-help-villages-washed-away-in-floods.

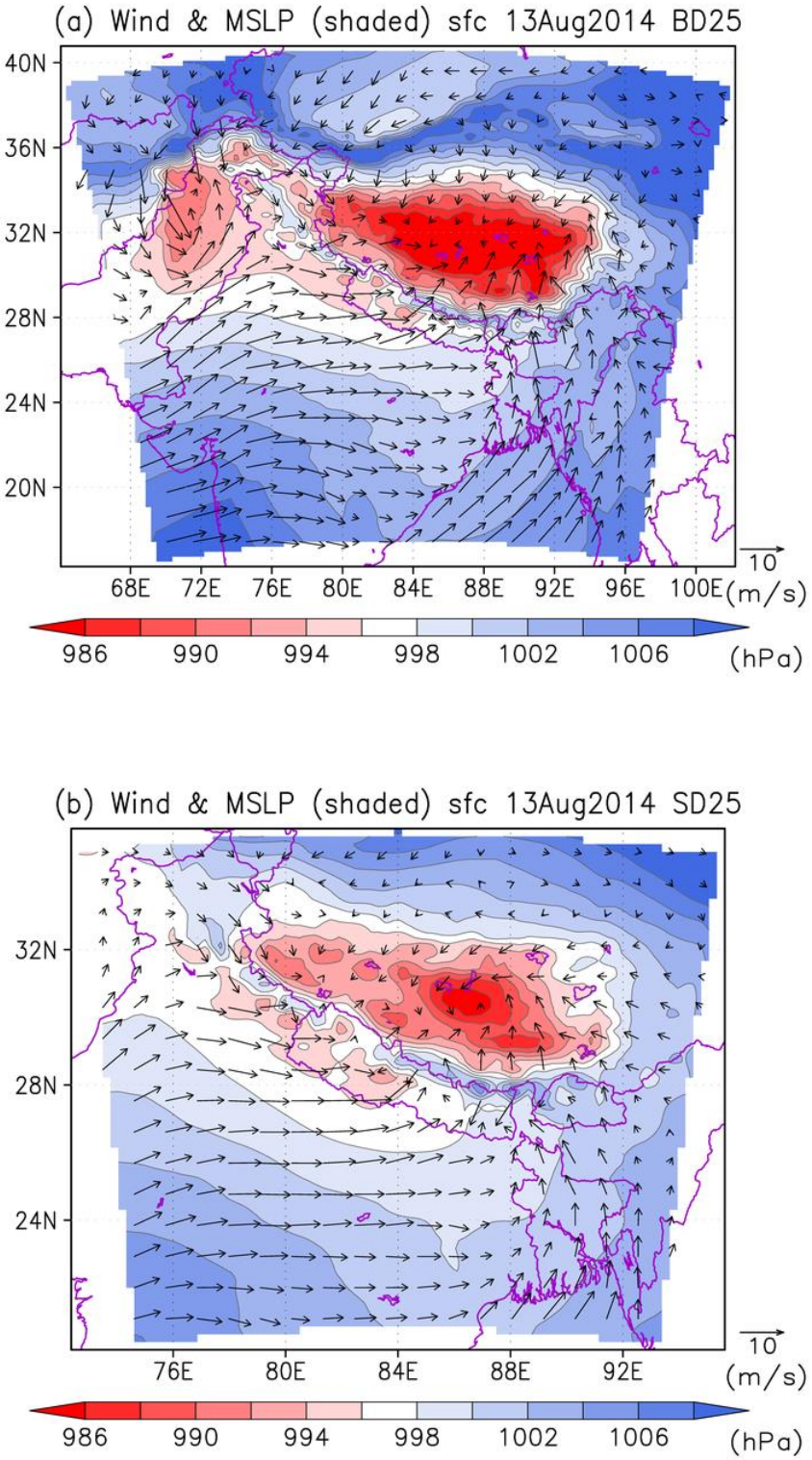


Figure C.1. Simulated mean seas level pressure (MSLP, Shaded) in hPa and wind at 10m (vectors) in m/s in (a) BD25 (b) SD25 on 13 August, 2014.

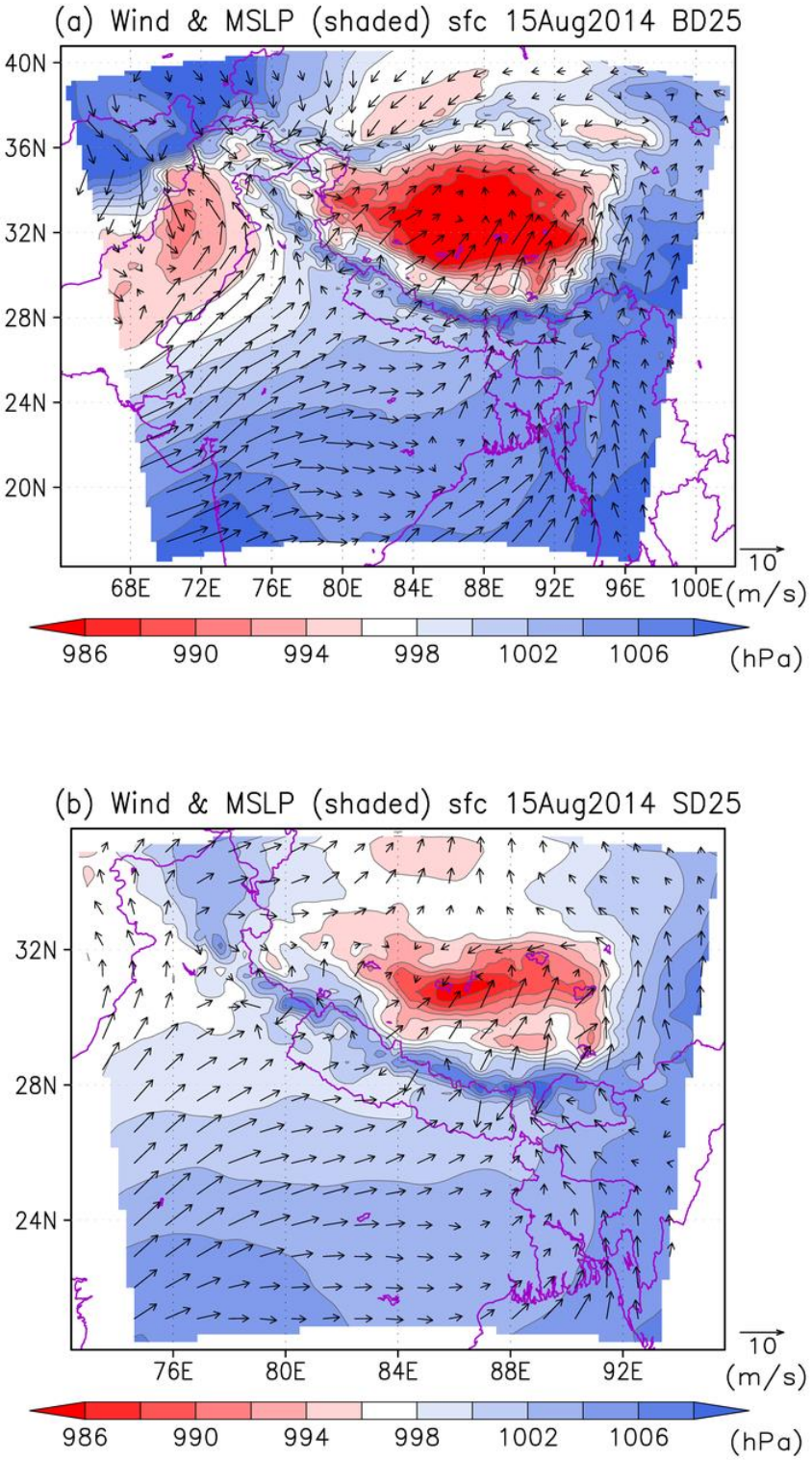


Figure C.2. Same as figure C.1 but on 15 August, 2014.

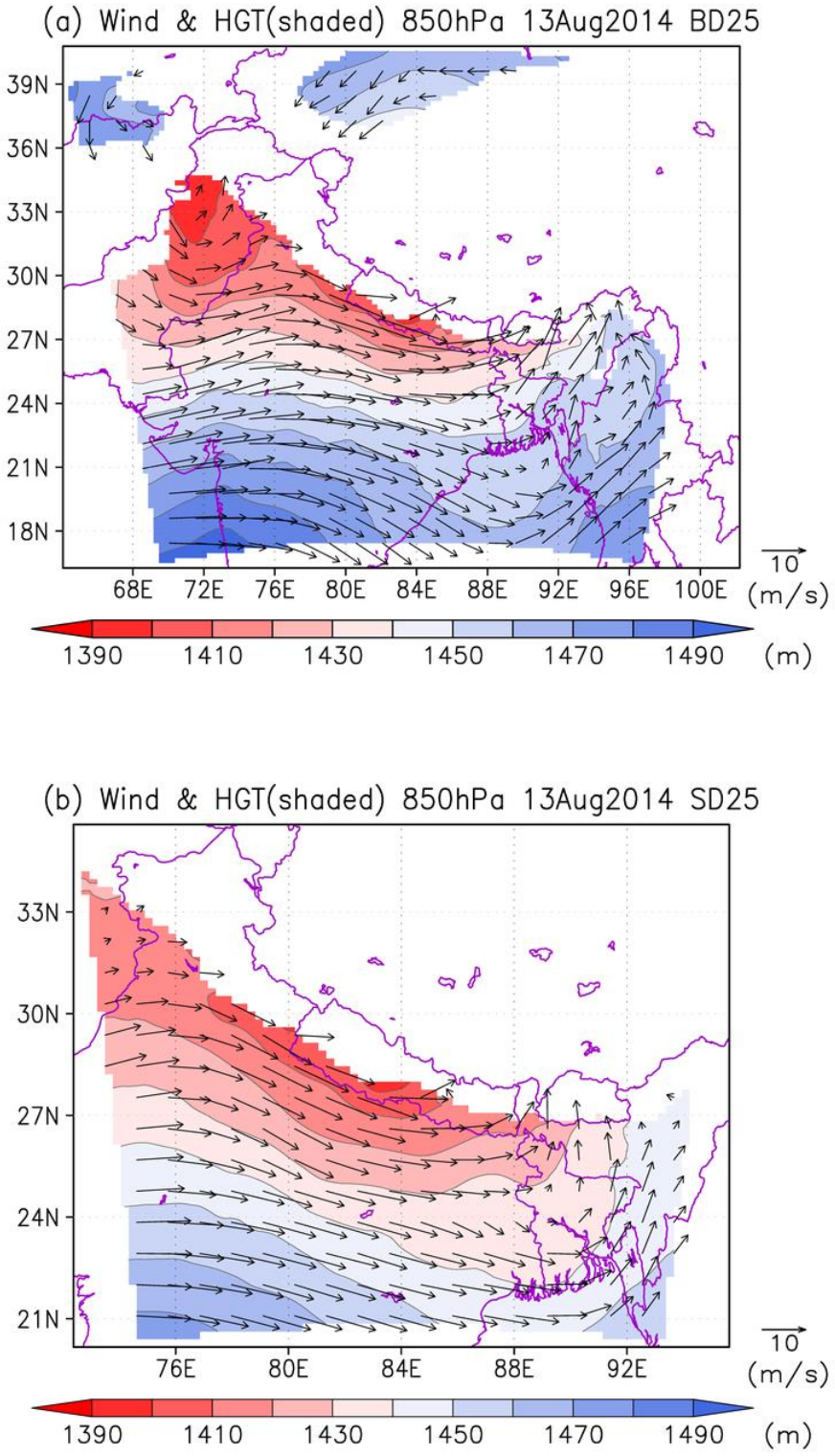


Figure C.3. Simulated geopotential height (HGT, Shaded) in m and wind (vectors) in m/s at 850 hPa in (a) BD25 (b) SD25 on 13 August, 2014.

