Statistical Downscaling for Hydro-climatic projections with CMIP5 simulations to assess Impact of Climate Change

Investigators:
Prof. Subimal Ghosh, IIT Bombay
Prof.Arup K Sarma, IIT Guwahati
Prof.Rajib Bhattacharya, IIT Guwahati
Prof.Manish Kumar Goyal, IIT Guwahati
Prof.P.P.Mujumdar, IISC Bangalore
Prof.Vimal Mishra, IIT Gandhinagar
Prof. Ashu Jain, IIT Kanpur
Prof. Shivam Tripathi IIT Kanpur

September 2020

The data obtained from the study is available on NIH Roorkee website. To obtain access to the data, please contact the Member Secretary, INCCC.

The contact details of Member Secretary. INCCC are as follows:

Shri R.P. Pandey,
Member Secretary (INCCC)
National Institute of Hydrology
Roorkee - 247667 (Uttarakhand), India.
Phone : 01332-249216 , 276220, 278478
Fax : +91-1332-272123
Email : inccc.nih@gmail.com, rppandey@gmail.com

Table of Contents

List of Figures

List of Tables

Abstract

Global warming, due to the enhanced greenhouse effect, is likely to have significant effects on the hydrological cycle. Impacts of climate change on hydrological cycle will affect nearly every aspect of human well-being. Assessing the impact of climate change on hydrology, essentially involves projections of climatic variables into local scale hydrologic variables and evaluation of risk of hydrologic extremes in future, for water resources planning and management. General Circulation Models (GCMs) and Earth System Models (ESMs) are tools, designed to simulate time series of climate variables globally for future. The spatial scale, on which these operate, is very coarse compared to that of any hydrologic process of interest. Direct use of the outputs of GCMs in hydrology is not desirable due to this limitations. Downscaling models [dynamical (physics based) and statistical (data driven)] are developed to address the limitations of GCMs by projecting high resolution climate variables, making use of coarse scale GCM simulations. These high resolution climate projections serve as key input not only in planning and management programs but also for obtaining future patterns of extreme events pertaining to different climate variables such as temperature, rainfall etc. These projections can play crucial role for country such as India (characterized by rainfed agriculture and high population density regions) in formulating different strategies regarding water food energy nexus, disaster mitigation planning etc. Hence, the research reported in this report contributes towards the main objective of obtaining high resolution temperature and rainfall projections for India using data driven models. Four methods have been used for the same. They are kernel regression-based approach, two approaches using neural networks and the Bias correction method. All the outputs are further merged and uncertainty modeling has been performed.

Chapter 1

Introduction

Climate is termed as long term statistics of atmospheric variables over a specific region, in contrast to weather which is a state of climate for a shorter time span. Standard averaging period for climate, which is usually considered for a scientific study, is 30 years [Intergovernmental Panel on Climate Change (IPCC), 2001]. Climate comprises of general patterns of weather conditions, seasons and weather extremes like hurricanes, droughts, or rainy periods. Two of the most important climate variables, which show maximum influence on ecology and determine the climate of a particular region, are temperature and precipitation. Climate undergoes changes with time on account of different forcings. The systematic change in the long term statistical properties of climate variable is referred to as 'climate change'. It brings a significant and lasting change in the statistical distribution of weather patterns over longer time scales. It may be a change in average weather conditions or in the distribution of weather around the average conditions. A change in the state of the climate can be identified (e.g. using statistical tests) in terms of changes in the mean and/or variability of its properties that persist for an extended period, typically decades or longer. Climate change may be because of different stimuli such as natural internal processes, external forcing, persistent anthropogenic changes in the composition of the atmosphere or in land use etc. This usage of climate change slightly differs from that in the United Nations Framework Convention on Climate Change (UNFCCC). As per UNFCCC, climate change refers to a change of climate, which is attributed directly or indirectly to human activity that alters the composition of the global atmosphere and it is in addition to natural climate variability, observed over comparable time periods [IPCC, 2007].

During recent years, development in anthropogenic activities such as deforestation, industrialization, burning of fossil fuels etc. has increased emission of greenhouse gases (GHG) (gases which trap outgoing heat radiations). These anthropogenic stimuli, along with natural and other external forcing, are proven to be additional triggers to initiate climate changes. Under the influence of GHG, climate changes are likely to be more intense as compared to those with natural climate variability. The phenomena such as global warming, ozone layer depletion are ominous implications of such changes. It is confirmed that the global warming,

due to the enhanced greenhouse emission, is likely to have significant effects on the hydrological cycle [IPCC, 1996]. One of the most important impacts on society due to future climatic changes will be changes in regional water availability [Buytaert et al., 2009]. IPCC appraises that the hydrological cycle would be intensified, with more evaporation and more precipitation. However, the extra precipitation will be unequally distributed over the globe. Such hydrologic changes will affect nearly every aspect of human well-being such as agricultural productivity, energy use to flood control, municipal and industrial water supply, fish, and wildlife management. The ill-effects of climate change can only be mitigated by designing and implementing effective planning measures at regional level. However, successful implementation of such measures will be highly dependent on falsifiability of future climatic projections at regional level (fine resolution), which serves as one of the key inputs in planning process. As such, there is a strong socio-economic value in predicting the potential effects of climate change affecting the local hydrological processes [e.g. Vergara et al., 2007]. This background serves as the motivation behind undertaking the research problem of apprising effects of climate change on different climate variables and impacts assessment. The present chapter provides brief information on the scientific aspects of climate, climate change, history of developments of climate models, future projections, extremes, research gaps, objectives of present study, and organization of the report.

1.1. Climate Forcing Mechanisms

Factors that shape the climate are called as 'Climate forcing mechanisms'. Such mechanisms are classified into (1) Internal forcing and (2) External forcing. Natural changes in the components of earth's climate system and their interactions are the causes of internal climate variability, or "Internal forcing". Ocean variability is a key component of internal forcing. Short-term fluctuations (years to a few decades) such as the El Niño-Southern Oscillation, the Pacific decadal oscillation, the North Atlantic oscillation, and the Arctic oscillation, represent climate variability rather than climate change. "External forcing" involves changes in solar irradiance or anthropogenic GHG emissions [IPCC, 2007]. The phenomenon of global warming, which consists of unequivocal and continuing rise in the average temperature of Earth's climate system is one of the examples of climate change because of GHG emissions [IPCC, 2007]. Most of the observed increase in global average temperatures since the mid-20th century is very likely due to the observed increase in anthropogenic GHG concentrations [IPCC, 2007]. The consequences of global warming are reflected in global as well as regional

climate in terms of changes in frequency, intensity, and duration of key climatic variables such as precipitation and atmospheric moisture, snow cover, extent of land and sea ice, sea level, and patterns in atmospheric and ocean circulation. Figure 1.1 depicts the variation of the global temperature over a period of 140 years. The plot also shows the average increase of 0.74º C in recent years.

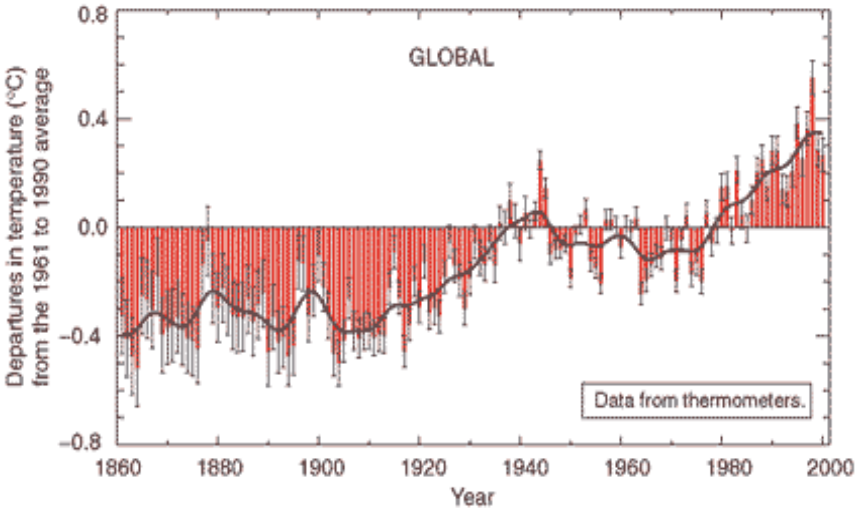


Figure 1.1: *Variation of the Earth's surface temperature for the past 140 years (IPCC 2007)*

(Red bars: Earth's surface temperature departures year by year, Black line: Average temperature, Black bars: Bias in temperature)

Therefore, study of climate change is absolutely necessary to understand its impacts on hydrological processes. Water resources are inextricably linked with climate. Global climatic change has serious implications over water resources and regional development [IPCC, 2001]. Increased evaporation (resulting from higher temperatures), combined with regional changes in precipitation characteristics (e.g., total amount, variability, and frequency of extremes), has the potential to affect mean runoff, frequency and intensity of floods and droughts, soil moisture, and water supplies for irrigation and hydroelectric generation [IPCC, 2001]. Hence, assessing the impact of climate change on hydrology would involves a) projections of climatic variables (e.g., temperature, humidity, mean sea level pressure etc) at a global scale b) downscaling of global scale climatic variables into local scale hydrologic variables and c) computations of risk of hydrologic extremes in future for water resources planning and management.

1.2. General Circulation Models

The history of climate modelling has its origin deeply rooted in past where it began with conceptual models, followed in the 19th century by mathematical models of energy balance and radiative transfer, as well as simple analog models. Advent of computer brought a revolution in the field of climate modelling. Since the 1950s, the principal tools of climate science have been computer simulation models of the global general circulation. From the 1990s to the present, a trend toward increasingly comprehensive coupled models of the entire climate system has dominated the field.

A General Circulation Model (GCM) is a mathematical model of the general circulation of a planetary atmosphere or ocean and based on the Navier–Stokes equations on a rotating sphere with thermodynamic terms for various energy sources (radiation, latent heat). These equations are the basis for complex computer programs, commonly used for simulating the atmosphere or ocean of the Earth. GCMs are widely applied for weather forecasting, understanding the climate, and projecting climate change. These computationally intensive numerical models are based on the integration of a variety of fluid dynamical, chemical, and sometimes biological equations. GCMs have been developed to simulate the present climate and have been used to project the change in future climate. The GCMs are broadly classified into atmospheric and oceanic GCMs (AGCM and OGCM), which are key components of GCMs. The AGCMs and OGCMs are coupled together to form atmosphere-ocean coupled general circulation model (AOGCM). With the addition of other components, such as sea ice and land-surface components, the AOGCM becomes the basis for a full climate model.

GCMs are the present day most credible tools in simulating the global climate systems due to increased level of greenhouse-gas concentration, and they provide current and future time series of climate variables for the entire globe [Prudhomme et al., 2002]. While GCMs demonstrate significant skill at the continental and hemispheric spatial scales and incorporate a large proportion of the complexity of the global system, they are inherently unable to represent local sub grid-scale features and dynamics [Wigley et al., 1990; Carter et al., 1994] especially fail to reproduce non-smooth fields such as precipitation [Hughes and Guttorp, 1994a, b]. Hence, while the impact of greenhouse gases on large-scale atmospheric circulation is well understood, regional changes in the hydrological cycle are far more uncertain in GCM simulations. The conflict between GCM performance at regional spatial scales and the needs of the regional-scale impact assessment is largely related to model resolution. While GCM

accuracy decreases at increasingly finer spatial scales, the needs of impacts assessment research conversely increase with higher-resolution [Hostetler, 1994]. To circumvent these problems, tools for generating high-resolution meteorological inputs are required for modelling hydrological processes. 'Downscaling' approaches have subsequently emerged as a means to bridge the gap between the large-scale atmospheric predictor variables and the local or station-scale meteorological series of interest.

1.3. Coupled Model Intercomparison Project (CMIP)

The Coupled Model Intercomparison Project (CMIP) was launched in late 1995 by the Climate Variability and Predictability (CLIVAR) Numerical Experimentation Group 2 [(NEG2) subsequently reconstituted as the World Climate Research Programme (WCRP)/CLIVAR Working Group on Coupled Models (WGCM)]. It is a framework for global coupled ocean-atmosphere general circulation models. CMIP began in 1995 by collecting output from model "control runs" in which climate forcing is held constant. Later versions of CMIP have collected output from an idealized scenario of global warming, with atmospheric CO_2 increasing at the rate of 1% per year until it doubles at about Year 70. CMIP output is available for study by approved diagnostic sub-projects. More recent phases of the project include more realistic scenarios of climate forcing for historical, paleoclimate and future.

CMIP3 is the third phase of CMIP, where Program for Climate Model Diagnosis and Intercomparison (PCMDI) volunteered to collect model output contributed by leading modeling centers around the world. Climate model output from simulations of the past, present and future climate was collected by PCMDI mostly during the years 2005 and 2006, and this archived data constitutes phase 3 of the Coupled Model Intercomparison Project (CMIP3). In part, the WGCM organized this activity to enable those outside the major modeling centers to perform research of relevance to climate scientists preparing the Fourth Asssessment Report (AR4) of IPCC. The report consists of different scenarios described in the IPCC Special Report on Emissions Scenarios [SRES, 2000]. The SRES scenarios are grouped into four scenario families (A1, A2, B1 and B2) that explore alternative development pathways, covering a wide range of demographic, economic and technological driving forces and resulting GHG emissions. The emissions projections are widely used in the assessments of future climate change, and their underlying assumptions with respect to socio-economic, demographic and technological change serve as inputs to many recent climate change vulnerability and impact

assessments. The A1 scenario storyline assumes a world of very rapid economic growth, a global population that peaks in mid-century and rapid introduction of new and more efficient technologies. A1 scenario is divided into three groups that describe alternative directions of technological change: fossil intensive (A1FI), non-fossil energy resources (A1T) and a balance across all sources (A1B). B1 scenario describes a convergent world, with the same global population as A1 scenario, but with more rapid changes in economic structures toward a service and information economy. B2 scenario describes a world with intermediate population and economic growth, emphasizing local solutions to economic, social, and environmental sustainability. A2 scenario describes a very heterogeneous world with high population growth, slow economic development and slow technological change. No likelihood has been attached to any of the SRES scenarios. The basis for CMIP3 scenarios lies in the carbon dioxide concentration at the end of 21st century.

CMIP5 [Taylor et al, 2012] are latest simulations, which have undergone several improvements in terms of physics and resolution as compared to CMIP3. CMIP5 promotes a standard set of model simulations in order to (1) evaluate how realistic the models are in simulating the recent past, (2) provide projections of future climate change on two time scales, near term (out to about 2035) and long term (out to 2100 and beyond), and (3) understand some of the factors responsible for differences in model projections, including quantifying some key feedbacks such as those involving clouds and the carbon cycle. CMIP5 simulations are mainly characterized by four emission scenarios known as 'Representative Concentration Pathways' (RCPs). These start from year 2006 and continue through the end of 2300.

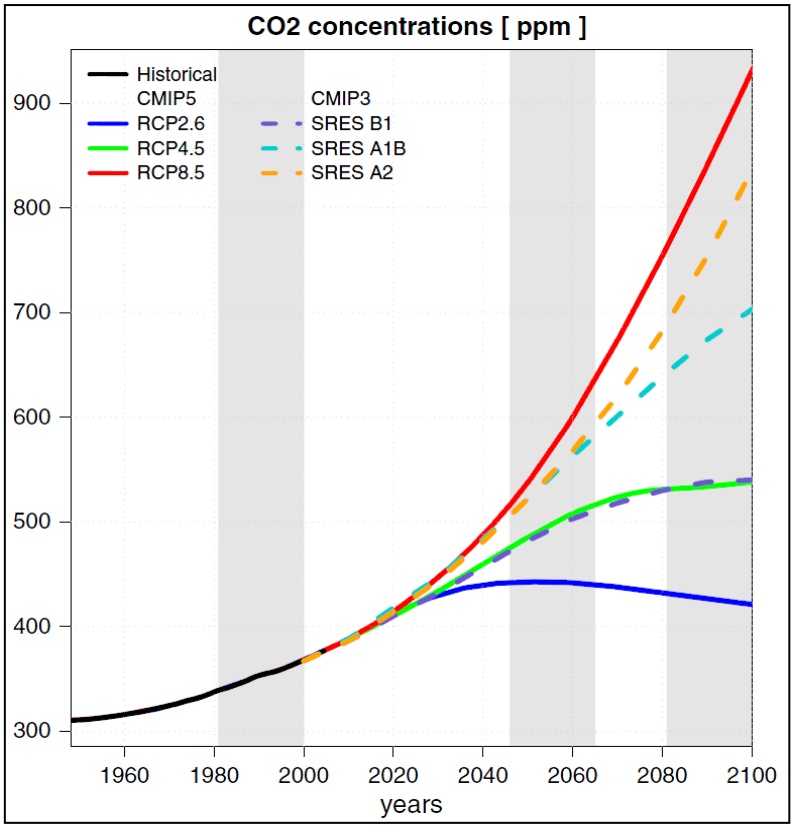


Figure 1.2: *Comparison of Carbon dioxide (CO_2) concentrations in ppm as used in the CMIP3 and CMIP5 historical and scenario simulations. [Sillmann et al, 2013]. The vertical shading indicates the reference period (1981–2000) and the two 20 year periods (2046–2065 and 2081–2100) considered in the analysis of future climate change*

The basis for RCP is approximate radiative forcing at year 2100 relative to pre-industrial conditions. Four scenarios that are termed as RCPs are (1) RCP2.6 (radiative forcing peaks at ~2.6 Wm-2 near 2100); (2) RCP4.5 (radiative forcing stabilizes at ~4.5 Wm-2 after 2100); (3) RCP6.0 (radiative forcing stabilizes at ~6 Wm-2 after 2100); (4) RCP8.5 (radiative forcing reaches at ~8.5 Wm-2 after 2100). Figure 1.2 shows comparison of CMIP3 and CMIP5 scenarios based on CO_2 concentrations variations. It is important to note that SRESB1 is the only scenario from CMIP3, which is comparable with RCP4.5 scenario from CMIP5, which is evident from overlapping time series. As compared to CMIP3 scenarios, CMIP5 scenarios show more spread in terms of CO_2 content at the end of 21st century

1.4. Climate Projections using Downscaling Techniques

The GCMs show different skill scores in simulations and projections of different climatic variables. The variables which show comparatively higher degree of spatial uniformity e.g. temperature or wind variables, are well simulated by GCMs. However, the variables like rainfall, which are highly affected by local parameters, are poorly simulated by GCMs. This is mainly because of the spatial resolution at which GCMs work. Also, for impacts assessment, high resolution projections of rainfall are required which GCMs cannot provide. GCMs have shown tremendous improvements over resolution from CMIP3 to CMIP5 in terms of resolution, underlying physics, and number of ensembles [Knutti and Sedláček, 2012] and in large scale drivers of precipitation [Ryu and Hayhoe, 2013]. However, simulations are not yet up to 'assessment-relevant' resolution [Knutti and Sedláček, 2012; Kumar et al., 2014]. Hence, downscaling techniques are deployed to obtained climate projections at high resolution. These techniques involved obtaining fine resolution projections either by developing physics based Regional Climate Model (RCM) which takes inputs from GCMs or by establishing statistical relationship between coarse scale climate variables (predictors) which are relatively well simulated by GCMs and local scale rainfall (predictand). First approach is known as 'Dynamical Downscaling' and the later approach is known as 'Statistical Downscaling'. Both the approaches have their pros and cons. Dynamical downscaling can be applied to obtain very high resolution projections; however, these are time consuming and require heavy computational facility.

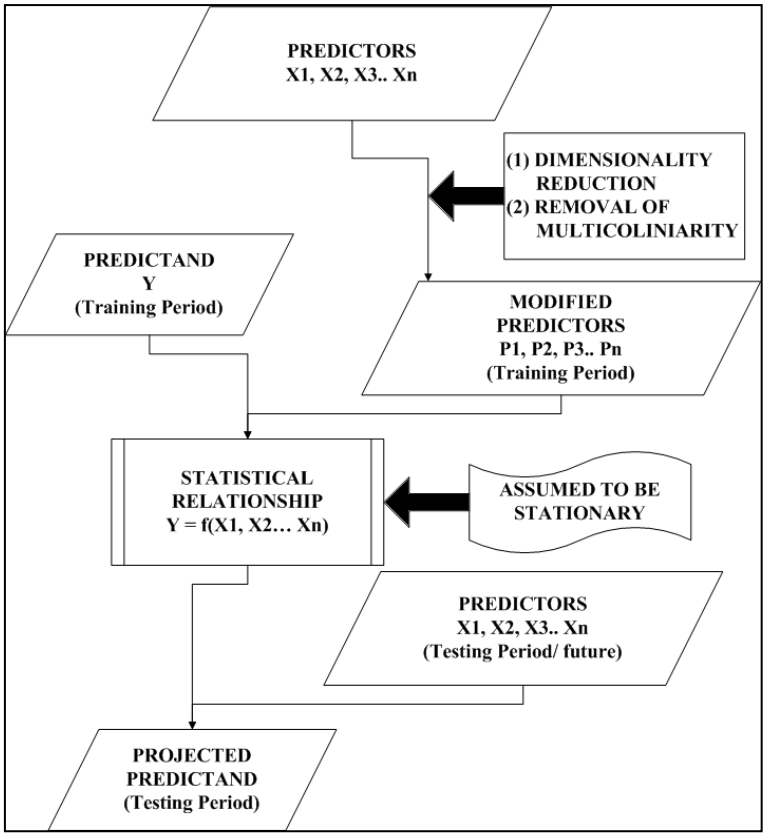


Figure 1.3: *Flowchart of Linear Regression based statistical downscaling methodology*

On the other hand, statistical downscaling techniques are very fast in obtaining projections, although, these techniques demand observed data availability over longer time scale. Also, the future projections are obtained using the same statistical relationship, which is established over past observed data. The credibility of statistical relationship, which is established over past data, is questioned under changing climatic conditions.

The statistical relationship between predictors and predictand can be as simple as linear regression to more complicated models. Predictand is the climate variables of interest at fine resolution and predictors are climate variables at coarse scale that influence predictand. Mathematically, predictors are independent variables $X= (X_1, X_2..)^T$, predictand is dependent variable (Y), and relationship is the function which links X and Y. Figure 1.3 shows the flowchart of the most basic, single site (at a single location), linear regression based downscaling methodology. The set of predictors $X= (X_1, X_2..)^T$ first undergoes mathematical operations which involve reduction of dimension and removal correlation among predictors. These are required for (1) reducing computational power and (2) fulfilling the assumption of

regression that the predictors are not correlated with each other, respectively. Let X' = (X_1', X_2'..) be the modified predictors (after mathematical operations). These form statistical relationship with the predictand 'Y'. This statistical relationship is assumed to be time invariant. Using future predictors that are simulated by GCMs and the established relationship (based on past observed data), future projections of predictand are obtained. The assumed stationarity in the relationship is major limitation of all data driven models.

Chapter 2

Statistical Downscaling with Kernel Regression

2.1. Introduction

Precipitation is any product of the condensation of atmospheric water vapor that falls under gravity in the form of drizzle, rain, sleet, snow and hail. It is one of the most important variables which are studied under the hydrological and meteorological sciences. It is also one of the most difficult variables to simulate mathematically because of complex geophysical processes involved at different spatial resolution, right from synoptic to regional/local scale. Indian Summer Monsoon Rainfall (ISMR) is an appropriate example to illustrate the complexity involved in understanding monsoon system. ISMR gets affected by teleconnections such as the Indian Ocean Dipole (IOD) and the El Niño/Southern Oscillation (ENSO), which have complementarily effect on the ISMR [Ashok et al., 2001]. Also spatial distribution of observed rainfall over India [Rajeevan and Bhate, 2008] shows high influence of local factors such as topography on ISMR. Hence, in order to simulate complex phenomenon such as ISMR, the model should have the capability to capture the large scale circulation as well as local scale parameter like topography, cloud formation etc. It is difficult to find both the qualities in a single GCM mainly because of the resolution at which it simulates the rainfall. Lack of incorporation of subgrid features such as topography, land use, and cloud physics bring down the skills of GCM to simulate ISMR [Wigley et al, 1990]. Secondly, coarse resolution simulations of rainfall by GCMs cannot be used for regional studies or impacts assessment. This is because regional studies/ impacts assessment require fine resolution rainfall data. Despite of these problems, it is necessary to perceive the future of ISMR as it plays a vital role in Indian economy. Indian economy is dominated by agriculture industry. India mainly receives rainfall in the month June, July, August, and September. The rainfall, received in these four months is the prime source of water, which is utilized over entire year. In this case, 'planning and management of water resources' is imperative for optimum usage of water. As ISMR is a major factor controlling water resources, agriculture, and ecosystems throughout India, it is one of the key inputs for climate change adaptation and policy making, accurate projections of ISMR are of paramount importance. Fine resolution rainfall projections for future serves as an essential input for planning purpose, which GCMs fail to provide. 'Downscaling' i.e. process

of obtaining fine resolution projections of climate variables, can be used to overcome the shortcomings of GCMs in simulating ISMR at fine resolution. Dynamical and statistical downscaling are two ways of obtaining fine resolution projections using GCM simulations. Some of the details about these methods are already discussed in previous chapters. Dynamical downscaling nests a Regional Climate Model (RCM) into the GCM to represent the atmospheric physics with a higher grid box resolution and local influencing factors (such as topography, land cover, cloud physics) within a limited area of interest. Using different RCMs such as 'Providing Regional Climates for Impacts Studies' (PRECIS) [Rupakumar et al., 2006; Krishnakumar et al., 2011], the 'RegCM3' [Moetasim et al., 2009], and the global atmosphere–ocean model 'European Center from Hamburg' (ECHAM5) as the driving model and the 'COSMO-CLM' as the regional model [Dobler and Ahrens, 2011]; dynamic downscaling techniques have also been applied for the projection of ISMR. Regional climate models have the advantage of capturing and explaining geophysical phenomena but are computationally expensive. On the other hand, statistical downscaling relies on data driven approaches that involve deriving empirical relationships that transform the large-scale features of GCM simulated climate variables (predictors) into regional-scale variables (predictand) such as rainfall. Statistical downscaling methods are computationally inexpensive and are useful if sufficient historical data is available for generating probability distribution functions and establishing statistical relationships.

Precipitation is a highly heterogeneous spatial phenomenon that normally occurs as a result of the complex interaction between various climatic variables, the ocean, and the landmass. Therefore, the spatial modelling of precipitation is a major research challenge. In this work, rainfall projections are obtained using the approach introduced by Kannan and Ghosh [2013]. This chapter provides complete details on selection of predictors, predictand, downscaling methodology, its application over India, and projections results.

2.2. Study Region and Data used

The entire exercise of obtaining rainfall projections using statistical downscaling is carried out on Indian landmass at 0.25 degree resolution. India mainly receives rainfall because of encounter of South Asian Summer Monsoon winds with different highlands. However, variations in local factors such as topography, land cover etc. result in spatially complex rainfall pattern. The complexity is in terms of variations in mean rainfall, encountered over a small

distance. This is mainly seen over hilly regions, windward side receives high rainfall whereas leeward side receives significantly low rainfall. Such complexities are not captured by GCMs and hence, different downscaling techniques are required to be deployed for getting high resolution rainfall projections.