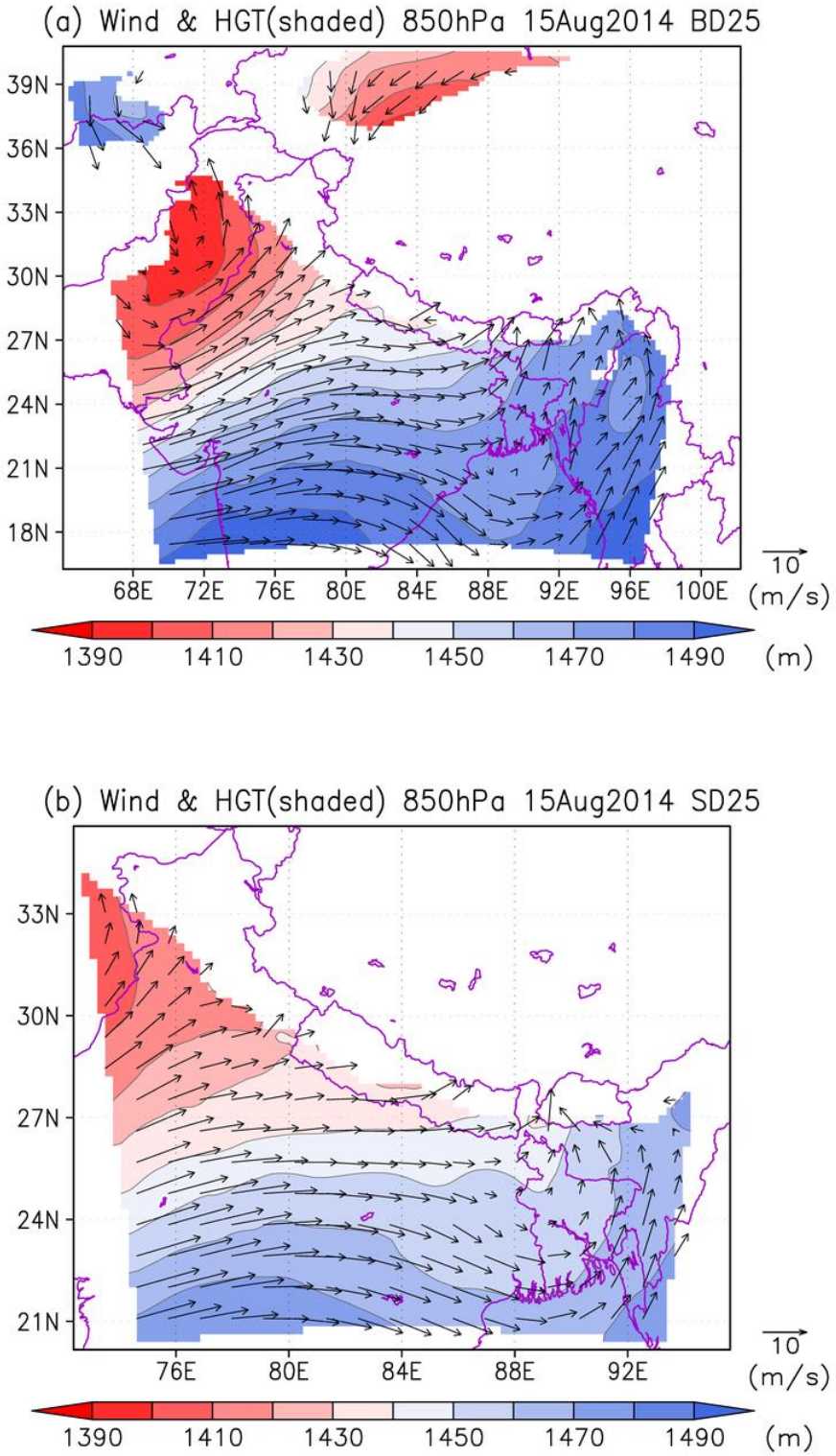


Figure C.4. Same as figure C.3 but on 15August, 2014.

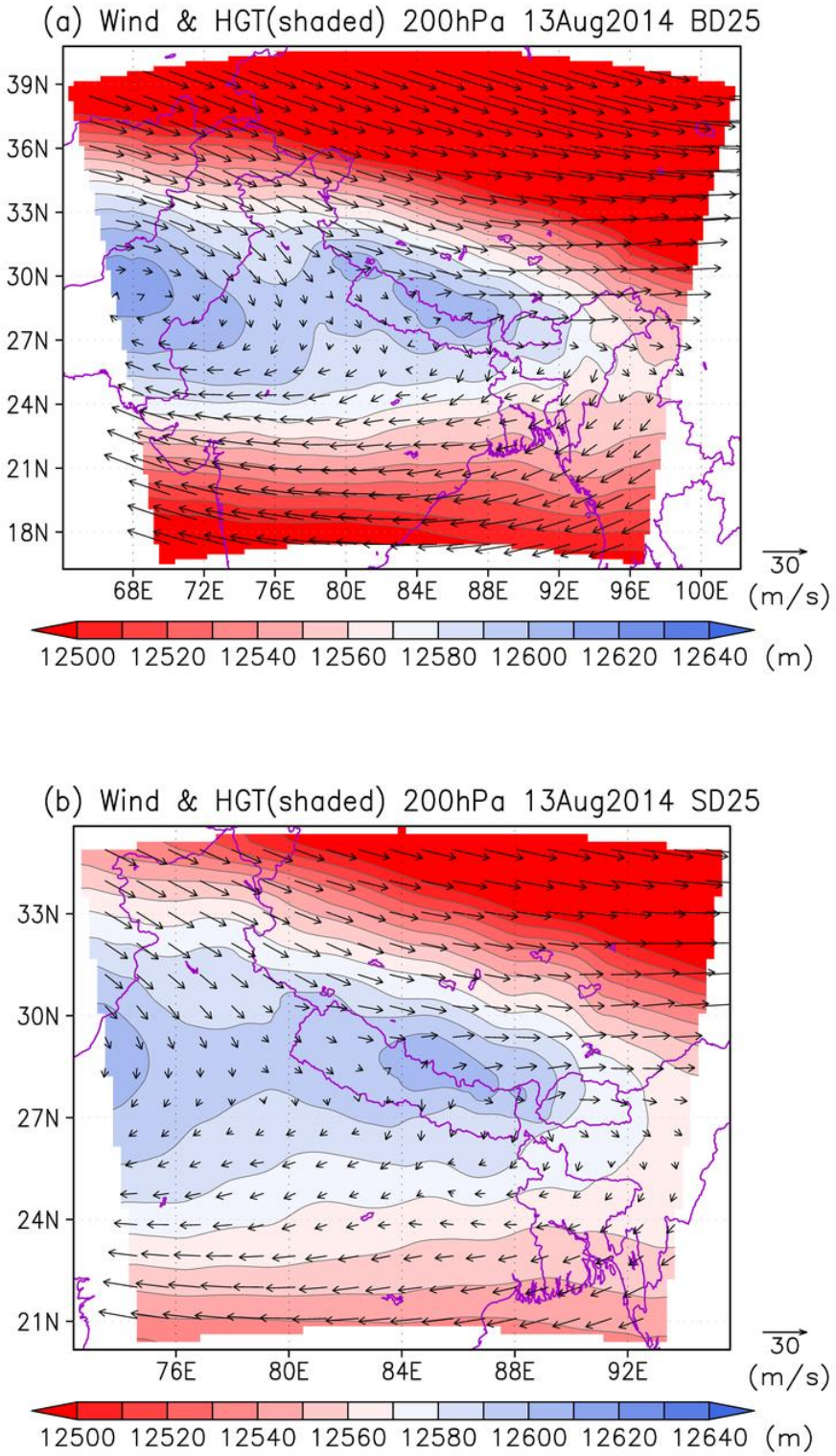


Figure C.5. Same as figure C.3 but at 200 hPa.