2.2.1. Selection of Predictors

Predictors are climatic variables that are well simulated by GCM**s** and that are used in statistical downscaling for simulating local scale hydro-meteorological variables of interest such as rainfall (predictand)**.** Ideally, the geophysical processes associated with rainfall are of a fine resolution. Therefore, coarse gridded GCMs fail to simulate them well. Predictors that directly affect rainfall processes are used in statistical downscaling as input variables. The choice of predictor variables is of the utmost importance when it comes to the accuracy of projected data. The selection of predictors should be dependent on the following criteria: (1) the data for the particular predictor should be available for the desired period; (2) the selected GCM should be capable of simulating the variable well; and (3) the predictor should show a good correlation with the predictand [Wilby et al., 1999]. Hewitson and Crane **[**1996] demonstrated how the downscaled projection of future change in mean rainfall and extreme events may significantly alter depending on the selection of predictor. Downscaled results can also depend on whether or not absolute or relative humidity is used as a predictor [Charles et al., 1999b]. The implication here is that while a predictor may or may not appear to be the most significant when developing the downscaling function for present climates, changes in that predictor for a future climate may be critical for determining climate change. Some estimation procedures, for example the stepwise regression, are not capable of recognizing the changes and exclude variables that may be vital for climate change. Sharma **[**2000a] adopted Partial Mutual Information (PMI) criterion in order to identify significant season-wise atmospheric predictors. Johnson and Sharma **[**2009] developed a new approach based on a Variable Convergence Score (VCS). Using this approach, variables are ranked based on the coefficient of variation for the ensemble. Dobler and Ahrens **[**2011] employed u-wind and v-wind in the form of wind shear at higher altitudes as predictors. Wind shear is basically the difference between the wind velocities at the 850 hPa and 200 hPa pressure levels. For the current study, the climatic variables described by Kannan and Ghosh **[**2011] are used as predictors viz. temperature, pressure, specific humidity, u-wind, and v-wind at the surface. The goal of present work is to project rainfall at a very high resolution for the entire 21st century. Pressure level variables are

not considered in this study as such they are not available for the entire period of 100 years of model simulation. However, rainfall projections using wind shears as additional predictors are performed for a single run in order to check whether or not the statistical downscaling technique provided any additional advantages in terms of computational power or accuracy over projections obtained using surface level variables.

2.2.2. Source of Predictors (Reanalysis Data)

The reanalysis project is an outgrowth of the Climate Data Assimilation System (CDAS) project undertaken by the National Center for Environmental Prediction/ National Center for Atmospheric Research (NCEP/NCAR). A frozen state-of-the-art analysis/forecast system is used to perform the data assimilation using past data [Kalnay et al., 1996]. The components of the assimilated datasets are the following: (1) global rawinsonde data, (2) a Comprehensive Ocean-Atmosphere Data Set (COADS) that comprises a collection of surface marine data, (3) aircraft data, (4) surface land synoptic data, (5) satellite sounder data, (6) Special Sensing Microwave/Imager (SSM/I) surface wind speeds, and (7) satellite cloud drift winds. The spectral statistical interpolation, a three dimensional variational analysis scheme, is used as an analysis module to assimilate the various data sets. The results are combined in order to obtain NCEP/NCAR reanalysis data for various climate variables. Climatic variables are categorized into three types [Kalnay et al., 1996]. Category 'A' variables are those that are strongly influenced by observations (e.g. zonal and meridional wind). Category ‘A’ variables are highly accurate. Category 'B' variables are influenced by the model and the observations, but are not as accurate as Category 'A' level variables (e.g. specific humidity). For Category 'C' variables, no observations directly affect the variable and they are completely generated by the model (e.g. precipitation flux). Additional information on the various types of NCEP/NCAR reanalysis variables can be obtained from [Kalnay et al., 1996]. For our work, the NCEP/NCAR reanalysis-I daily data for surface air temperature, mean sea level pressure, specific humidity, zonal wind velocity, and meridional wind velocity for a region delimited by the latitudes 5°-40°N and longitudes 60°-120°E, surrounding the entire study area for a period of 30 years from 1971-2000,is utilized for the bias correction, the training, and the validation of the downscaling model.

The data was obtained for the important climatic variables from five GCMs (CMIP5) for the emission scenarios (SRESs) prescribed by the Intergovernmental Panel on Climate Change

(IPCC). The following Global Circulation Models (GCMs) were selected for the current analysis:

1. CCCMA CanESM2;
2. CNRM CM5;
3. MPI ESM MR;
4. MPI ESM LR and;
5. BNU ESM

From each GCM, following predictors were taken for rainfall projection:

1. Near-surface Air Temperature (TAS);
2. Air Temperature at 850hpa pressure level (T850);
3. Air Temperature at 500hpa pressure level (T500);
4. Eastward Near-surface Wind Velocity (UAS);
5. Eastward Wind Velocity at 850hpa pressure level (U850);
6. Northward Near-surface Wind Velocity (VAS);
7. Northward Wind Velocity at 850hpa pressure level (V850);
8. Specific Humidity at 850hpa pressure level (Q850);
9. Sea level Air Pressure (PSL); and
10. Geo-potential Height at 500hpa pressure level

More details about the original CMIP5 dataset can be found at http://cmippcmdi.llnl.gov/cmip5/.

Selection of GCMs is important for a specific application. The above mentioned GCMs and predictors were selected for the rainfall projection because of the following reasons:

1. All the GCMs listed above provide data for the desired analysis period that is a historical period (19512005) and a future period (2006-2100 for RCP4.5 and RCP8.5 climate scenarios)
2. All the GCMs listed above can simulate all the predictors listed above

3. The predictors listed above have good correlation with the predictand (i.e. rainfall)

2.3. Kernel Regression based statistical downscaling methodology

Statistical downscaling is the methodology by which coarse resolution predictors are linked to the fine resolution predictand using a statistical relationship. For the current study, the methodology developed by Kannan and Ghosh [2013] is applied with required modifications. Figure 2.1 provides a flowchart depicting the stepwise mathematical operations performed on the predictors and predictand for the rainfall projections. The GCM simulated predictors and the observed rainfall, as a predictand, undergo different mathematical operations before actually becoming statistically linked. The predictors undergo a bias correction operation where the systematic error is removed using a quantile based remapping technique [Li et al., 2010]. The bias corrected predictors go through a principal component analysis (PCA) that involves the application of orthogonal transformation on a set of correlated predictor variables, producing principal components. The resulting principal components are dimensionally reduced and uncorrelated to one another. Principal components carry almost the same variability as that of the original data. Hence, the PCA helps to reduce both dimensionality and multicollinearity. A reduction in the dimensions also results in a reduction in the computational effort.

A K-means clustering technique is applied in order to individually derive **t**he daily rainfall states for the seven Indian zones**.** The step helped to reserve cross-correlation amongst rainfall for multiple grids in one zone. The daily rainfall states and the bias corrected predictors, which all undergo principal component analysis, are key inputs to the kernel regression model for establishing the statistical relationship for the training period. Assuming that the relationship holds for the future, future states can be generated with the help of pre-established relationships and predictors for the future period. By applying a nonparametric kernel regression, rainfall is projected at each node.

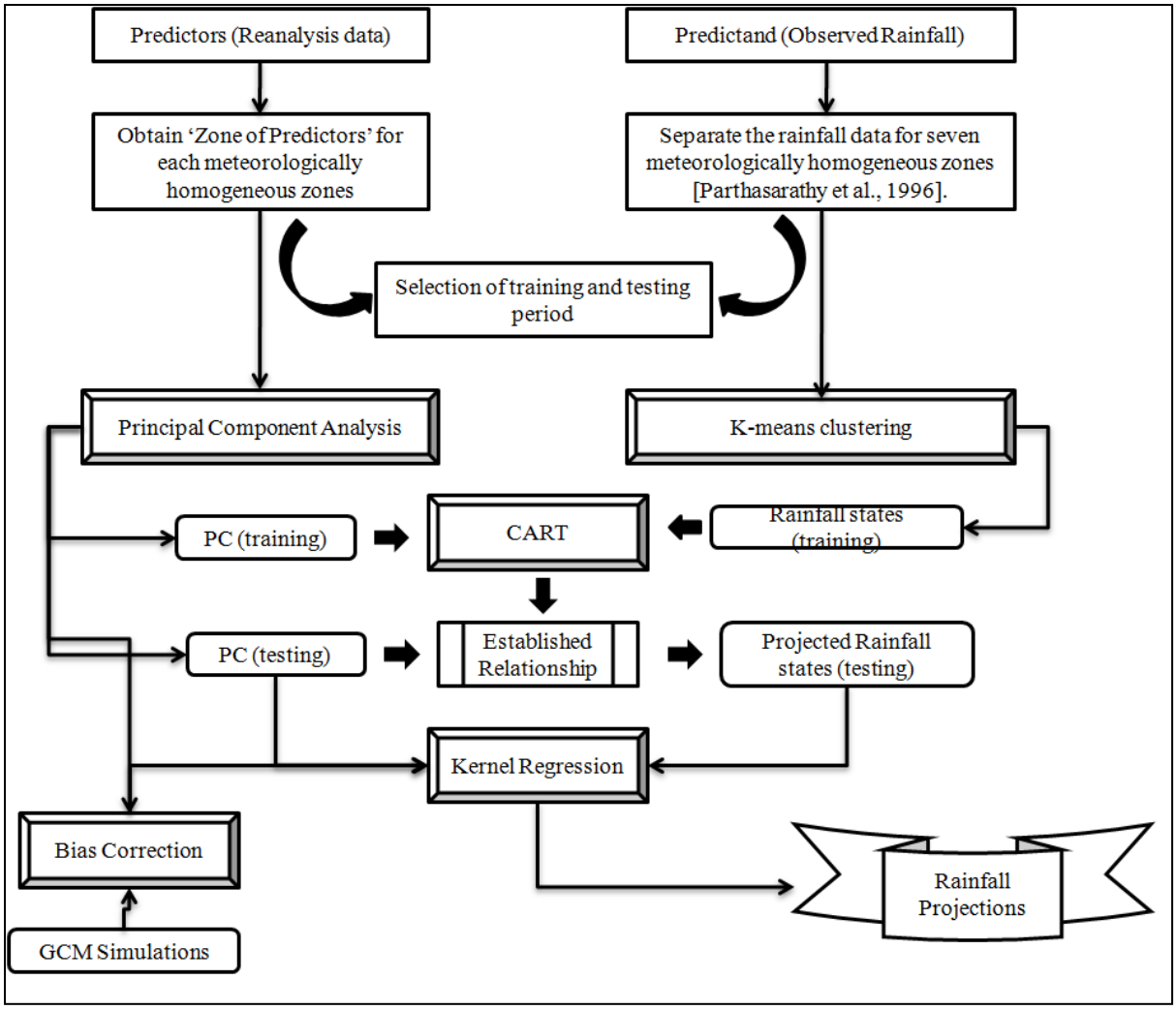


***Figure 2.1**: A flowchart for the Multisite Statistical Downscaling Model. Flowchart enlists various mathematical operations that are performed on predictors (the GCM simulated climate variables) and the predictand (rainfall) which take part in statistical downscaling as inputs. The current statistical downscaling model is a combination of the daily weather state generator and the transfer function method.*

Each of the mathematical operations discussed above are illustrated in the flowchart in Figure 2.1 are summarized in the subsequent text.

Due to incomplete knowledge regarding geophysical processes, assumptions are made in the development of the GCM in terms of the parameterizations and the empirical formula. As a result of these assumptions, the GCM may not simulate climate variables accurately and there is a systematic difference, known as bias, between the observed and simulated climate variables

for almost all of the GCMs. Different techniques of bias correction are already discussed in previous chapter.

The bias correction methodology developed by Li et al. [2010] demands that the reference data (observed data/reanalysis data) time series and the GCM simulated variable time series be at the same location in terms of latitude and longitude**,** so that their cumulative distribution functions (CDFs) can be compared. The GCM simulated data used for this study is at a 3.75º × 3.75º (lat × long) resolution. The NCEP/NCAR reanalysis data is at a 2.5º × 2.5º (lat × long) resolution. Therefore, using the bilinear interpolation technique, first the GCM simulated data is scaled to the NCEP/NCAR resolution, then the bias correction is performed**.**

Use of high dimensional correlated data in the present downscaling framework is computationally expensive, and the high correlation between the variables may result in multicollinearity (high correlation among the explanatory variables). Hence, there is a need to convert the highly correlated multidimensional predictor into a set of uncorrelated variables with reduced dimensionality. Principal component analysis (PCA) is used for this purpose. The following section presents the methodology and algorithm used in PCA.

Bias corrected data cannot be directly used for simulations since it suffers from two major impediments viz. multidimensionality and multicollinearity**.** The multidimensionality problem, normally referred to as the Curse of Dimensionality, arises while dealing with high dimensional space. The problem leads to an increase in the degree of sparseness in data that may impact outputs where a statistical significance is required. Multidimensionality also causes a drastic increase in the computational time required for the statistical analysis (e.g. multivariate regression). On the other hand, if the dimensions are reduced without considering the internal data pattern and its variability, the accuracy of the model output will get hampered. The identification of the pattern of multidimensional data is a tough task, especially when the comfort of visual representations is not available. A 'trade off' always exists between multi-dimensions, computational power, and accuracy. Multicollinearity indicates the presence of a high correlation among two or more predictor variables in cases of multiple regression analyses, and may lead to erratic changes in the coefficient estimation for small changes in data. Therefore, it is necessary to identify these problems and to tackle them before the predictors actually get involved in the statistical downscaling. PCA is a powerful mathematical tool that is mainly used for identifying patterns in multidimensional data and reducing the number of dimensions without reducing the variability of the original data to the best possible

extent. For the current study, PCA is used to obtain principal components from the large scale climate variables acting as predictors in the statistical relationship.

The following steps are used to obtain the principal components of the predictor variables:

(1) Standardize the predictors with training period mean and standard deviation (n).

(2) Compute the covariance matrix, τ, of the standardized variables.

(3) Compute the eigen values (λ) and eigen vectors from the covariance matrix, τ.

(4) Arrange the eigen vectors in such a way that the first eigen vector corresponds to the largest eigen value and in general the k^{th} eigen vector (e_{ki}) to the k^{th} largest eigen value λ_k. The k^{th} principal component in time t (pc_{kt}) is computed as:

$$pc_{kt} = \sum_i e_{ki}\, n_t^i \qquad \ldots (2.1)$$

(5) The percentage of total variance ωk explained by the kth principal component is given by:

$$\omega_k = \frac{\lambda_k}{\sum_{m=1}^{M} \lambda_m} \times 100 \qquad \ldots (2.2)$$

where, M is the dimensionality of the original data set

There is not a single clear criterion that can be used to choose the number of principal components that are best retained in a given circumstance especially for atmospheric data concerned. Many principal component selection rules are available in literature [e.g. Preisendorfer, 1988]. One of the selection rule, namely, Kaiser's rule [Jolliffe, 1972] suggest a method to retain the principal components accounting for more than the average amount of the total variance explained. Jolliffe [2002] suggests that $70\% \leq \omega_{crit} \leq 90\%$ may often be a reasonable range in order to be considered sufficient. In this study, number of principal components which explain 98% of variability of original data, are retained.

The meteorological homogeneous zones identified by the IMD are shown in Figure 2.3a [Parthasarathy et al., 1996]. Each zone is treated as an individual entity, and for each zone a corresponding region is fixed that is regular in shape (rectangle or square) and large enough to completely encompass the zone**.** Rainfall in a particular zone is assumed to be influenced more by the predictors in the selected region surrounding the zone than the predictors that are outside

the region. The regions are fixed after examining the correlation contour plots of the coefficient of correlation between the area averaged observed rainfall time series in the given zone and the individual predictor time series available at a 2.5° resolution after the bias correction, as seen in Figure 2.2(c-g). The validity of the assumption is depicted in terms of the correlation plot.

Stepwise details of the PCA are explained using the central zone as an example. Figure 2.3b indicates the central zone as well as the region of predictors for the central zone. In the central zone, there are 64 nodes in the region of the predictors (Figure 2.2b). At each node there are five predictors, implying that for rainfall projections in the central zone, number of variables available as predictors 64×5 = 320. Considering all of the 320 predictors for projecting rainfall in the central zone will lead to the problems of multicollinearity and multidimensionality. Therefore, PCA is applied in order to negotiate these problems.

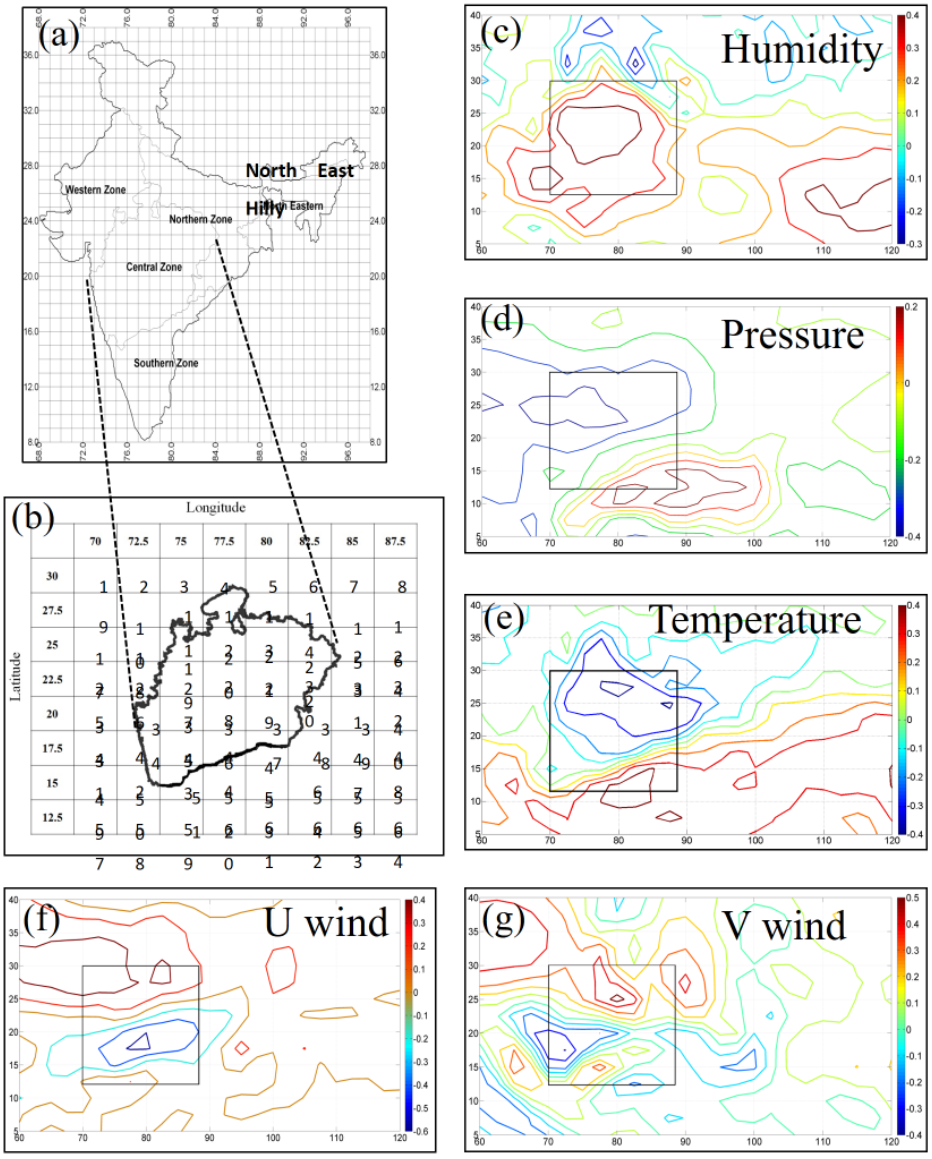
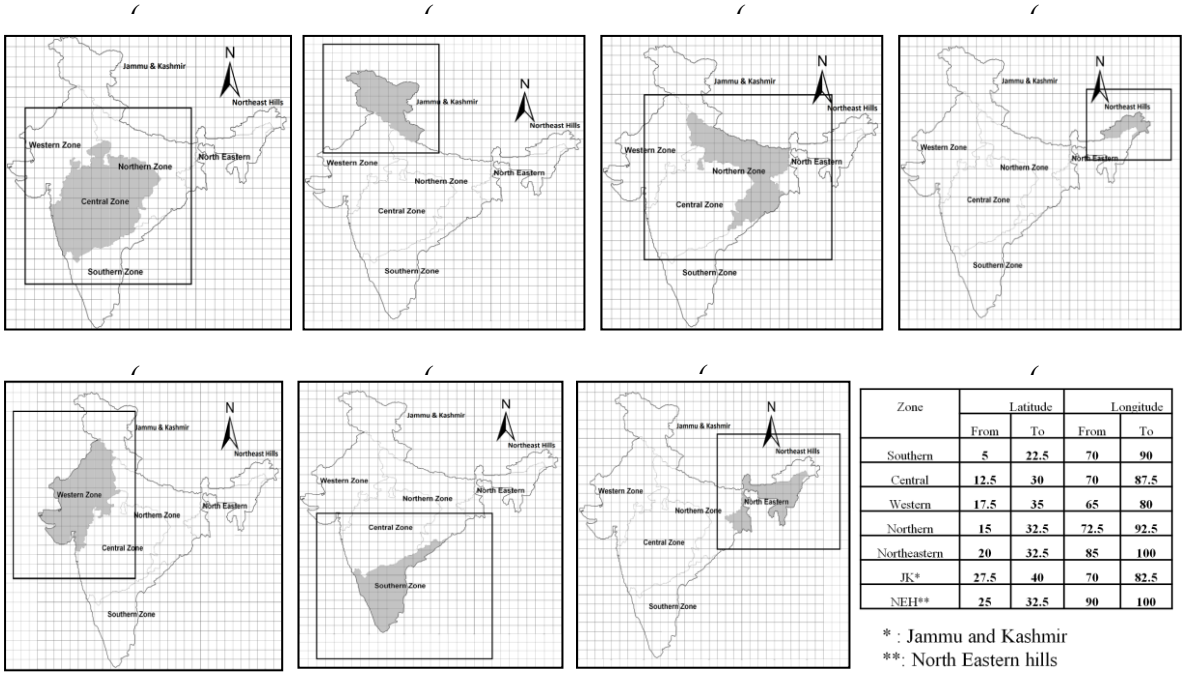


***Figure 2.2**: Correlation analysis for selection of zone of predictors. Spatial extent of the predictor selection for meteorologically homogeneous subdivisions in India (a) and is illustrated for central zone (b). The correlation contours between the zonal averaged rainfall time series, and (c) Humidity (d) Pressure (e) Temperature (f) Uwind and (g) Vwind are utilized; and a common region is selected from all plots (as shown by the black rectangle) in such a manner that it completely encompasses central zone and contains the region having a high correlation contours for all predictors.*

Figure 2.3 provides the region encompassing each zone, assumed to be the region of predictors for the corresponding zone. Each meteorologically homogeneous zone is highlighted with a gray shade and the region of predictors is shown with a black rectangle (refer to Figure 2.3 (a-g). The exact locations of the region of predictors in terms of latitude and longitude boundaries are shown in Figure 2.3h in the form of a table. From this point forward, these regions are referred to as the region of predictors.



Zone	Latitude		Longitude	
	From	To	From	To
Southern	5	22.5	70	90
Central	12.5	30	70	87.5
Western	17.5	35	65	80
Northern	15	32.5	72.5	92.5
Northeastern	20	32.5	85	100
JK*	27.5	40	70	82.5
NEH**	25	32.5	90	100

* : Jammu and Kashmir
**: North Eastern hills

Figure 2.3*: The region of predictors (as shown with black rectangles) for each meteorologically homogeneous zone (as shown with gray shade) are illustrated (a) Central, (b) Jammu and Kashmir, (c) North, (d) Northeast hills, (e) Western, (f) South, and (g) Northeast. The extent of the region of predictors in terms of latitude and longitudes are detailed in (h).*

Clustering is the process of grouping objects into clusters such that the objects from the same cluster are similar and objects from different clusters are dissimilar. Objects can be described in terms of measurements (e.g. attributes, features) or by relationships with other objects (e.g. pairwise distance, similarity). Clustering does not require assumptions about category labels that tag objects with prior identifiers. Therefore, clustering is an unsupervised data classification technique used to group together feature vectors; which are close to one another in a multidimensional feature space, to uncover some inherent structure which the data possess.

Capturing the cross-correlation amongst the rainfall at nearby rain gauge stations or nearby nodes (in the case of gridded data) is one of the research problems in multisite statistical downscaling. Cross correlation is the correlation among the rainfall time series at two or more rain gauge stations. Rainfall is a spatial phenomenon. Thus, it is highly unlikely that two rain gauge stations that are in the proximity of one another receive rainfall with magnitudes that are significantly different. However, for single site statistical downscaling (i.e. projecting rainfall at individual stations), the influence of rainfall occurring at nearby points on the location where downscaling is performed is neglected. As a result, the model failed to capture the spatial pattern of rainfall that could be visualized using cross-correlation plots. Mehrotra and Sharma **[**2005] introduced the concept of the weather state in order to tackle the multisite problem. The weather state captures the spatial distribution of rainfall on a particular day over a selected region. Kannan and Ghosh **[**2011] adopted an unsupervised data classification technique, K-means clustering, in order to group together the rainfall data at all nodes in one zone and to assign a state in order to represent the data. The same approach is used in this study to negotiate the multisite problem. K-means clustering is a method of cluster analysis that aims to partition 'n' observations into k clusters for which each observation belongs to the cluster with the nearest mean. In the current study, K-means clustering is used to deal with the spacial variation of rainfall in a particular zone. In order to preserve the same imprint of spatial variation in further analysis and to maintain the cross-correlation among the grid points within a particular zone**,** the k-mean clustering technique is adopted. The technique reads the observed rainfall values for all nodes in a zone on any day, clusters them, and provides one representative value that is referred to as the state for that day. Th**e** step is important, provides the representative rainfall category for a particular day, and is linked to the predictors for establishing the statistical relationship. The total number of clusters formed from the data should be optimized to represent the true classification that is inherently present in the data**.** Clustering is used to classify multi-site rainfall for each zone into different states (in the form of clusters). Clusters can be considered as groups into which the data is divided. Each cluster has a corresponding state to which it is associated. Hence, the state can be viewed as an identity number assigned to each cluster. The observed rainfall data for each zone is clustered using the K-means clustering technique**.** The total number of clusters presented in the data is determined to be three after comparing the Dunn's index and the Silhouette index obtained for different combinations of numbers of clusters (the number of clusters varied from 2 to 11)**.** The highest values for these indices, indicating the optimum number of clusters in the data, are obtained when the data is classified into two clusters**.** However, two clusters imply a dry wet day

classification. Therefore, in order to capture the spatial patterns of non-zero multisite rainfall [Kannan and Ghosh, 2013], the optimum number of clusters is identified as three for which the second highest values of the cluster validity indices are obtained**.** Since, three cluster formations is selected based on rainfall in all of the nodes, any zone is assigned one of the three state values (1, 2, and 3).

The Classification and Regression Trees (CART) is a supervised classification based model that builds classification trees for categorical dependent variables and regression trees for predicting continuous dependent variables. This technique uses historical data to construct decision trees. The Decision trees are then used to classify new data. In general, supervised classification is a machine learning technique for learning a function from training data. CART uses training sample for building the decision trees. The task of CART is then to predict the value of the function in terms of categorical/ continuous value for any valid input object. Classes in the training sample may be provided by user or calculated in accordance with some exogenous rule. In the present study k-means clustering is applied to cluster multisite rainfall data into rainfall classes/ states containing categorical values. These categorical values are used in the training sample for building the decision trees.

Modeling the occurrence of daily rainfall state involves construction of a classification tree. CART analysis is a form of tree building technique with binary recursive partitioning. The tree building technique classifies objects or predicts outcomes by selecting from a large number of variables in determining the outcome. The term “binary” implies that the learning set (training vectors), represented by a “node” in a decision tree, can only be split into two groups. Thus, each node can be split into two child nodes, in which case the original node is called as a parent node. The term “recursive” refers to the fact that the binary partitioning process can be applied over and over again. Thus, each parent node can give rise to two child nodes and, in turn, each of these child nodes may themselves be split, forming additional children. The term “partitioning” refers to the fact that the dataset is split into sections or partitioned.

The advantages of CART are the following:

(1) CART makes no distributional assumption of any kind, either on dependent or on independent variables.

The predictor variables used in CART can be a mixture of categorical, interval, and continuous.

(2) CART is not affected by outliers, colinearities, hetero-scedasticity or distributional error structures that affect parametric procedures.

(3) CART has the ability to detect and reveal interactions among the predictors.

(4) CART is invariant under monotonic transformation of independent variables and

(5) CART effectively deals with higher-dimensional data.