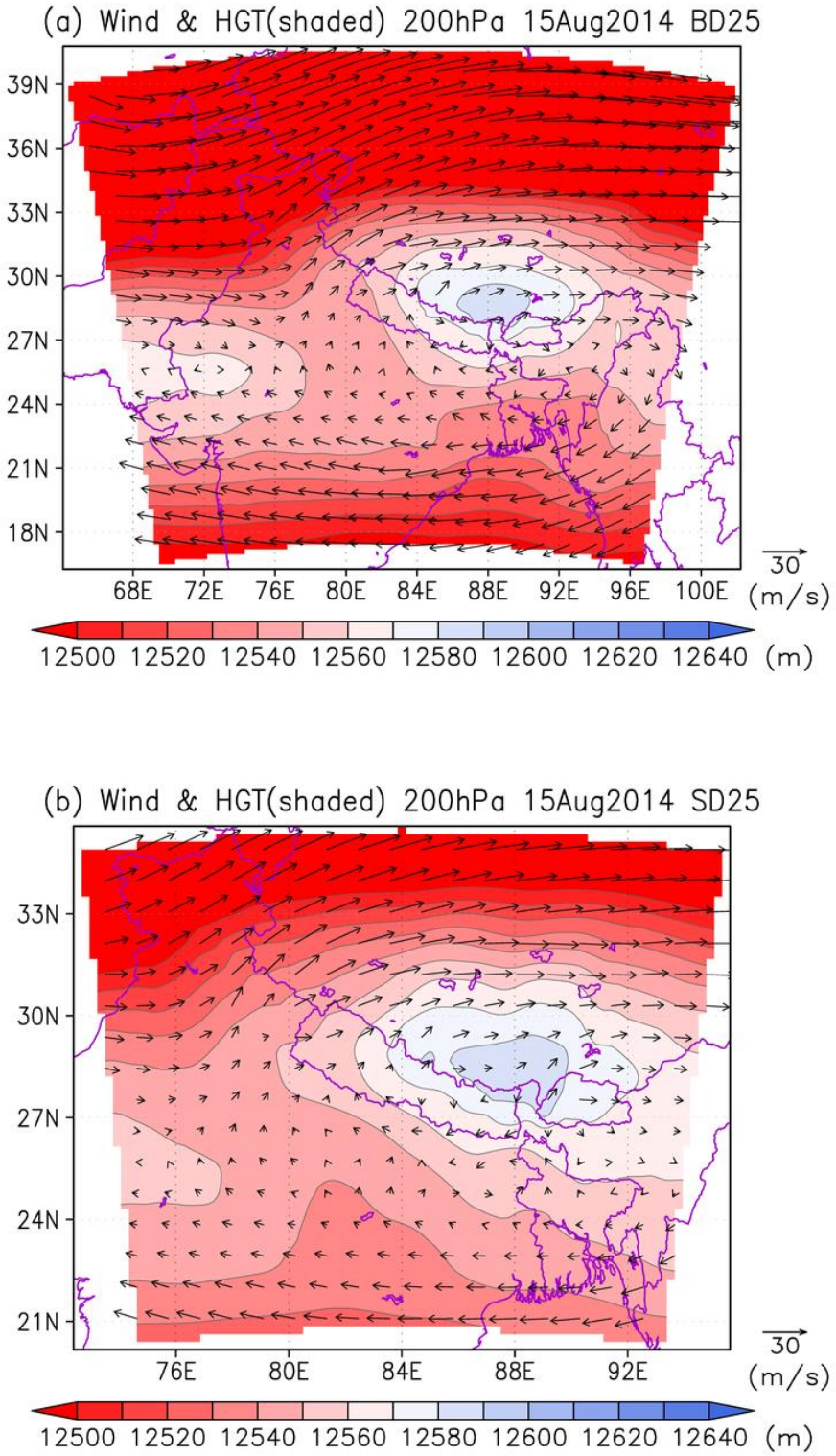


Figure C.6. Same as figure C.4 but at 200 hPa.

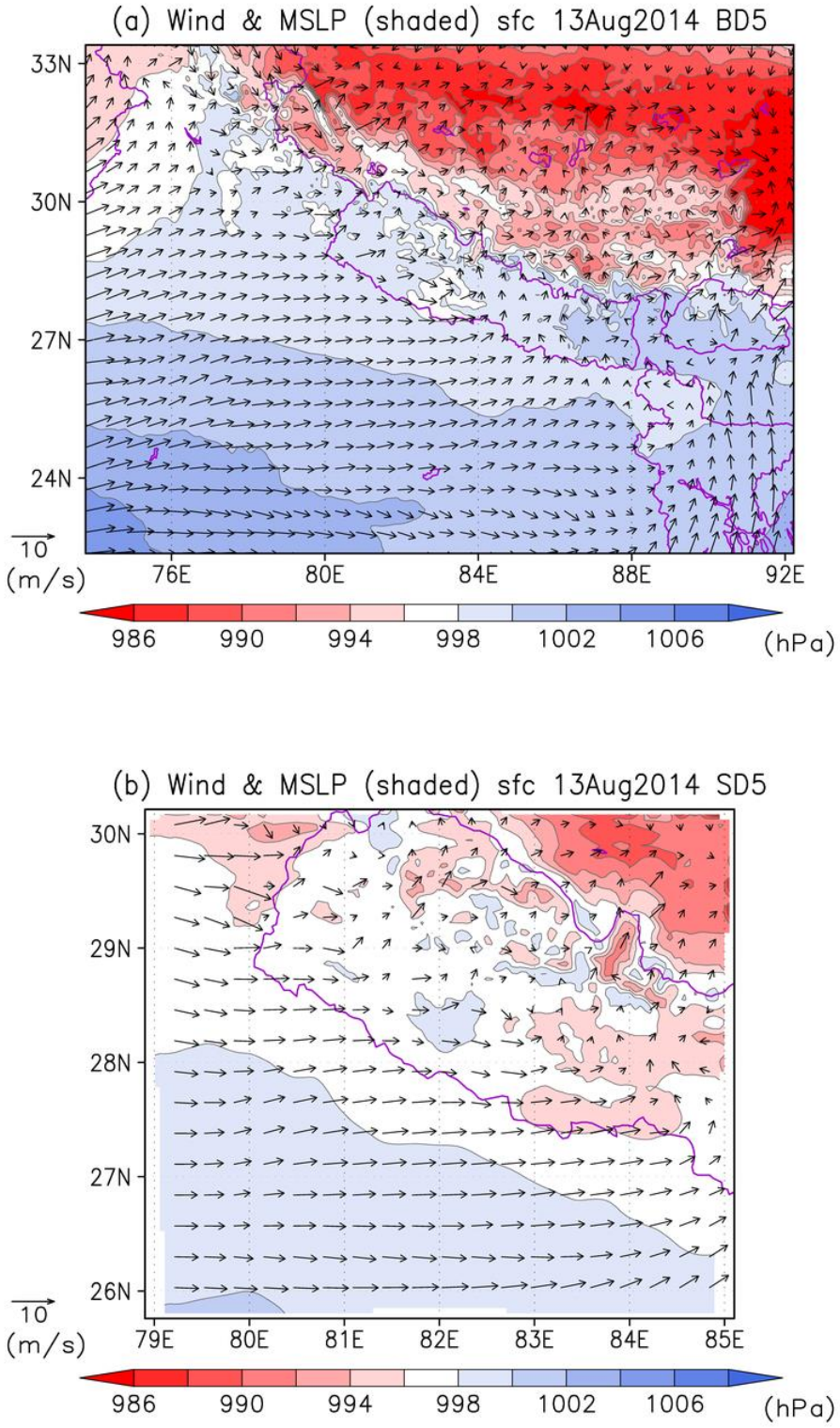


Figure C.7. Simulated mean seas level pressure (MSLP, Shaded) in hPa and wind at 10m (vectors) in m/s in (a) BD5 (b) SD5 on 13 August, 2014.

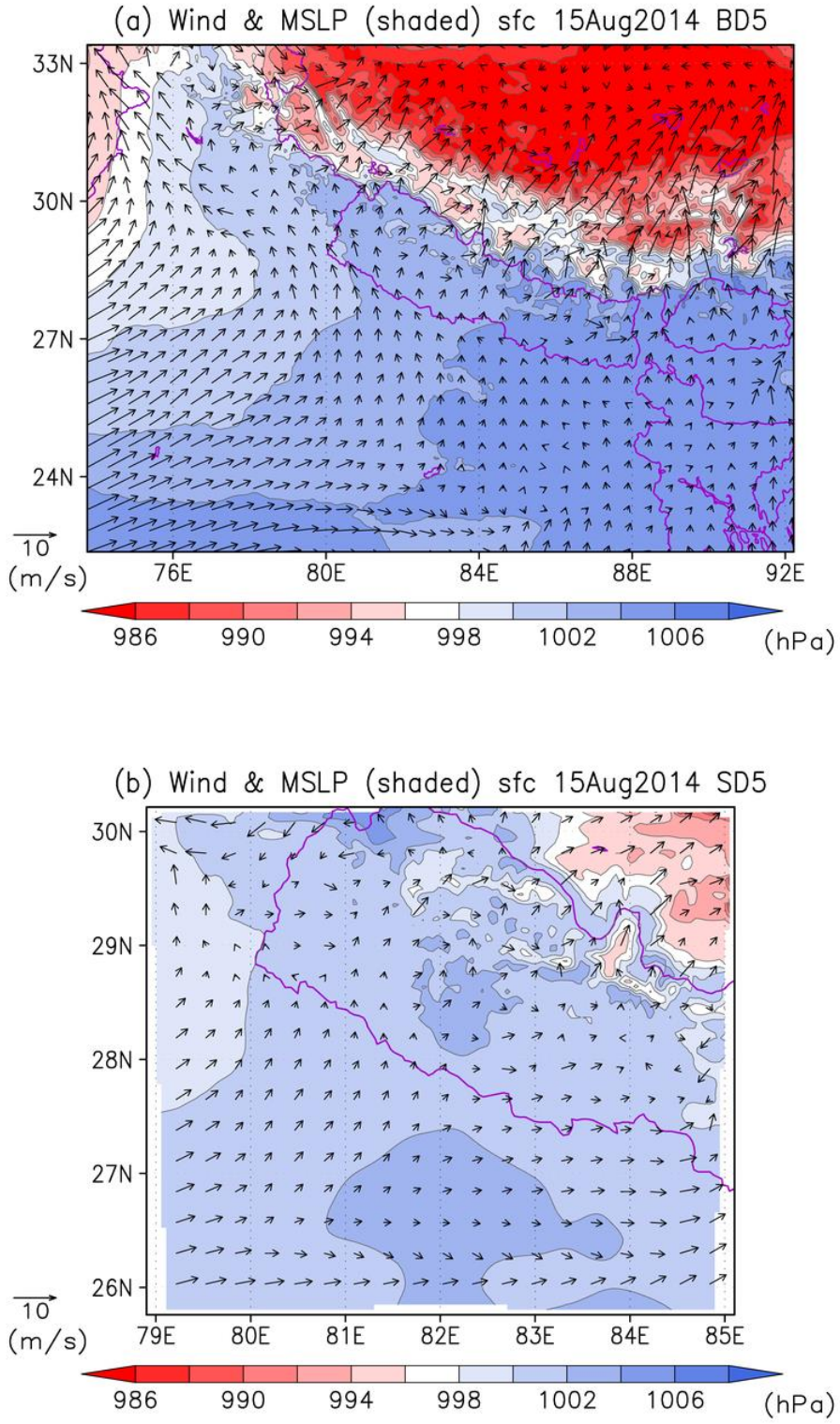


Figure C.8 Same as figure C.7 but on 15 August, 2014.

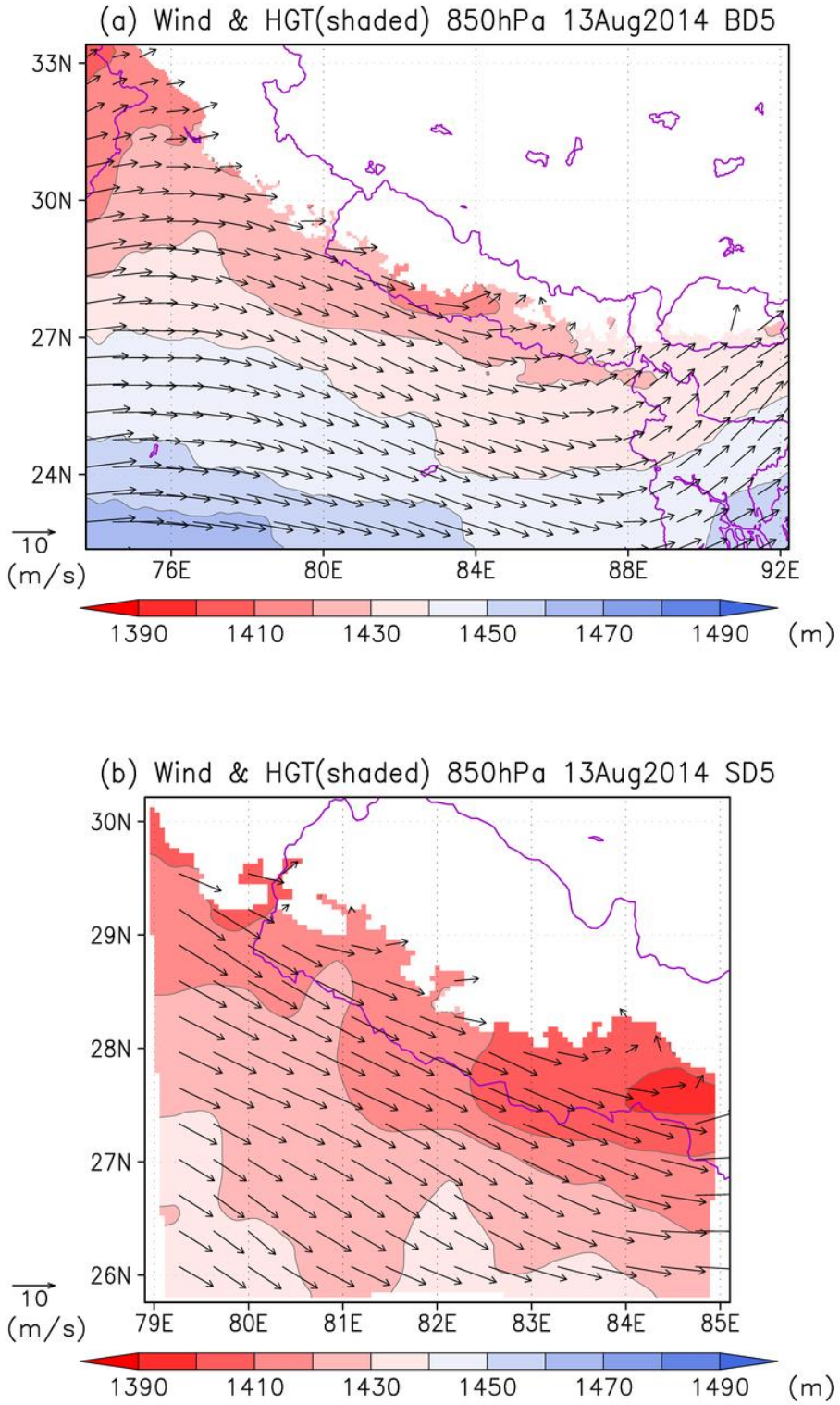


Figure C.9. Simulated geopotential height (HGT, Shaded) in m and wind (vectors) in m/s at 850 hPa in (a) BD5 (b) SD5 on 13 August, 2014.

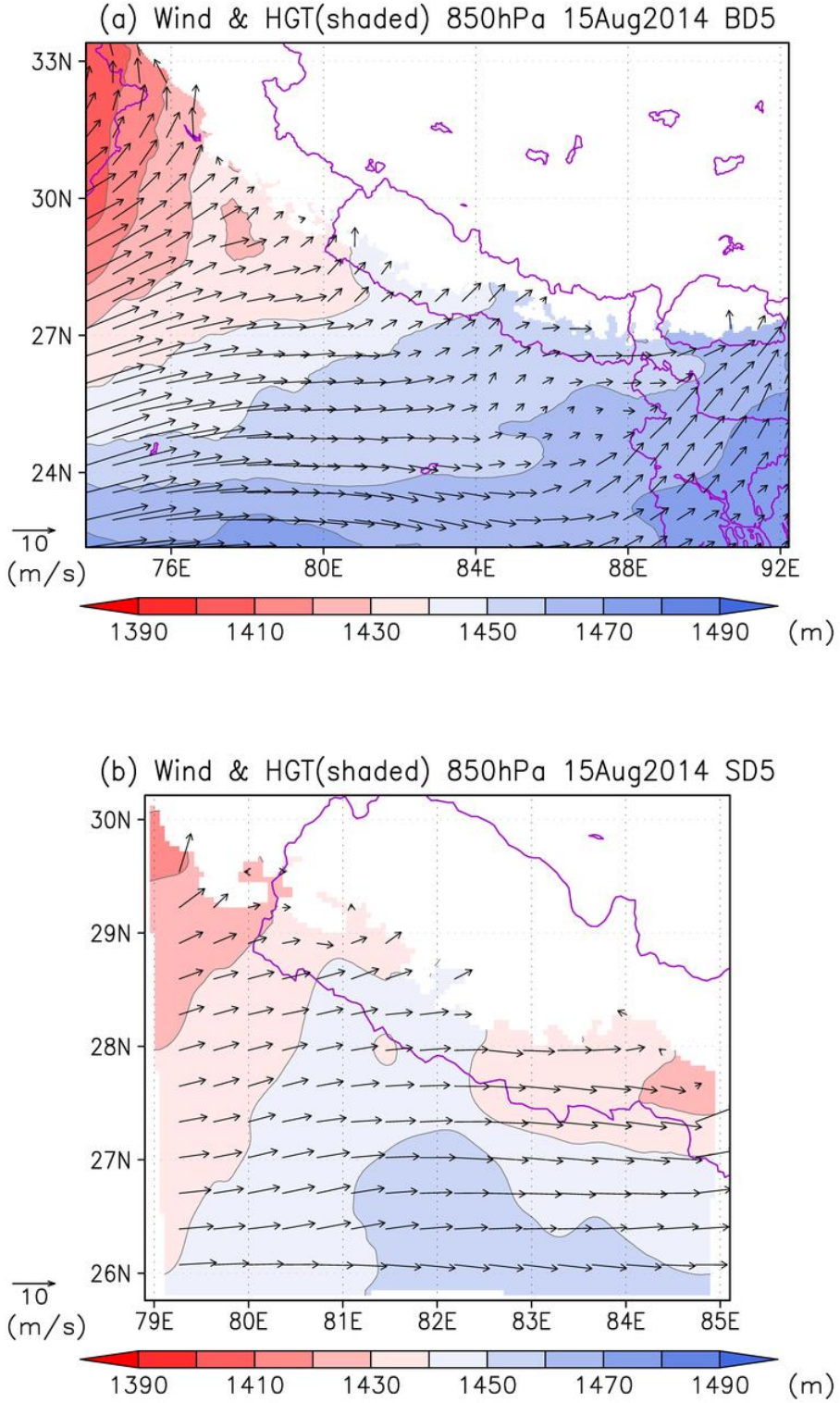


Figure C.10. Same as figure C.9 but on 15August, 2014.

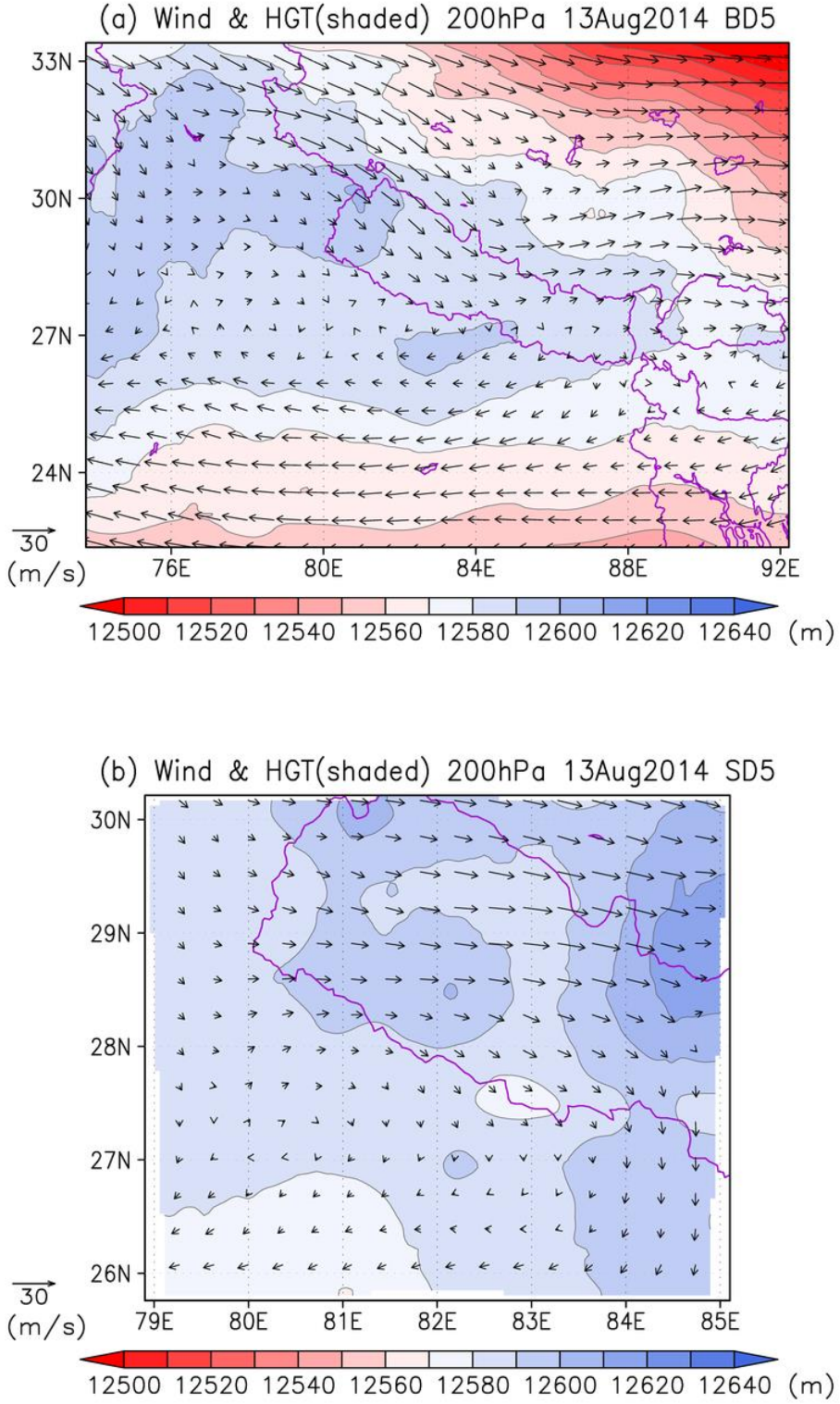


Figure C.11. Same as figure C.9 but at 200 hPa.