The following section gives details of modeling the occurrence of daily river basin-wise rainfall states.

Nonparametric statistical downscaling techniques such as the k-nearest neighborhood are widely used for multisite rainfall projections [Lall and Sharma, 1996; Yates et al., 2003; Mehrotra et al., 2004; Mehrotra and Sharma, 2006]. Any nonparametric regression estimator is generally a smoothing filter that projects the predictand for a desired set of predictors by applying weights to the other predictand in the neighboring region of the one desired. In general, a weight function is deployed in order to fulfill this task, which assigns heavy weights to the nearby data and very low weights to data that is far away. Multivariate kernel regression is a nonparametric technique in statistics that is used for estimating the conditional expectation of a random variable. The objective is to capture the nonlinear relationship between a pair of random variables X and Y [Wand and Jones, 1995]. Therefore, for the present case, by considering the nonlinear relationship between daily rainfall and the predictors, this technique is adopted**.** Mathematically, the general form for the conditional expectation of the **k**ernel **r**egression can be written as follows:

$$E(Y|X) = m(X) = \frac{\int y\, f(y|x)}{f_x(x)} \quad \text{... (2.3)}$$

where Y is the set of predictands at the node level, X is the set of principal components, f(y|x) is the conditional probability density function (PDF) of Y given X = x, and $f_X(x)$ is the marginal PDF of X. Nadaraya and Watson **[**1964] provided an estimator by replacing the multivariate PDF with kernel density estimates.

$$m_h(x) = \frac{\widehat{\sum_{i=1}^{n} K_h(x - X_i) Y_i}}{\sum_{i=1}^{n} K_h(x - X_i)} \quad \text{... (2.4)}$$

Where $m_h(x)$ is the expected value Y for a condition of Xi = x, and K_h is the kernel with bandwidth h.

2.4. Results

The results from the statistical downscaling model were compared with the observed dataset provided by Indian Meteorological Department. Here we provide comparison of statistically downscaled rainfall using Kernel Regression from CCCMA CanESM2 with IMD observed rainfall. Figure 2.4 Observed (IMD) Mean Rainfall (1976-2005) provides mean IMD rainfall averaged from 1976-2005 and Figure 2.5 Projected Mean Rainfall using GCM Predictors (1976-2005) provides projected rainfall using Kernel Regression. Figure 2.6 provides Difference between the two.

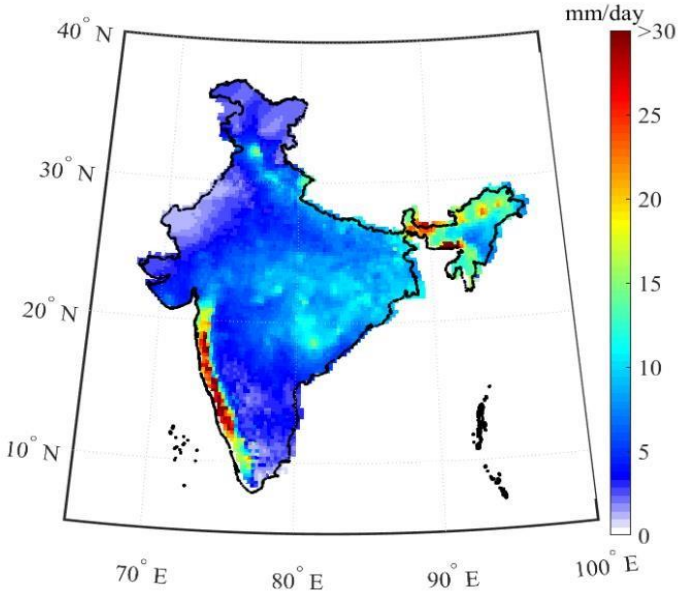


***Figure 2.4**: Observed (IMD) Mean Rainfall (1976-2005)*

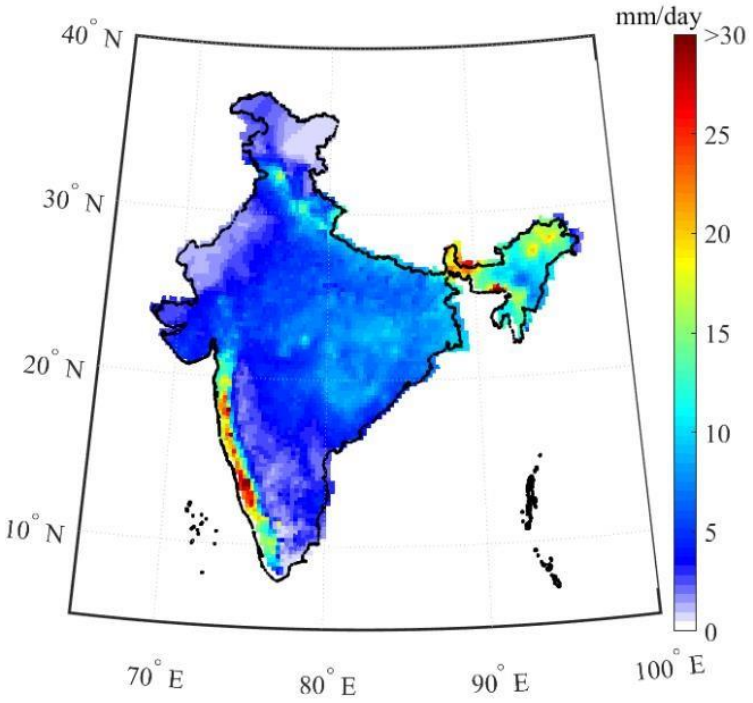


***Figure 2.5**: Projected Mean Rainfall using GCM Predictors (1976-2005)*

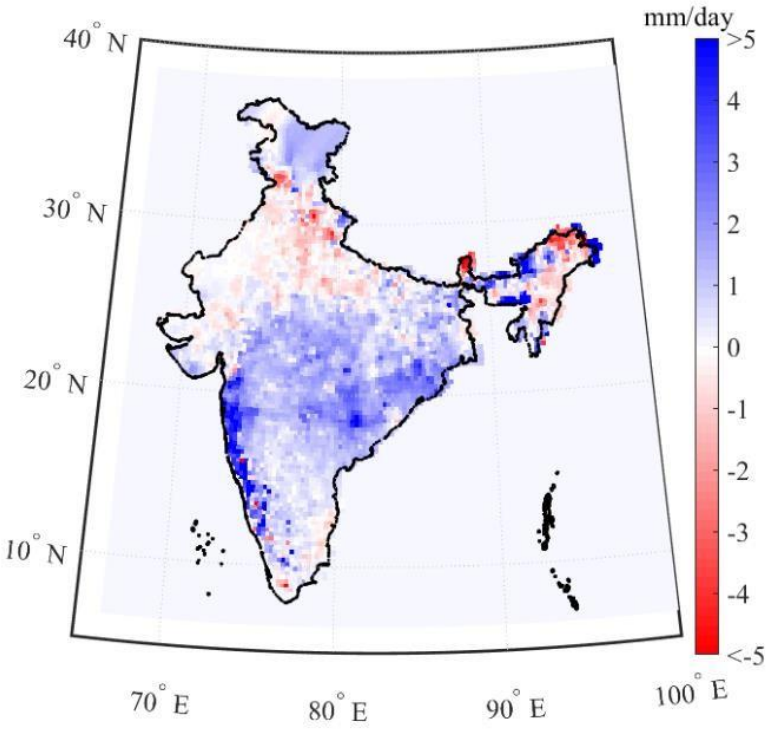


***Figure 2.6**: Errors in mean rainfall (Observed - simulated)*

Chapter 3

Statistical Downscaling for Hydro-climatic projections with CMIP5 simulations to assess the Impact of Climate Change

3.1. Introduction

This document described the GRNN methodology used for the statistical downscaling of the GCM data. The GRNN model is used for predicting the future rainfall for RCP4.5 and RCP8.5. The projection is available in a gridded format for further use in river basin scale hydrologic studies.

3.2. Methodology

Generalized Regression Neural Network (GRNN), formulated by Donald F. Specht in Specht (1991), is neural networks having a one-pass learning algorithm. This algorithm can be used for any regression problem even it involves a complex and non-linear relationship of the input and the target variables. The output is estimated using the weighted average of the output of the training dataset, where the weight is calculated using the Euclidean distance between the training data and test data. If the distance is large, then the weight will be very less, and on the other hand, if the distance is small, it will put more weight on the output. A schematic diagram of the GRNN architecture is presented in Fig. 3.1. GRNN configuration consists of four layers. The first layer consists of input neurons, the second layer is pattern neurons with activation function $\varphi_i(d_i)$, the third summation layer, and the final layer is the output layer.

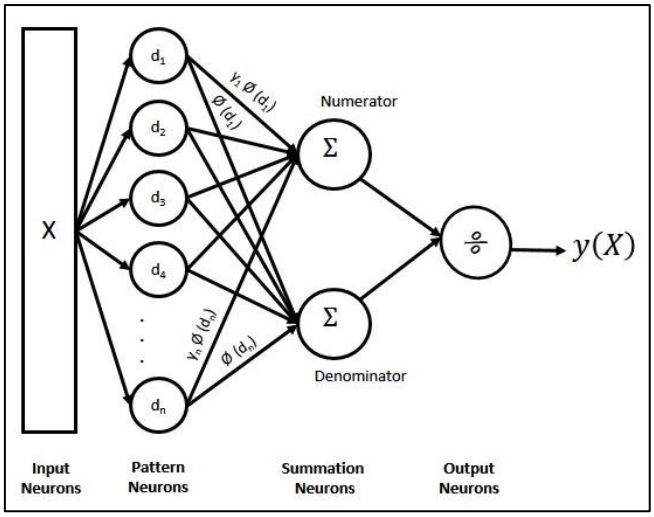


Figure 3.1*: Schematic diagram of generalized regression neural network architecture*

Here, $X = [x_1, x_2, \ldots x_n]$ are inputs and $y_1, y_2, \ldots y_n$ are targets. The difference between two vectors d_i is calculated as,

$$d_i = (X - X_i)^T (X - X_i) \quad \text{..... (3.1)}$$

The activation function,

$$\varphi_i(d_i) = exp\left(-\frac{d_i}{2\sigma^2}\right) \quad \text{..... (3.2)}$$

Numerator = $\sum_{i=1}^{n} y_i \varphi_i(d_i)$, Denominator = $\sum_{i=1}^{n} \varphi_i(d_i)$,

$$y(X) = \frac{Numerator}{Denominator}$$

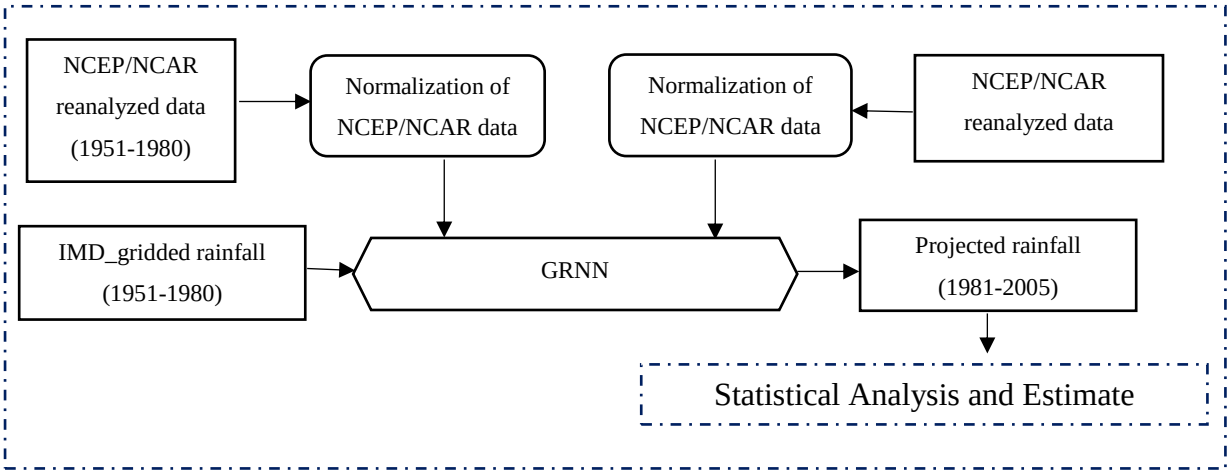
$$\Rightarrow \; y(X) = \frac{\sum_{i=1}^{n} y_i \varphi_i(d_i)}{\sum_{i=1}^{n} \varphi_i(d_i)} \quad \text{.....(3.3)}$$

$y(X)$ is the predicted output from the model of input X. In this project, bias-corrected PCA analyzed GCM outputs is used as an input 'X' and the IMD original rainfall 'y'is used as the target.

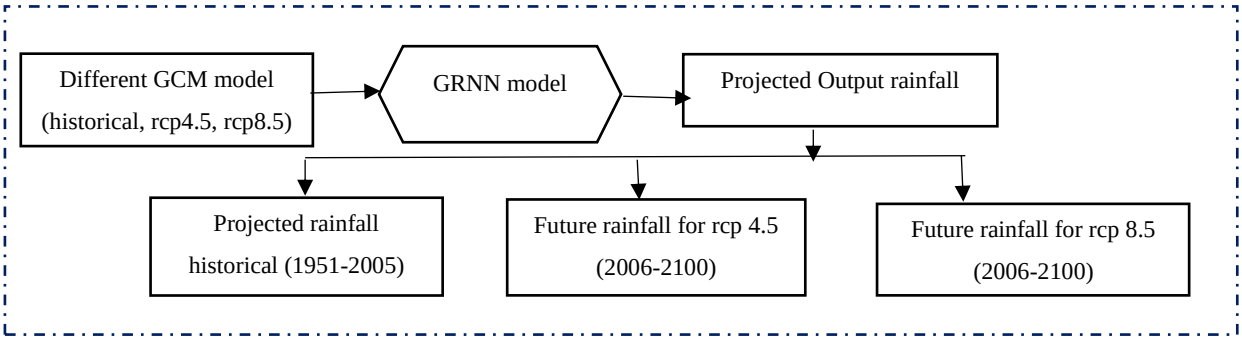
3.3. Model setup

NCEP/NCAR reanalyzed data has been used as Input data for the period of 1951-1980, and the IMD original rainfall data is used as a target of our GRNN model. The optimal spread value is

obtained using the testing period (1981-2005) data. Fig. 3.2 shows the flow chart of the model used for the prediction of future rainfall.