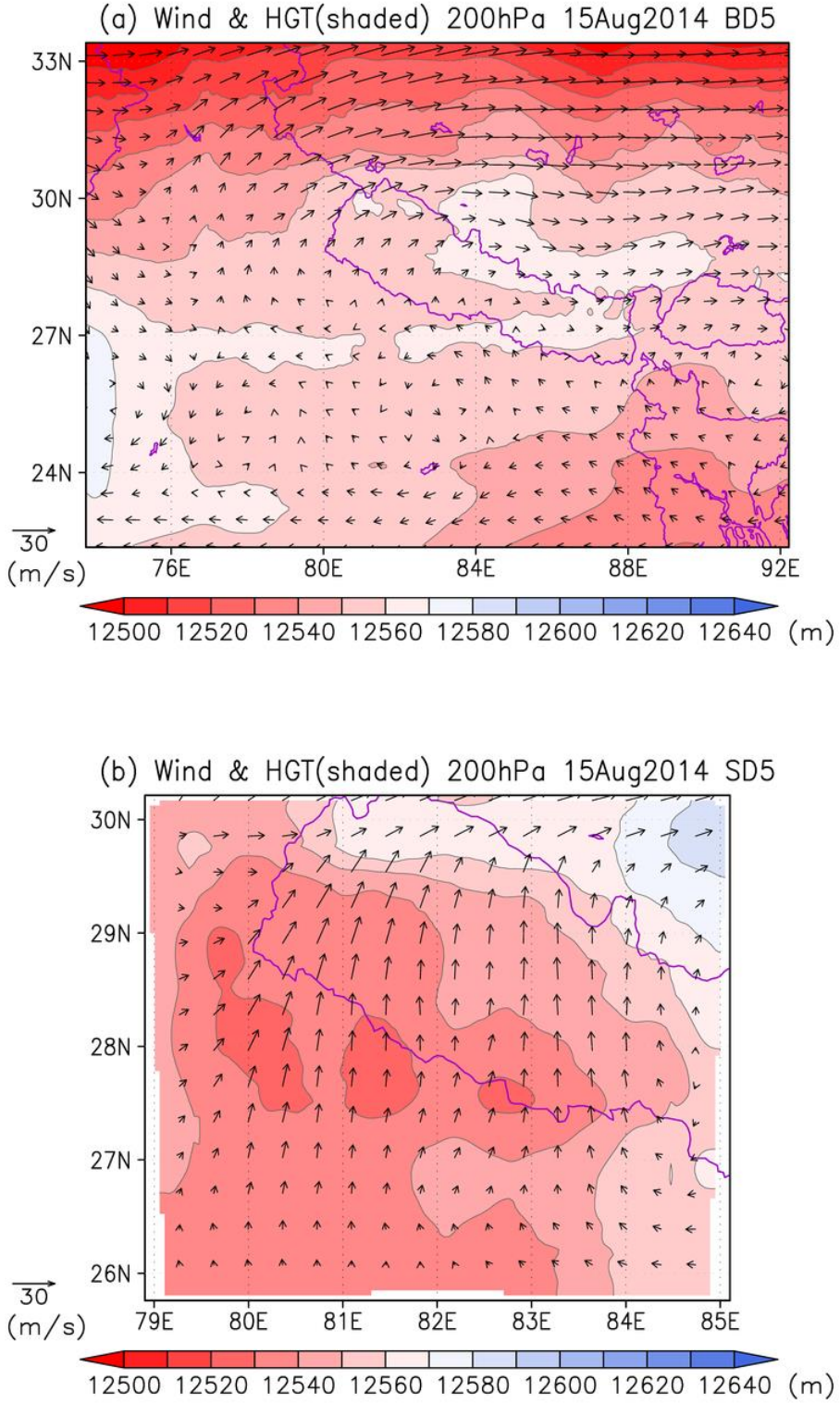


Figure C.12. Same as figure C.10 but at 200 hPa.

Reference

Ageta, Y., N. Natio, M. Nakawa, K. Fujita, K. Shankar, A. P. Pokhrel, and D. Wangda, 2001: Study project on the recent rapid shrinkage of summer-accumulation type glaciers in the Himalayas, 1997-1999. *Bulletin of Glaciological Research*, **18**, 45-49.

Alves, J. M. B., F. C. Vasconcelos Jr, R. R. Chaves, E. M. Silva, J. Servain, A. A. Costa, S. S. Sombra, A. C. B. Barbosa, and A. C.S. dos Santos, 2016: Evaluation of the AR4 CMIP3 and the AR5 CMIP5 Model and Projections for Precipitation in Northeast Brazil. *Frontiers in Earth Science*, **4**, 44, doi: 10.3389/feart.2016.00044

Andermann, C., S. Bonnet and R. Gloaguen, 2011: Evaluation of precipitation data sets along the Himalayan front. *Geochemistry Geophysics Geosystems*, **12**, Q07023, doi: 10.1029/2011GC003513

Anders, A. M., G. H. Roe, B. Hallet, D. R. Montgomery, N. J. Finnegan, and J. Putkonen, 2006: Spatial patterns of precipitation and topography in the Himalaya. *Geological Society of America, Special Paper*, **398**, 39-53, doi: 10.1130/2006.2398(03)

Annamalai, H., K. Hamilton, and K. R. Sperber, 2007: The South Asian summer monsoon and its relationship with ENSO in the IPCC AR4 simulations. *Journal of Climate*, **20**, 1071–1092.

Arakawa, O., and A. Kitoh, 2011: Intercomparison of the Relationship between Precipitation and Elevation among Gridded Precipitation Datasets over the Asian Summer Monsoon Region. *Global Environmental Research,* **15**, 109-118.

Baidya, S. K., M. L. Shrestha, and M. M. Sheikh, 2008: Trends in daily climatic extremes of temperature and precipitation in Nepal. *Journal of Hydrology and Meteorology*, **5** (1), 38-51.

Barros A. P., G. Kim, E. Williams, and S. W. Nesbitt, 2004: Probing orographic controls in the Himalayas during the monsoon using satellite imagery. *Natural Hazards Earth System Sciences*, **4**, 29-51.

Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525**, 47–55, doi:10.1038/*nature*14956.

Beniston, M., 2003: Climatic Change in Mountain Regions: A Review of Possible Impacts. *Climatic Change*, **59**, 5–31.

Bi et al., 2013: The ACCESS coupled model: description, control climate and evaluation. *Australian Meteorological and Oceanographic Journal*, **63**, 41-64.

Bougeault, P., and Coauthors, 2010: The THORPEX Interactive Grand Global Ensemble (TIGGE). *Bulletin of the American Meteorological Society*, ***91***, 1059–1072, doi:10.1175/2010BAMS2853.1

Bowler, N. E., A. Arribas, K. R. Mylne, K. B. Robertso, and S. E. Beare, 2008: The MOGREPS short-range ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, **134,** 703–722.

Buizza, R., J. R. Bidlot, N. Wedi, M. Fuentes, M. Hamrud, G. Holt, and F. Vitart, 2007: The new ECMWF VAREPS (Variable Resolution Ensemble Prediction System). *Quarterly Journal of the Royal Meteorological Society*, **133**, 681–695.

Buizza, R., and T. N. Palmer, 1998: Impact of ensemble size on ensemble prediction. *Monthly Weather Review*, **126**, 2503–2518.

Buizza, R., M. Leutbecher, L. Isaksen, and J. Haseler, 2010: Combined use of EDA-and SV-based perturbations in the EPS, *ECMWF Newsletter*, **123**, 22–28.

Castro, C. L., R. A. Sr Pielke, and G. Leoncini, 2005: Dynamical downscaling: assessment of value retained and added using the Regional Atmopsheric Modeling System (RAMS). *Journal of Geophysical Research*, **110,** 1–21, doi: 10.1029/2004JD004721.

Chadwick, R., I. Boutle and G. Martin, 2013: Spatial Patterns of Precipitation Change in CMIP5: Why the Rich Do Not Get Richer in the Tropics, *Journal of Climate*, **26**, 3803-3822, doi: 10.1175/JCLI-D-12-00543.1

Chalise, S.R., 2002: Management of water resources for poverty alleviation in the Hindu-Kush Himalayas. In: FRIEND 2002- Regional Hydrology. IAHS Publication **274**, 27–34.

Chan, S. C., E. J. Kendon, H. J. Fowler, S. Blenkinsop, C. A.T. Ferro, and D. B. Stephenson, 2013: Does increasing the spatial resolution of a regional climate model improve the simulated daily precipitation? *Climate Dynamics*, **41**, 1475–1495.

Charron, M., G. Pellerin, L. Spacek, P. L. Houtekamer, N. Gagnon, H. L. Mitchell, and L. Michelin, 2010: Toward Random Sampling of Model Error in the Canadian Ensemble Prediction System. *Monthly Weather Review*, **138**, 1877–1901, doi: 0.1175/2009MWR3187.1

Chen, M., P. Xie, J. E. Janowiak, and P. A. Arkin, 2002: Global Land Precipitation: A 50-yr Monthly Analysis Based on Gauge Observations, *Journal of Hydrometeorology*, **3**, 249-266.

Collier, E., and W. W. Immerzeel: High-resolution modeling of atmospheric dynamics in the Nepalese Himalaya. *Journal of Geophysical Research*, **120**, 9882–9896, https://doi.org/10.1002/2015JD023266, 2015.

Collins, M., K. A. Rao, K. Ashok, S. Bhandari, A. K. Mitra, S. Prakash, R. Srivastava and A. Turner, 2013: Observational challenges in evaluating climate models. *Nature Climate Change*, **3**, 940-941.

Cui, B., Z. Toth, Y. Zhu, and D. Hou, 2012: Bias correction for global ensemble forecast. *Weather and Forecasting*, **27**, 396–410.

Cruz, F. T., and H. Sasaki, 2017: Simulation of present climate over Southeast Asia using the Non-hydrostatic Regional Climate Model. *SOLA*, **13**, 13−18.

Dahal, R. K. and S. Hasegawa, 2008: Representative rainfall thresholds for landslides in the Nepal Himalaya. *Geomorphology*, **100** (3-4), 429–443.

Dahal, R. K. 2012: Rainfall-induced Landslides in Nepal. *International Journal of Japan Erosion Control Engineering*, **5** (1), 1–8.

Das, S., R. Ashrit, and M. W. Moncrieff, 2006: Simulation of a Himalayan cloudburst event. *Journal of Earth System Science*, **115**(3), 299–313.

Dee, D. P., S. M. Uppala, A. J. Simmons, and coauthor, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597, doi: 10.1002/qj.828.

Dobler, A., and B. Ahrens, 2010: Analysis of the Indian summer monsoon system in the regional climate model COSMO-CLM. *Journal of Geophysical Research*, **115**, D16, doi: 10.1029/2009JD013497.

Duan, A., J. Hu and A. Xiao, 2013: The Tibetan Plateau Summer Monsoon in the CMIP5 Simulations. *Journal of Climate*, **26**, 7747-7766, doi: 10.1175/JCLI-D-12-00685.1

Duan, K., T. Yao and L. G. Thompson, 2006: Response of monsoon precipitation in the Himalayas to global warming, *Journal of Geophysical Research*, **111** (D19110), doi:10.1029/2006JD007084

Durai, V. R., and R. Bhardwaj, 2014: Forecasting quantitative rainfall over India using multi-model ensemble technique. *Meteorology and Atmospheric Physics*, **126**, 31–48.

Fennessy, M. J., J. L. Kinter III, B. Kirtman, L. Marx, S. Nigam, E. Schneider, J. Shukla, D. Straus, A. Vernekar, X. Xue, and J. Zhou, 1994: The simulated Indian monsoon: A GCM sensitivity study. *Journal of Climate*, **7**, 33-43.

Feser, F., B. Rockel, H. von Storch, J. Winterfeldt, and M. Zahn, 2011: Regional climate models add value to global model data: a review and selected examples. *Bulletin of the American Meteorological Society*, **92**, 1181–1192, doi: 10.1175/2011BAMS3061.1.

Fritsch, J. M., and Coauthors, 1998: Quantitative precipitation forecasting: Report of the eighth prospectus development team. U.S. Weather Research Program. *Bulletin of the American Meteorological Society*, **79** (2), 285–299.

Fujinami, H., D. Hatsuzuka, T. Yasunari, T. Hayashi, and Coauthor, 2011: Characteristic intraseasonal oscillation of rainfall and its effect on interannual variability over Bangladesh during boreal summer. *International Journal of Climatology*, **31**, 1192–1204.

Gadgil, S. and P. Joseph, 2003: On breaks of the Indian monsoon. *Proceedings of the Indian Academy of Sciences, Earth and Planetary Sciences*, ***112*** (4), 529-558.

Giorgi, F., and M. R. Marinucci, 1996: An Investigation of the Sensitivity of Simulated Precipitation to Model Resolution and Its Implications for Climate Studies. *Monthly Weather Review*, **124**, 148–166.

Giorgi, F., and L. O. Mearns, 1999: Introduction to special section: Regional climate modeling revisited. *Journal of Geophysical Research*, **104** (D6), 6335–6352, doi: 10.1029/98JD02072.

Gofa F., D. Tzeferi, and A. Raspanti, 2013: Using synoptic classification to evaluate an operational weather forecasting system. In: Helmis C., and P. Nastos (eds) Advances in Meteorology, Climatology and Atmospheric Physics. Springer Atmospheric Sciences. Springer, Berlin, Heidelberg pp. 110.

Hamill, T. M., 2012: Verification of TIGGE multimodel and ECMWF reforecast-calibrated probabilistic precipitation forecasts over the contiguous United States. *Monthly Weather Review*, **140**, 2232–2252, doi:10.1175/MWR-D-11-00220.1.

Harris, I., P. D. Jones, T. J. Osborn and D.H. Lister, 2014: Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *International Journal of Climatology*, **34**: 623–642. doi: 10.1002/joc.3711.

Haslinger, K., I. Anders, and M. Hofstätter, 2013: Regional climate modelling over complex terrain: an evaluation study of COSMO-CLM hindcast model runs for the Greater Alpine Region. *Climate Dynamics*, 40, 511–529.

He, X., N. W. Chaney, M. Schleiss, and J. Sheffield, 2016: Spatial downscaling of precipitation using adaptable random forests. *Water Resources Research*, **52**, 8217–8237, doi:10.1002/2016WR019034.

Houze, R. A. Jr., D. C. Wilton, and B. F. Smull, 2007: Monsoon convection in the Himalayan region as seen by the TRMM Precipitation Radar. *Quarterly Journal of the Royal Meteorological Society*, **133**(627), 1389–1411.

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

Ikawa, M. and K. Saito, 1991: Description of a non-ydrostatic model developed at the Forecast Research Department of the MRI. Technical Reports of the MRI, 28, 238

Iorio, J. P., P..B. Duffy, B. Govindasamy, S. L. Thompson, M. Khairoutdinov, and D. Randall, 2004: Effects of model resolution and subgrid-scale physics on the simulation of precipitation in the continental United States. *Climate Dynamics*, **23**, 243–258.

Inoue, T. and H. Ueda, 2011: Delay of the First Transition of Asian Summer Monsoon under Global Warming Condition, *SOLA*, **7**, 081-084, doi:10.2151/sola.2011-021

IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Stocker, T. F., D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley (eds.), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535pp.

Jha, B., Z. Z. Hu and A. Kumar, 2014: SST and ENSO variability and change simulated in historical experiments of CMIP5 models, *Climate Dynamics*, **42** (7), 2113–2124, doi:10.1007/s00382-013-1803-z.

Ji, Z., and S. Kang, 2012: Double-Nested Dynamical Downscaling Experiments over the Tibetan Plateau and Their Projection of Climate Change under Two RCP Scenarios. *Journal of the Atmospheric Sciences*, **70**, 1278-1290, doi: 10.1175/JAS-D-12-0155.1.

Jiao Y., and D.Caya, 2006: An investigation of summer precipitation simulated by the Canadian regional climate model. *Monthly Weather Review* 134: 919–932, doi: 10.1175/MWR3103.1.

Johnson, C., and R. Swinbank, 2009: Medium-range multimodel ensemble combination and calibration. *Quarterly Journal of the Royal Meteorological Society*, **135**, 777–794.

Jones, R. G., J. M. Murphy, and M. Noguer, 1995: Simulation of climate change over Europe using a nested regional climate model. I: assessment of control climate, including sensitivity to location of lateral boundaries, *Quarterly Journal of the Royal Meteorological Society*, **121**, 1413–1449.

Jourdain, N. C., A. S. Gupta, A. S. Taschetto, C. C. Ummenhofer, A. F. Moise and K. Ashok, 2013: The Indo-Australian monsoon and its relationship to ENSO and IOD in reanalysis data and the CMIP3/CMIP5 simulations. *Climate Dynamics*, **41**, 3073-3102, doi:10.1007/s00382-013-1676-1.

Kadel, I., 2012: A Study on Precipitation over the Central Himalaya, Nepal. MSc. thesis, Tohoku University, Japan.

Kain, J. S., and J. M. Fritsch, 1990: A one-dimensional entraining/detraining plume model and its application in convective parameterization. *Journal of the Atmospheric Sciences*, 47, 2784–2802

Kain, J. S., 2004: The Kain-Fritsch Convective Parameterization: An Update. *Journal of Applied Meteorology*, **43**, 170–181

Kalnay, E., S. J. Lord, and R. D McPherson, 1998: Maturity of operational numerical weather prediction: Medium range. *Bulletin of the American Meteorological Society*, **79**, 2753–2769.

Karmacharya, J., R. Jones, W. Moufouma-Okia, and M. New, 2016: Evaluation of the added value of a high-resolution regional climate model simulation of the South Asian summer monsoon climatology. *International Journal of Climatology*, **37**, 3630–3643, https://doi.org/10.1002/joc.4944.

Karki, R., S. Hasson, U. Schickhoff, T. Scholten, and J. Böhner, 2017a: Rising Precipitation Extremes across Nepal. *Climate*, **5**(1), 4, doi:10.3390/cli5010004

Karki, R., S. Sassan, L. Gerlitz, U. Schickhoff, T. Scholten, and J. Böhner, 2017b: Quantifying the added value of convection-permitting climate simulations in complex terrain: a systematic evaluation of WRF over the Himalayas. *Earth System. Dynamics*, **8**, 507–528.

Kipling, Z., C. Primo, and A. Charlton-Perez, 2011: Spatiotemporal behavior of the TIGGE medium-range ensemble forecasts. *Monthly Weather Review*, 139, 2561–2571.

Kitoh, A. and H. Endo, 2016: Changes in precipitation extremes projected by a 20-km mesh global atmospheric model, *Weather and Climate Extremes*, **11**, 41-52.

Knutti, R., D. Masson, and A. Gettelman, 2013: Climate model genealogy: Generation CMIP5 and how we got there. *Geophysical Research Letters*, **40**, 1194–1199.

Kobayashi, S., Y. Ota, Y. Harada, A. Ebita, M. Moriya, H. Onoda, K. Onogi, H. Kamahori, C. Kobayashi, H. Endo, K. Miyaoka, and K. Takahashi , 2015: The JRA-55 Reanalysis: General Specifications and Basic Characteristics. *Journal of the Meteorological Society of Japan*, **93**, 5–48, doi:10.2151/jmsj.2015-001.

Kreft, S., and D. Eckstein, 2013: *Global Climate Risk Index 2014*, Germanwatch e.V., Germany.

Krishnamurti, T. N., C. M. Kishtawal, T. LaRow, D. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil, and S. Surendran, 1999: Improved skills for weather and seasonal climate forecasts from multi-model superensemble. *Science*, **285**, 1548–1550.