Development of GRNN Model



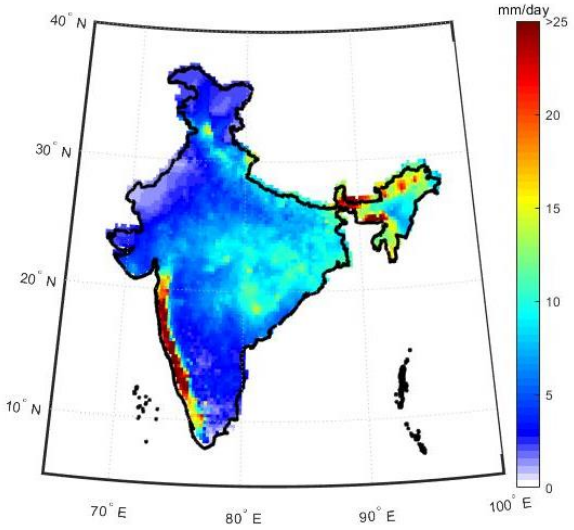
Prediction

***Figure 3.2**: Flow chart of our proposed technique*

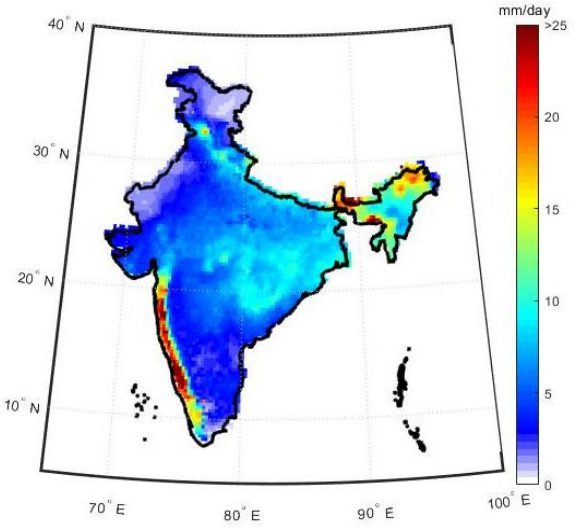
3.4. Results and discussions

The GRNN model has been run for all the seasons for the NCEP_NCAR dataset with an optimized spread value of 0.5, the regression coefficients of IMD original rainfall, and projected rainfall for all zones found more than 0.6. The projections for the 5 different GCMs are obtained for the entire India. The obtained results are validated using various statistical checks. The capabilities of the model in terms of accuracy and spatial variations are validated by projecting rainfall using NCEP/NCAR reanalysis climate variables as predictors and by comparing the results to the observed data.

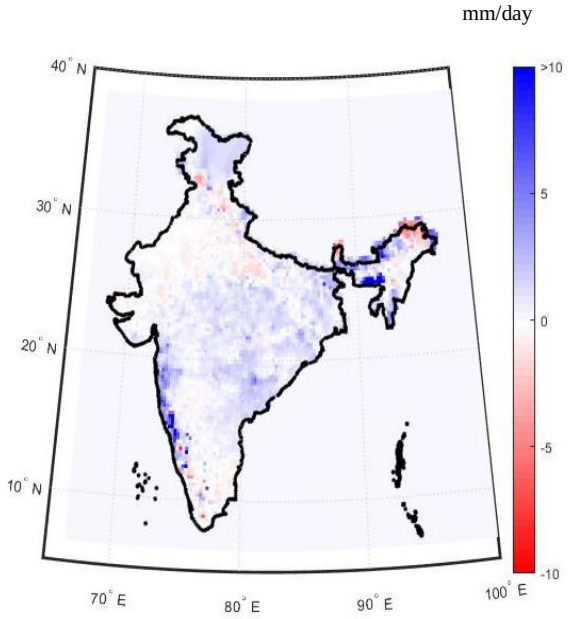
The training period is selected as 1951–1980, and the testing period is selected as 1981–2005. The spatial distribution of mean rainfall shown by the observed data is well replicated by projected mean rainfall (Fig. 3.3a and 3.3b). Although the maximum absolute difference between observed mean rainfall and projected mean rainfall is approximately 10 mm, for most of the parts of India, the difference is around 3 mm (Fig. 3.3c). As compared to other sections of India, the difference is greater where actual mean rainfall is greater. Here we have produced the results (Fig. 3.4) for the CCCma-CanESM2.



(a) IMD observed rainfall (1981-2005)



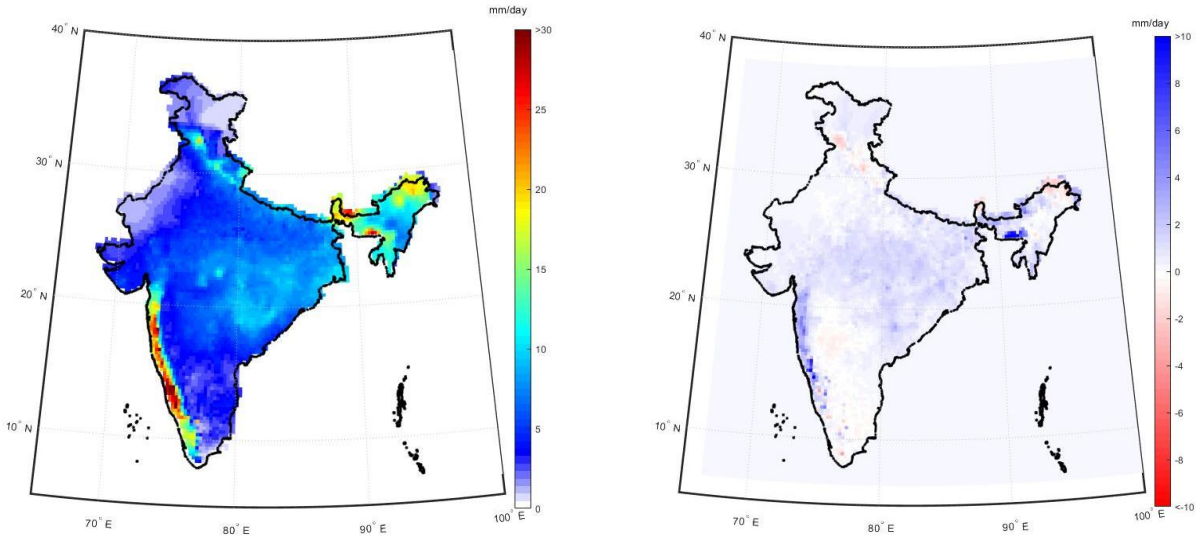
(b) Projected historical rainfall (1981-2005)



(c) Resulting Error

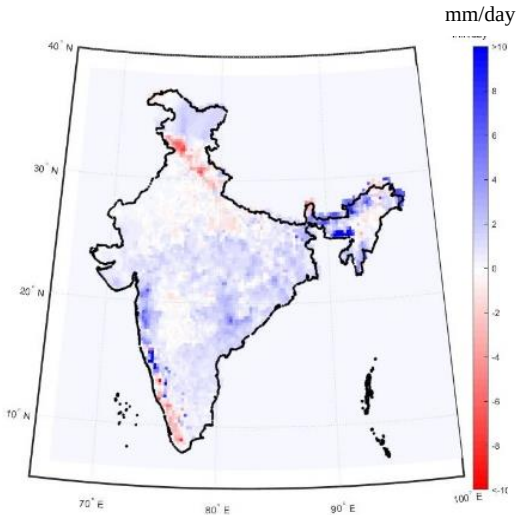
***Figure 3.3**: Comparison of statistical properties of observed rainfall and projected rainfall (using NCEP/ NCAR reanalysis data) for entire India (1981–2005). Mean of (a) observed rainfall data (b) projected rainfall data shows good match in magnitude as well as spatial variability resulting in low error values (c) with ranges around -3 to 3 mm for the majority of the country.*

Change in rainfall projection of GCM models for the future represented in three divisions, near_future (2010-2039), mid_future (2040-2069), far_future (2070-2099) as compared to the historical period (1976-2005) both for RCP 4.5 and RCP8.5 are presented in Fig. 3.5.

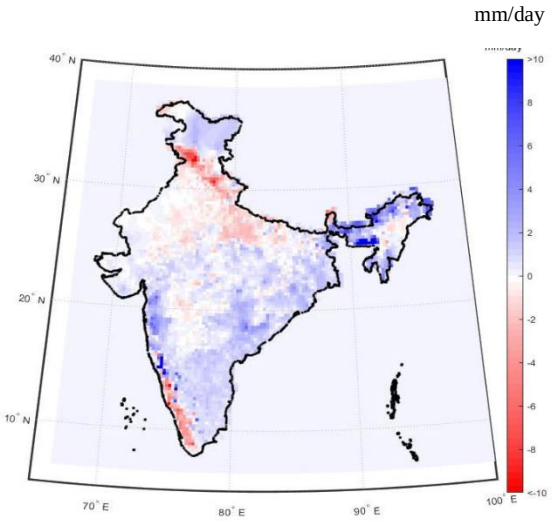


(a) Projected historical rainfall(1981-2005) (b) Resulting Error with respect to IMD rainfall

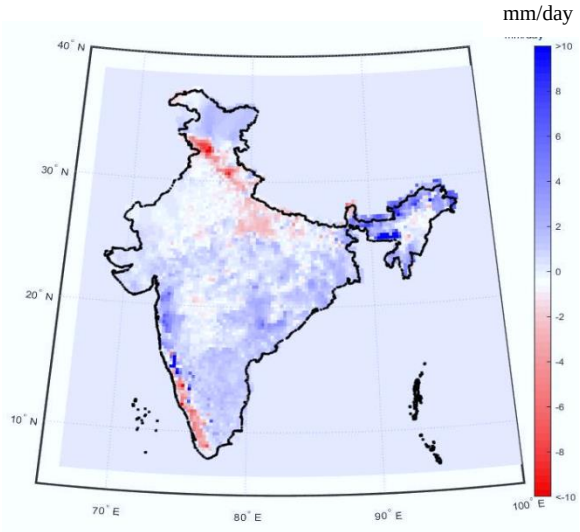
***Figure 3.4**: Statistical representation of CCCma_CanESM2 GCM for the monsoon season (historical).*



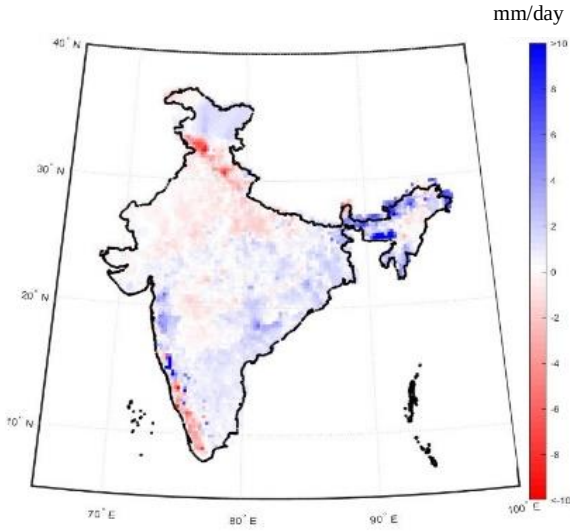
(a) Change in rainfall during near_future



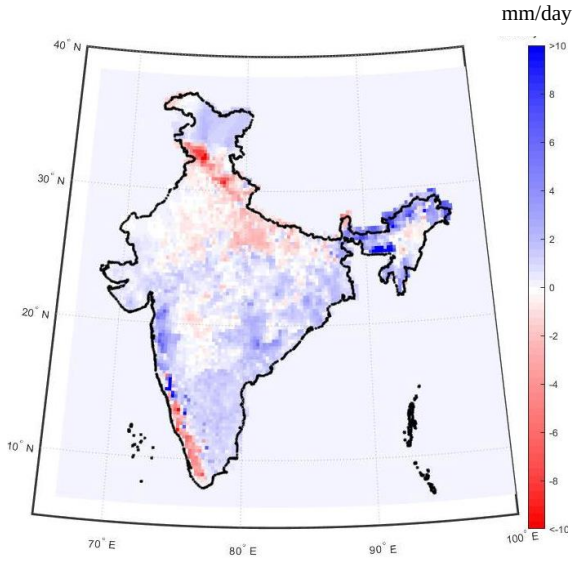
(b) Change in rainfall during mid_future



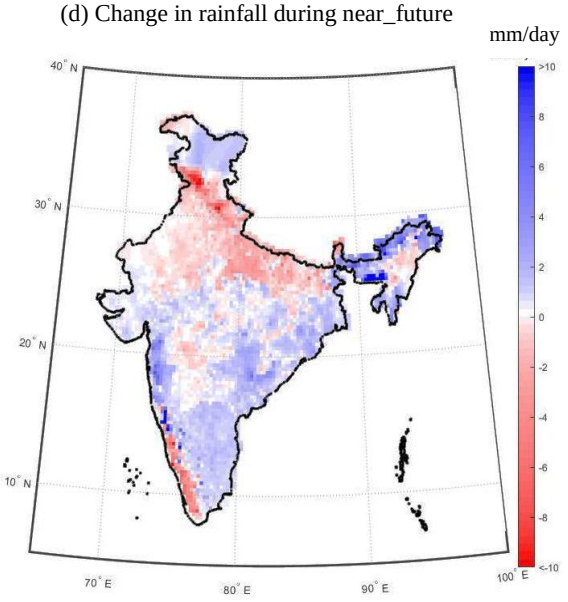
(c) Change in rainfall during far_future



(d) Change in rainfall during near_future



(e) Change in rainfall during mid future



(f) Change in rainfall during far future

Figure 3.5*: Model performance for future, changes in the projected rainfalls with respect to IMD original rainfall (1976-2005) in terms of climate change at RCP4.5 and RCP 8.5, respectively*

Note: This study projected future precipitation using the GRNN model. Other parameters like temperature, pressure, humidity, etc. have been projected by correcting the bias of NCEP/NCAR reanalyzed data. For the bias-corrected data, IITB's dataset may be referred.