Krishnamurti, T., A. D. Sagadevan, A. Chakraborty, A. Mishra, and A. Simon, 2009: Improving multimodel weather forecast of monsoon rain over China using FSU superensemble. *Advances in Atmospheric Sciences*, **26**(5), 813–839, doi:10.1007/s00376-009-8162-z.

Krishnamurthy, V. and J. Shukla, 2000: Intraseasonal and Interannual Variability of Rainfall over India. *Journal of Climate*, **13**, 4366-4377.

Kumar, A., A. K. Mitra, A. K. Bihra, G. R. Iyengar, and V. R. Durai, 2012: Multi-model ensemble (MME) prediction of rainfall using neural networks during monsoon season in India. *Meteorological Applications*, **19**, 161–169.

Kumar, S., V. Merwade, J. L. Kinter III and D. Niyogi, 2013: Evaluation of Temperature and Precipitation Trends and Long-Term Persistence in CMIP5 Twentieth-Century Climate Simulations. *Journal of Climate*, **26**, 4168-4185.

Kulkarni, A., S. Patwardhan, K. Krishna Kumar, K. Ashok, and R. Krishnan, 2013: Projected Climate Change in the Hindu Kush–Himalayan Region By Using the High-resolution Regional Climate Model PRECIS, *Mountain Research and Development*, **33**(2), 142-151.

Lang, T. J. and A. P. Barros, 2002: An Investigation of the Onsets of the 1999 and 2000 Monsoons in Central Nepal. *Monthly Weather Review*, **130**, 1299-1316.

Lee, J. Y. and B. Wang, 2014: Future change of global monsoon in the CMIP5. *Climate Dynamics*, **42** (1), 101-119, doi: 10.1007/s00382-012-1564-0.

Leduc, M. and R. Laprise, 2009: Regional climate model sensitivity to domain size, *Climate Dynamics*, **32**, 833–854 doi:10.1007/s00382-008-0400-z.

Li, X., and M. Ting, 2015: Recent and future changes in the Asian monsoon-ENSO relationship: Natural or forced? *Geophysical Research Letters*, **42**, 3502-3512, doi:10.1002/2015GL063557

Lin, R., T. Zhou and Y. Qian, 2014: Evaluation of Global Monsoon Precipitation Changes based on Five Reanalysis Datasets. *Journal of Climate*, **27**, 1271- 1289, doi: 10.1175/JCLI-D-13-00215.1

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, **20**, 130-141.

Lorenz, E. N., 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus*, **17**, 321-333.

Maharjan, S., and R. P. Regmi, 2015: Numerical prediction of extreme precipitation over a truly complex terrain of Nepal Himalaya. *Journal of Institute of Science and Technology*, **20**(1), 15 – 19.

Matsueda, M., and H. Endo, 2011: Verification of medium-range MJO forecasts with TIGGE. *Geophysical Research Letters*, **38**, L11801, doi:10.1029/2011GL047480.

Matsueda, M., and T. Nakazawa, 2015: Early warning products for severe weather events derived from operational medium-range ensemble forecasts. *Meteorological Applications*, **22**, 213–222.

McBride, J. L. and E. E. Ebert, 2000: Verification of Quantitative Precipitation Forecasts from Operational Numerical Weather Prediction Models over Australia. *Weather and Forecasting*, **15**, 103-121.

Mehran, A., A. AghaKouchak, and T. J. Phillips, 2014: Evaluation of CMIP5 continental precipitation simulations relative to satellite-based gauge-adjusted observations. *Journal of Geophysical Research Atmos.*, **119**, 1695–1707, doi:10.1002/2013JD021152.

Menegoz, M., H. Gallee, and H. W. Jacobi, 2013: Precipitation and snow cover in the Himalaya: from reanalysis to regional climate simulations, *Hydrology and Earth System Sciences*, **17**, 3921-3936, doi:10.5194/hess-17-3921-2013.

Menon, N., 2009: Rainfall Uncertainty and Occupational Choice in Agricultural Households of Rural Nepal. *The Journal of Development Studies*, **45** (6), 864-888, doi: 10.1080/00220380902807387

Menon, A., A. Levermann, J. Schewe, J. Lehmann and K. Frieler, 2013: Consistent increase in Indian monsoon rainfall and its variability across CMIP5 models. *Earth System Dynamics*, **4**, 287-300, doi: 10.5194/esd-4-287-2013

Mirza, M. M. Q., 2011: Climate change, flooding in South Asia and implications. *Regional Environmental Change*, **11**(1), 95-107, doi: 10.1007/s10113-010-0184-7

Mishra, V., 2015: Climatic uncertainty in Himalayan water towers. *Journal of Geophysical Research*, **120**, 2689-2705, doi: 10.1002/2014JD022650

Mishra, V., D. Kumar, A. R. Ganguly, J. Sanjay, M. Mujumdar, R. Krishnan and R. D. Shah, 2014: Reliability of regional and global climate models to simulate precipitation extremes over India, *Journal of Geophysical Research*, **119** (15), 9301–9323, doi: 10.1002/2014JD021636

Molteni, F., R. Buizza, T. N. Pamer, and T. Petroliagis, 1996: The ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, **122**, 73–119.

Murphy, A. H., 1973: A new vector partition of the probabilistic score. *Journal of Applied Meteorology*, **12**, 595–600.

Nakanishi, M., and H. Niino, 2004: An improved Mellor-Yamada level 3 model with condensation physics: Its design and verification. *Boundary-Layer Meteorology*, **112**, 1-31, doi: 10.1023/B:BOUN.0000020164.04146.98

Nayava J. L., 1980. Rainfall in Nepal. *The Himalayan Review*, Nepal Geographical Society, **12**, 1-18.

Ngo-Duc, T., and coauthors, 2017: Performance evaluation of RegCM4 in simulating extreme rainfall and temperature indices over the CORDEX-Southeast Asia region. *International Journal of Climatology*, **37**(3), 1634–1647.

Norris, J., L. M. V. Carvalho, C. Jones, F. Cannon, B. Bookhagen, E. Palazzi, and A. A. Tahir, 2016: The spatiotemporal variability of precipitation over the Himalaya: evaluation of one-year WRF model simulation, *Climate Dynamics*, 1–26, doi:10.1007/s00382-016-3414-y.

Ogata, T., H. Ueda, T. Inoue, M. Hayasaki, A. Yoshida, S. Watanabe, M. Kira, M. Ooshiro and A. Kumai, 2014: Projected Future Changes in the Asian Monsoon: A Comparison of CMIP3 and CMIP5 Model Results. *Journal of the Meteorological Society of Japan*, **92** (3), 207-225, doi:10.2151/jmsj.2014-302

Palazzi, E., J. von Hardenberg, and A. Provenzale, 2013: Precipitation in the Hindu-Kush Karakoram Himalaya: Observations and future scenarios, *Journal of Geophysical Research*, **118**, 85-100, doi: 10.1029/2012JD018697.

Palazzi E., J. von Hardenberg, S. Terzago and A. Provenzale, 2015: Precipitation in the Karakoram-Himalaya: a CMIP5 view. *Climate Dynamics*, **45** (1-2), 21-45, doi: 10.1007/s00382-014-2341-z

Panday, P. K., J. Thibeault and K. E. Frey, 2015: Changing temperature and precipitation extremes in the Hindu Kush-Himalayan region: an analysis of CMIP3 and CMIP5 simulations and projections. *International Journal of Climatology*, **35** (10)**,** 3058-3077. doi: 10.1002/joc.4192

Panthi, J., P. Dahal, M. L. Shrestha, S. Aryal, N. Y. Krakauer, S. M. Pradhanang, T. Lakhankar, A. K. Jha, M. Sharma and R. Karki, 2015: Spatial and Temporal Variability of Rainfall in the Gandaki River Basin of Nepal Himalaya. *Climate*, **3**(1), 210-226. doi:10.3390/cli3010210.

Pope, V. D., and R. A. Stratton, 2002: The processes governing horizontal resolution sensitivity in a climate model. *Climate Dynamics*, **19**, 211–236, doi: 10.1007/s00382-001-0222-8.

Prakash, S., A. K. Mitra, I. M. Momin, E. N. Rajagopal, S. Basu, M. Collins, A. G. Turner, K. Achuta Rao, and K. Ashok, 2015: Seasonal intercomparison of observational rainfall datasets over India during the southwest monsoon season. *International Journal of Climatology*, **35,** 2326-2338. doi:10.1002/joc.4129

Prasanna, V., 2016: Assessment of South Asian Summer Monsoon Simulation in CMIP5-Coupled Climate Models during the Historical Period (1850–2005), *Pure and Applied Geophysics*, **173**, 1379-1402, doi: 10.1007/s00024-015-1126-6

Rajbhandari, R., A. k. Shrestha, S. Nepal and S. Wahid, 2016: Projection of Future Climate over the Koshi River Basin Based on CMIP5 GCMs. *Atmospheric and Climate Sciences*, **6**, 190-204.

Rajeevan, M., S. Gadgil and J. Bhate, 2010: Active and break spells of the Indian summer monsoon, *Journal of Earth System Science*, **119** (3), 229-247.

Rajendran, K., S. Sajani, C. B. Jayasankar and A. Kitoh, 2013: How dependent is climate change projection of Indian summer monsoon rainfall and extreme events on model resolution? *Current science*, **104**(10), 1409-1418.

Ramesh K. V. and P. Goswami P, 2014: Assessing reliability of regional climate projections: the case of Indian monsoon. *Scientific Reports*, **4**, 4071, doi: 10.1038/srep04071.

Rauscher, S.A., E. Coppola, C. Piani, and F. Giorgi, 2010: Resolution effects on regional climate model simulations of seasonal precipitation over Europe. *Climate Dynamics*, **35**, 685–711, doi: 10.1007/s00382-009-0607-7.

Reynolds, R.W, T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M.G. Schlax, 2007: Daily high-resolution-blended analyses for sea surface temperature. *Journal of Climate*, **20**, 5473–5496, doi: 10.1175/2007JCLI1824.1.

Richardson, D. S., 2000: Skill and relative economic value of the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, **126**, 649–667.

Roxy, M. K., K. Ritika, P. Terray, R. Murtugudde, K. Ashok and B. N. Goswami, 2015: Drying of Indian subcontinent by rapid Indian Ocean warming and a weakening land-sea thermal gradient. *Nature Communication*, **6**, 7423. doi: 10.1038/ncomms8423.

Roy Bhowmik, S. K., and V. R. Durai, 2010: Application of multimodel ensemble techniques for real-time district level forecasts in short range time scale over Indian region. *Meteorology and Atmospheric Physics*, **106**, 19–35.