Chapter 4

MLP-ANN based Statistical Downscaling

4.1. General

'Downscaling' techniques are used to overcome the inability of GCMs to simulate climatic events at finer resolutions. Downscaling is the process by which coarser-resolution GCM outputs are translated into finer-resolution climate information, so that they better account for regional climatic influences, such as local topography. Downscaling uses the outputs of the GCM as inputs and simulates the values of other climatic events that operate at finer resolutions. There are primarily three methods of downscaling, namely: change factor methods, statistical downscaling, and dynamic downscaling. The statistical downscaling methods involve developing relationships between local synoptic situations and large-scale climate simulations. This usually leads to a very close representation of the observed climate in the statistical downscaling model. The model can then be used to infer local-scale changes from large-scale changes generated by GCMs. In this project, we have focussed on statistical downscaling using Kernel regression and ANNs. The details of the development of the Kernel regression and GRNN based statistical downscaling models and associated results are presented in the previous chapters of this report. In this chapter, we present the results of the multi-layer perceptron (MLP) type ANN models developed for statistical downscaling of the climatic variables to estimate the rainfall for entire India. The inputs and output data remain the same as presented earlier, the only difference is the type of ANN model developed in this chapter.

The MLP-ANNs show a good performance in downscaling of rainfall in arid regions where observed rainfall is meagre (Ahmed et al. 2015). The MLP-ANNs are efficient in solving regression and classification type of problems by establishing the non-linear relationships between inputs and output data and hence are most suited in prediction and forecasting of hydrologic and climatic variables (Vidyarthi and Jain, 2020). It is known that the random climatic variables (predictors) can be very well linked to observed rainfall through ANN models (Laddimath and Patil, 2016) and hence an ANN can serve as a good tool for downscaling. This study investigated the use of MLP-ANNs for statistical downscaling of the

Indian rainfall using the GCM outputs. The specific objectives of the present study are summarized below.

1. Development of ANN model using NCEP and IMD observed data for the historical period.
2. Projection of the rainfall data for the 21st century using the developed ANN model using five different GCMs outputs (also known as Predictors).

The data employed in this study are the NCEP predictors, and the IMD observed rainfall data for historical period (1951-2005); and five GCMs predictors for the projection period (2006-2100) as described previously. This section of the chapter sets the context and background and includes the research problem and objectives. The second section provides the details of the MLP type ANN models employed in this study. Section 4.3 presents the model development, while Section 4.4 reflects the results and discussion of this work. The summary and conclusion are presented in Section 4.5.

4.2. MLP-ANNs

An Artificial Neural Network (ANN) is a netted structure exhibiting a learning process analogous to the human brain (Kleber et al., 2010). ANNs are essentially of two kinds feed-forward networks in which the information flows only in a forward direction and the feed-back networks in which the information is allowed to flow in both forward and backward directions. The multi-layer percentron (MLP) type ANN architecture is a feed-forward type ANN architecture and is the most popularly employed tool among engineers and scientists. The basic building block of an ANN is an *artificial neuron* (simply referred to as *neuron*) that tries to mimic the behaviour/ functioning of a *biological neuron* in the human brain. The neurons represent an aggregation function that assimilates all inputs coming to them and an activation function which produces an output depending on whether the aggregated input crosses a particular threshold or not. The governing/ activation function could be of any kind ranging from a simple quadratic error function to complex functions (source: accessed www.skymid.ai on 4th July, 2019).

In the current study, the ANN structure consisted of three layers, namely: an input layer, a hidden layer, and an output layer. Figure 4.1 depicts the structure of a typical MLP-ANN used in engineering and sciences. The input layer receives the input variables, the hidden layer establishes the relationship, and the output layer gives the processed output (Maind and

Wankar, 2014). The ANNs are able to learn from the data fed into them for training (Murata et al., 1993). This makes them proficient at identifying relationships. Coulibaly et al. (2000) and Maier et al. (2010) have reviewed the ANN-based modelling in hydrology and reported that a large number of studies used feed-forward MLP type ANNs.

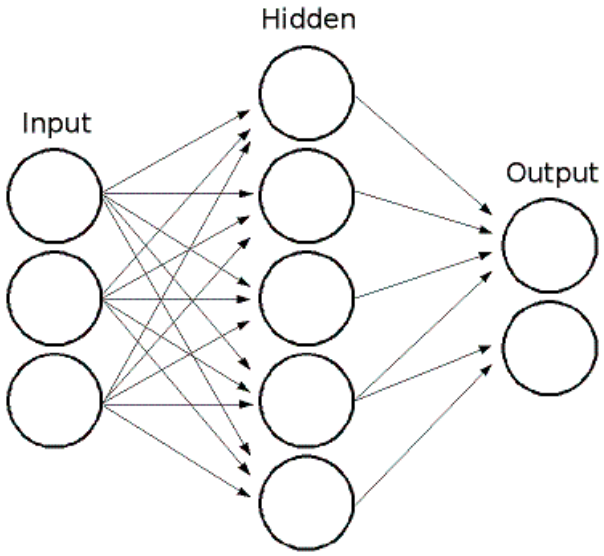


***Figure 4.1**: Structure of an MLP-ANN*

The training algorithm used in this method is the standard back propagation algorithm based on a recently developed optimization technique called as Adaptive Moment Estimation Algorithm (ADAM). The ADAM uses Momentum (β_1, β_2 and ε) and Adaptive Learning Rates (α) to converge faster as compared to the traditional gradient-decent method. β_1 and β_2 are the learning rate decays for the first and second moments, respectively, and ε is epsilon. The full mathematical derivation of this algorithm can be found in Kingma and Ba (2015). There are issues with the traditional gradient-decent back propagation algorithm (BPA) e.g. its tendency to get stuck in local optima. In other words, the final weights may be trapped over highly complex error surface and also may over fit the data (Bishop, 1995). Also the traditional BPA is painfully slow in convergence. The ADAM helps in overcoming these problems. The hyper parameter optimization using Grid-Search Algorithm based on Mean Absolute error is used for selecting the optimum number of hidden neuron (i.e., for architecture selection).

4.3. Model Development

In this section, the model development of MLP-ANN has been discussed. In the first step, a statistical relationship was established between the PCA analysed bias corrected NCEP

predictors and the observed IMD rainfall. For this, the data for the historical period was used. Once the ANN model was trained and tested, the rainfall projection is done by using the five different GCM predictors for the non-historic period (2006-2100). A flow Chart for depicting the procedure is given in Figure 4.2.

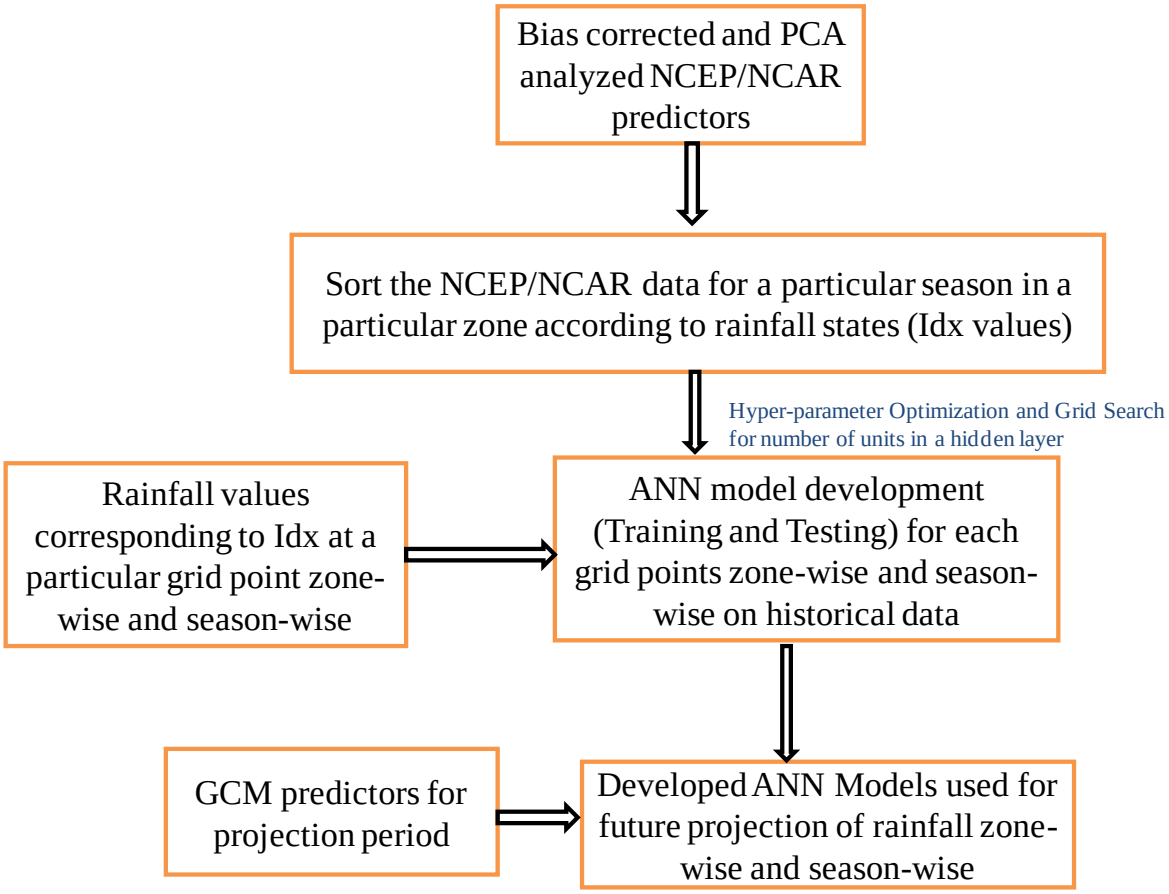


***Figure 4.2**: Flow Chart depicting the methodology*

STEPS:

i. ***Development of relation between NCEP Predictors (Ps) and IMD Rainfall data (R):*** Develop relation at each grid point for each rainfall state in a season, R = *f* (Ps), where R is the IMD rainfall and Ps are the NCEP predictors. R is the rainfall in mm corresponding to a particular rainfall state. The rainfall states are used only for segregation of the rainfall data. If three states are there, we need to develop three models for each state at each grid point. A sample excel sheet for a particular grid point can be prepared as shown below:

NCEP predictors for a Zone								IMD rainfall at a grid point	Rainfall states for a Zone
P1	P2	P3	P4	P5	P6	P7	---- P8	IMD Rain	IDx
								C1	1
								C2	3
								C3	2
								C4	2
								C5	1

Ps are the NCEP predictors and do not change on grid points in a zone. C1, C2…, C5 are the continuous values of the IMD rainfall data at a grid point. The IDx also do not change on grid points in a zone. Thus, only the IMD Rain data changes for each grid point. IDx for training is calculated by K-mean clustering for 3 clusters, while IDx for testing and projection periods comes from the CART model by using predictors of the NCEP predictors for testing period and using predictors of projection period, respectively. After segregation on the basis of IDx column, we develop three ANN models, one corresponding to each state. In a nutshell, we develop three ANN models at each grid point for three different rainfall states for each zone (by training and testing), and create a final time series of modeled rainfall, and observed IMD rainfall.

ii. ***Projection of rainfall using GCM data:*** For the GCM data, the IDx can be obtained from the CART model using the projected GCM predictors. The IDx gives the idea which rainfall state the projected rainfall would belong to and so using the developed ANN model for the corresponding state, the rainfall values can be projected.

The monthly mean and standard deviation of daily rainfall obtained from ANN model are compared with the monthly mean and standard deviation of the IMD rainfall data.

Study Area and data

As mentioned previously, the study area is entire India and the analysis is carried out on all the seven zones of India: Northern, Southern, Central, Western, North-East (NE), North-East Himalaya (NEH), and, Jammu and Kashmir (J&K). The NCEP predictors, and the IMD observed rainfall data for historical period (1951-2005); and five GCMs predictors for the projection period (2006-2100) are used. The predictors are climatic variables well simulated by the GCM which are used in downscaling for predicting local-scale hydrological processes like rainfall (Salvi et al., 2013). The predictors are obtained after PCA analysis of a larger set containing variables (Benestad et al., 2016). This helps in dimensionality reduction and makes

process computationally less expensive. However, PCA must not be done randomly as this might lead to the loss of essential variables (Salvi et al., 2013). Some of the predictors are: TAS-Near surface air temperature, which implies the temperature of the air near the surface, T850-Temperature of air at 850 hPa pressure, etc. The results of downscaling are very sensitive to the choice of predictors (Hewitson and Crane, 1996). All of the predictors mentioned here are the NCEP (National Centre for Environmental Prediction) predictors. All of these predictors are available for the historic period (1951-2005) and were bias corrected. A description of the climatic variables is given in the Table 4.1. The rainfall values for all the four seasons namely: JJAS (June, July, August, and September), MAM (March, April, and May), ON (October and November) and DJF (December, January and February) were obtained from IMD for the historic period at 0.5°×0.5° resolution. The Figure 4.3 shows one of the predictors after bias correction from one of the GCMs (CNRMCM5) and the same from NCEP for the historic period (1951-2005) for the ON season. This shows that the bias corrected GCM predictors follow the same trend with NCEP predictors and hence projected GCM predictors can be used for the rainfall projection in absence of NCEP predictors for future periods.