Rummukainen, M., 2010: State-of-the-art with regional climate models, *WIREs Climate Change*, **1**, 82–96.

Rupa Kumar, K., A. K. Sahai, K. Krishna Kumar, S. K. Patwardhan, P. K. Mishra, J. V. Revadekar, K. Kamala and G. B. Pant, 2006: High-resolution climate change scenarios for India for the 21st century. *Current Science*, **90**(3), 334-345.

Sabeerali, C. T., S. A. Rao, A. R. Dhakate, K. Salunke and B. N. Goswami, 2015: Why ensemble mean projection of south Asian monsoon rainfall by CMIP5 models is not reliable? *Climate Dynamics*, **45** (1), 161-174.

Saito, K., J. Ishida, K. Aranami, T. Hara, T. Segawa, M. Narita, and Y. Honda, 2007: Nonhydrostatic Atmospheric Models and Operational Development at JMA. *Journal of the Meteorological Society of Japan*, **85B**, 271–304

Schneider U., A. Becker, P. Finger, A. Meyer-Christoffer, M. Ziese, and B. Rudolf, 2014: GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theoretical and Applied Climatology*, **115**, 15-40, doi:10.1007/s00704-013-0860-x

Sharmila S., S. Joseph, A. K. Sahai, S. Abhilash and R. Chattopadhyay, 2015: Future projection of Indian summer monsoon variability under climate change scenario: An assessment from CMIP5 climate models. *Global and Planetary Change*, **124**, 62–78.

Shrestha, A., 2016: Cloudbursts in the Nepal Himalayas: Interaction between the Indian monsoon and extratropics. PhD thesis, University of Wisconsin-Madison, USA.

Shrestha, A. B., C. P. Wake, J. E. Dibb, P. A. Mayewski 2000: Precipitation fluctuations in the Nepal Himalaya and its vicinity and relationship with some large scale climatological parameters. *International Journal of Climate*, **20**(3), 317–327.

Shrestha, A. B. and R. Aryal, 2011: Climate change in Nepal and its impact on Himalayan glaciers. *Regional Environmental Change*, **11** (1), 65–77.

Shrestha, D., and R. Deshar, 2015: Spatial variations in the diurnal pattern of precipitation over Nepal Himalayas. *Nepal Journal of Science and Technology*, **15**, 57-64.

Shrestha, P., A. P. Dimri, A. Schomburg, and C. Simmer, 2015: Improved understanding of an extreme rainfall event at the Himalayan foothills – a case study using COSMO. *Tellus*, 67, 26031, doi: http://dx.doi.org/10.3402/tellusa.v67.26031.

Shrestha, M. L., 2000: Interannual variation of summer monsoon rainfall over Nepal and its relation to Southern Oscillation Index. *Meteorology and Atmospheric Physics*, **75**, 21-28.

Sigdel, M., and M. Ikeda, 2010: Spatial and temporal analysis of drought in Nepal using standardized precipitation index and its relationship with climate indices. *Journal of Hydrology and Meteorology*, **7**(1), 59-74.

Sigdel, M., and M. Ikeda, 2012: Summer Monsoon Rainfall over Nepal Related with Large-Scale Atmospheric Circulations, *Journal of Earth Science Climate Change*, **3**:112, doi:10.4172/2157-7617.1000112

Smith, T.M., R.W. Reynolds, T.C. Peterson, and J. Lawrimore, 2008: Improvements NOAAs Historical Merged Land–Ocean Temp Analysis (1880–2006). *Journal of Climate*, 21, 2283–2296.

Solman, S. A., E. Sanchez, P. Samuelsson, R. P. da Rocha, L. Li, J. Marengo, N. L. Pessacg, A. R. C. Remedio, S. C. Chou, H. Berbery, H. Le Treut, M. de Castro, and D. Jacob, 2013: Evaluation of an ensemble of regional climate model simulations over South America driven by the ERA-Interim reanalysis: model performance and uncertainties. *Climate Dynamics*, **41**, 1139–1157, doi: 10.1007/s00382-013-1667-2.

Sperber K. R. and T. N. Palmer, 1996: Interannual Tropical Rainfall Variability in General Circulation Model Simulations Associated with the Atmospheric Model Intercomparison Project, *Journal of Climate*, **9,** 2727-2750.

Sperber, K. R., H. Annamalai, I.-S. Kang, A. Kitoh, A. Moise, A. Turner, B. Wang, and T. Zhou, 2013: The Asian summer monsoon: an intercomparison of CMIP5 vs. CMIP3 simulations of the late 20th century. *Climate Dynamics*, **41**, 2711-2744, doi: 10.1007/s00382-012-1607-6.

Su, F., X. Duan, D. Chen, Z. Hao and L. Cuo, 2013: Evaluation of the Global Climate Models in the CMIP5 over the Tibetan Plateau. *Journal of Climate*, **26**, 3187-3208.

Su, X., H. Yuan, Y. Zhu, Y. Luo and Y. Wang, 2014: Evaluation of TIGGE ensemble predictions of Northern Hemisphere summer precipitation during 2008–2012. *Journal of Geophysical Research*, **119**, 7292–7310.

Swinbank, R., and Coauthors, 2016: The TIGGE project and its achievements. *Bulletin of the American Meteorological Society*, **97**, 49–67, doi:10.1175/BAMS-D-13-00191.1

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, **106**, 7183-7192.

Taylor, K.E., R.J. Stouffer and G.A. Meehl, 2012: An Overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, **93**, 485-498, doi:10.1175/BAMS-D-11-00094.1

The HadGEM2 Development Team, 2011: The HadGEM2 family of Met Office Unified Model climate configurations, *Geoscientific Model Development*, **4**, 723-757, doi:10.5194/gmd-4-723-2011

Tong, K., F. Su, D. Yang, L. Zhang and Z. Hao, 2014: Tibetan Plateau precipitation as depicted by gauge observations, reanalyses and satellite retrievals, *International Journal of Climatology*, **34**, 265-285, doi: .1002/joc.3682

Toreti, A., P. Naveau, M. Zampieri, A. Schindler, E. Scoccimarro, E. Xoplaki, H. A. Dijkstra, S. Gualdi and J. Luterbacher, 2013: Projections of global changes in precipitation extremes from Coupled Model Intercomparison Project Phase 5 models. *Geophysical Research Letters*, **40**, 4887-4892, doi:10.1002/grl.50940.

Tracton, M. S., and E. Kalnay, 1993: Operational ensemble prediction at the National Meteorological Center: Practical aspects. *Weather and Forecasting*, **8**, 379–398.

Turner, A. G., P. M. Inness and J. M. Slingo, 2005: The role of the basic state in the ENSO-monsoon relationship and implications for predictability. *Quarterly Journal of the Royal Meteorological Society*, **131**, 781–804.

Turner, A. G. and A. Annamalai, 2012: Climate change and the South Asian summer monsoon. *Nature Climate Change*, **2**, 587–595, doi:10.1038/nclimate1495.

Vannitsem, S., and F. Chomé, 2005: One-way nested regional climate simulations and domain size. *Journal of Climate*, **18,** 229–233, doi: 10.1175/JCLI3252.1.

Wang S.-Y., J.-H. Yoon, R. R. Gillies, and C. Cho, 2013: What Caused the Winter Drought in Western Nepal during Recent Years? *Journal of Climate*, **26**, 8241-8256. doi: http://dx.doi.org/10.1175/JCLI-D-12-00800.1.

Wang, Y., H. Qian, J.-J. Song, and M.-Y. Jiao, 2008: Verification of the T213 global spectral model of China National Meteorology Center over the East-Asia area. *Journal of Geophysical Research*, **113**, D10110, doi:10.1029/2007JD008750.

Warner, T.T., R.A. Peterson, and R.E. Treadon, 1997: A tutorial on lateral boundary conditions as a basic and potentially serious limitation to regional numerical weather prediction. *Bulletin of the American Meteorological Society*, **78,** 2599–2617.

Webster, P. J., V. O. Magana, T. N. Palmer, J. Shukla, R. A. Tomas, M. Yanai and T. Yasunari , 1998: Monsoons: Processes, predictability, and the prospects for prediction. *Journal of Geophysical Research*, **103** (C7), 14451–14510, doi:10.1029/97JC02719

Wehner, M. F., R. L. Smith, G. Bala, and P. Duffy, 2010: The effect of horizontal resolution on simulation of very extreme US precipitation events in a global atmosphere model. *Climate Dynamics*, **34**, 241–247, doi: 10.1007/s00382-009-0656-y.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. 2nd Edition. Academic Press, San Diego, California.

WMO, 2008: Recommendations for the verification and intercomparison of QPFs and PQPFs from operational NWP models. WWRP 2009–1, Revision 2, WMO TD No. 1485, 37pp.

Wu, W., A. H. Lynch, and A. rivers, 2005: Estimating the Uncertainty in a Regional Climate Model Related to Initial and Lateral Boundary Conditions. *Journal of Climate*, **18**, 917-933.

Xu, J., R. E. Grumbine, A. Shrestha, M. Eriksson, X. Yang, Y. Wang, and A. Wilkes, 2009: The Melting Himalayas: Cascading Effects of Climate Change on Water, Biodiversity, and Livelihoods. *Conservation Biology*, **23**(3), 520-530, doi: 10.1111/j.1523-1739.2009.01237.x.

Yasunaga, K., H. Sasaki, Y. Wakazuki, T. Kato, C. Muroi, A. Hashimoto, S. Kanada, K. Kurihara, M. Yoshizaki, and Y. Sato, 2005: Performance of the long-term integrations of the Japan Meteorological Agency nonhydrostatic model with use of the spectral boundary coupling method, *Weather and Forecasting*, **20**, 1061–1072.

Yatagai, A., K. Kamiguchi, O. Arakawa, A. Hamada, N. Yasutomi, and A. Kitoh, 2012: APHRODITE: Constructing a Long-Term Daily Gridded Precipitation Dataset for Asia Based on a Dense Network of Rain Gauges. *Bulletin of the American Meteorological Society*, 93, 1401–1415.

Yuan, H., S. L. Mullen, X. Gao, S. Sorooshian, J. Du, and H. H. Juang, 2005: Verification of probabilistic quantitative precipitation forecasts over the southwest United States during winter 2002/03 by the RSM ensemble system. *Monthly Weather Review*, **133**, 279–294.

Zhong, Z., 2006: A possible cause of a regional climate model's failure in simulating the east Asian summer monsoon. *Geophysical Research Letters*, 33(24), L24707, doi:10.1029/2006GL027654