Table 4.1: The various predictors involved in the analysis

Variable/Predictor	**Abbreviation**
Near surface air temperature	TAS
Air temperature at 850hpa	T850
Air temperature at 500hpa	T500
Eastward near surface wind velocity	UAS
Eastward wind velocity at 850hpa	U850
Northward near surface wind velocity	VAS
Northward wind velocity at 850hpa	V850
Specific humidity at 850hpa	Q850
Sea level air pressure	PSL
Geopotential height at 500hpa	Z500

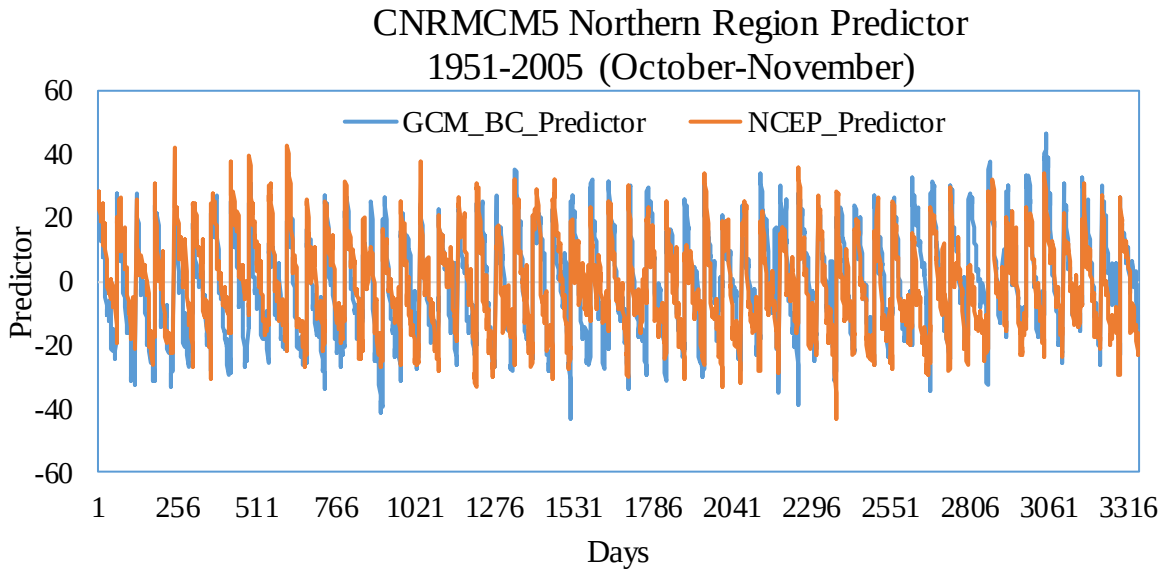


***Figure 4.3**: Comparison between the two predictors*

The rainfall values were not used directly to train the ANN model because such an input would result in a considerable loss of accuracy in the model. The rainfall values were clustered into groups denoted by numbers called rainfall states (IDx values: here 1, 2 and 3). For the clustering of rainfall values of historical period K-means clustering was used. To generate future rainfall states Classification and Regression Tree (CART) was used. The Figure 4.4 shows a general scheme of downscaling process devised by Kannan and Ghosh (2013). The scheme presented in Figure 4.4 was strictly followed while developing the model in this study. It is to be noted that the pre-processing of the data including PCA analysis, bias correction and rainfall states determination using CART was performed by another collaborators of this project (IIT Bombay). In this study, instead of using the Kernel Regression, MLP-ANN is used as the downscaling technique.

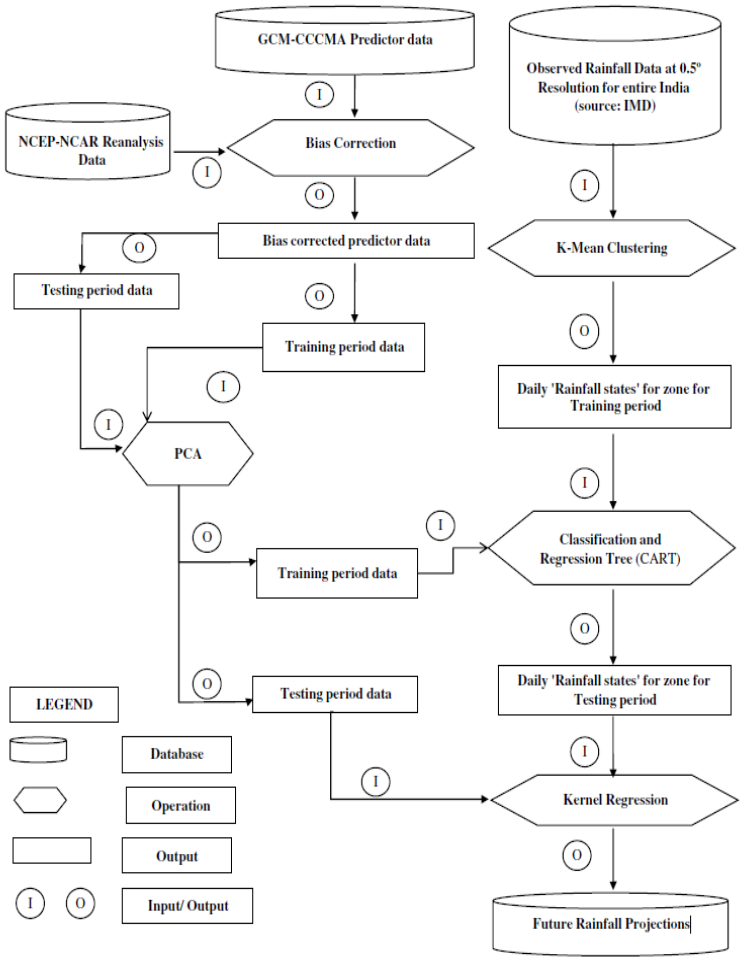


***Figure 4.4**: A general scheme of downscaling process devised by Kannan and Ghosh (2013)*

The PCA analysed, zone-wise, and season-wise input variables were obtained from IIT Bombay. Some examples are shown in the Table 4.2.

Table 4.2: Example of the Zone-wise and Season-wise Input variables

Regions	No. of Predictors for ON after PCA	No. of Predictors for MAM after PCA
Northern	16	15
Southern	20	17
Western	15	14

Central	17	15
Northeast	11	10
NEH	6	7
JNK	9	9

At the single grid point, three different ANN models were developed corresponding to three different rainfall indices (Id_x). The optimum architectures were in the form of Np-Nn-1, where, Np, Nn are the number of predictors and number of total hidden neurons, respectively. The optimum values of the ADAM parameters, α, β_1, and β_2 for obtaining the optimum weights were found in the range of 0.001 to 0.1, while the ε was of the order of 10^{-8} at different grid points and at different zones.

4.4. Results and Discussions

This current section illustrates the various findings of the study. The findings include the rainfall values for the entire 21st century with the help of GCM predictors. The findings include the statistics of the rainfall generated after the ANN models development.

The ANN models were trained using the NCEP predictors and the observed rainfall obtained from IMD. The data spans the entire historic period (1951-2005). Correlation between observed rainfalls and ANN simulated rainfalls was found to be an important factor for model efficiency. The correlation is represented in graphical form here. The correlation graphs for Northern Zone of India for visualizing the validation of the ANN model is produced for the 4 different seasons: JJAS, MAM, ON, and DJF in Figure 4.5 to Figure 4.12.

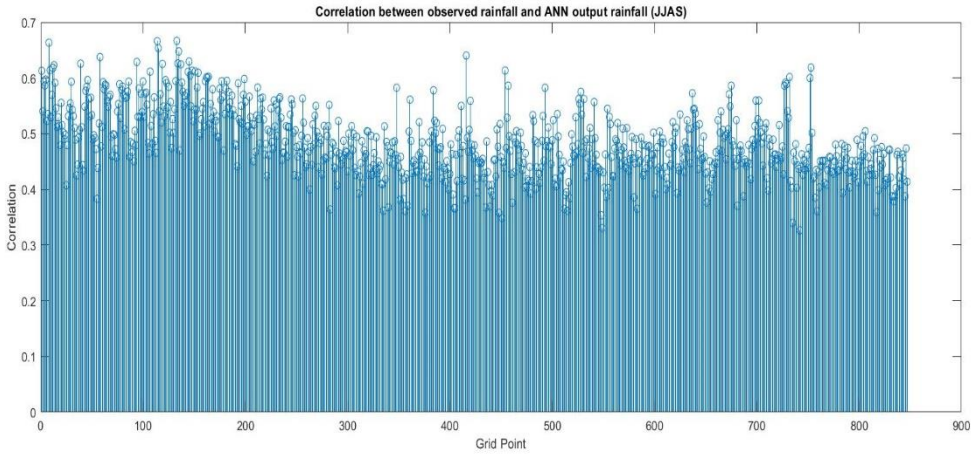


***Figure 4.5**: Correlation for JJAS-Training for Northern Zone of India*

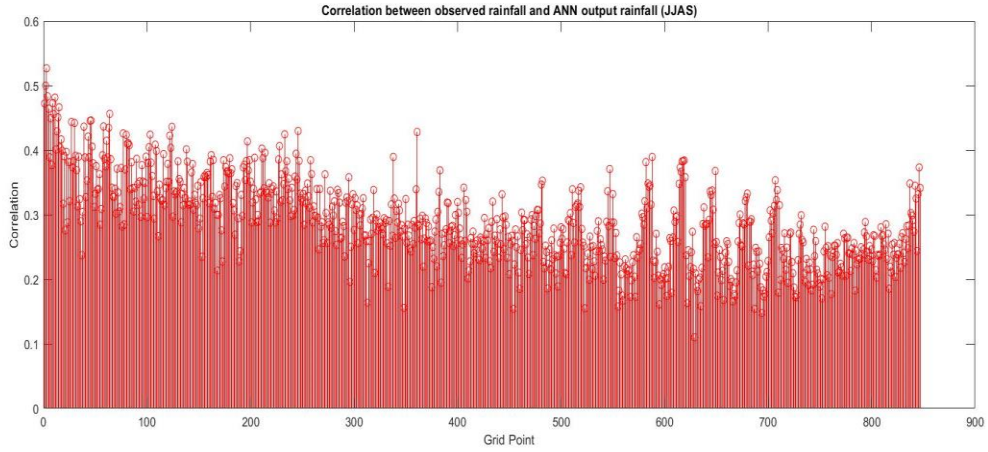


***Figure 4.6**: Correlation for JJAS-Testing for Northern Zone of India*

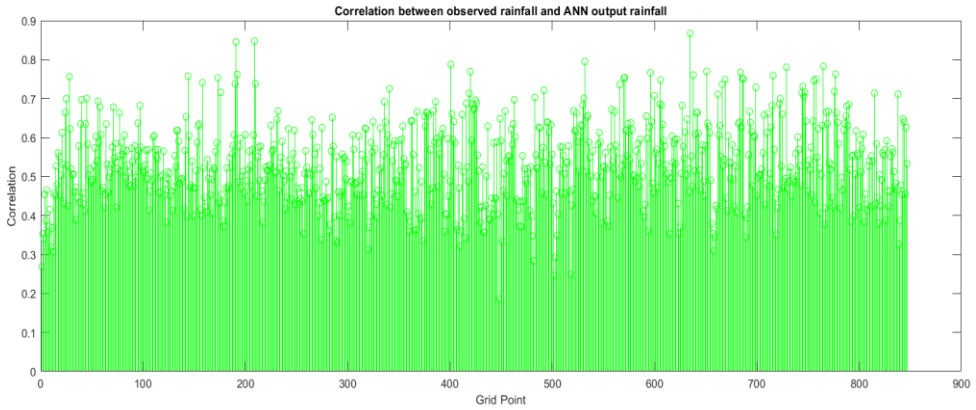


***Figure 4.7**: Correlation for MAM-Training for Northern Zone of India*

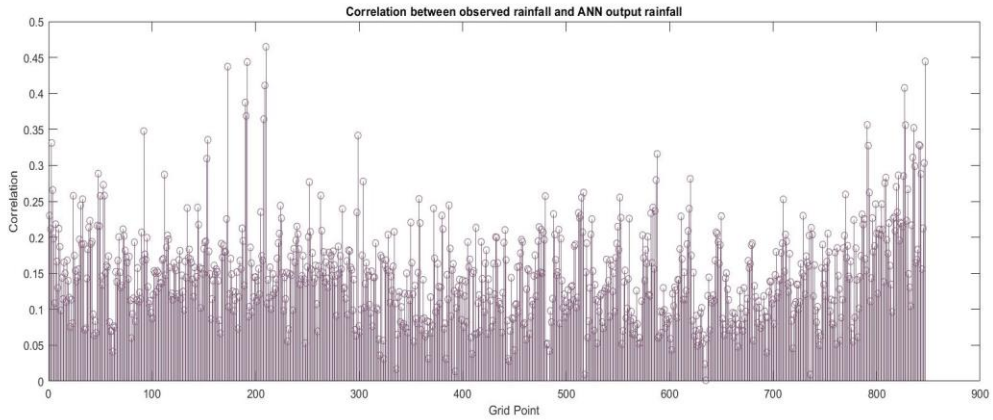


***Figure 4.8**: Correlation for MAM-Testing for Northern Zone of India*

The Figures 4.5, 4.6, 4.7 and 4.8 show the correlation values for the training and testing periods for the seasons JJAS and MAM for Northern Zone of India. For JJAS, during the training period most grid points show correlation greater than 0.4 while for the testing period, the corresponding value is 0.2. Similarly for MAM, during the training period most grids show a correlation greater than 0.4 while for the testing period the corresponding value is 0.1.

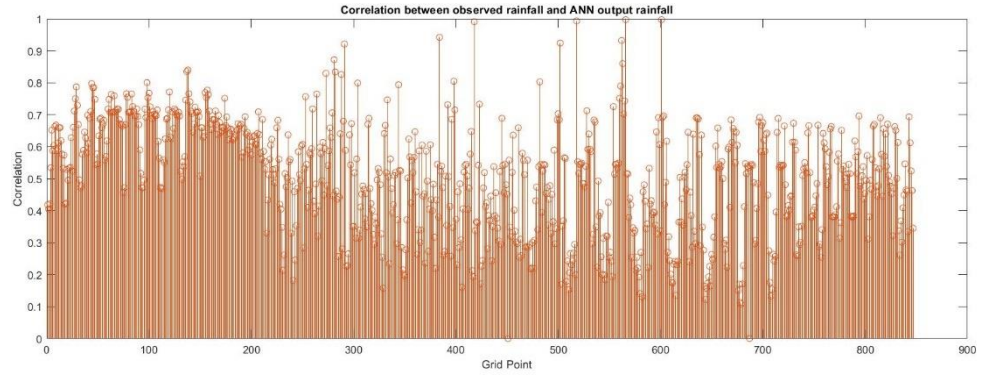


***Figure 4.9**: Correlation for ON-Training for Northern Zone of India*

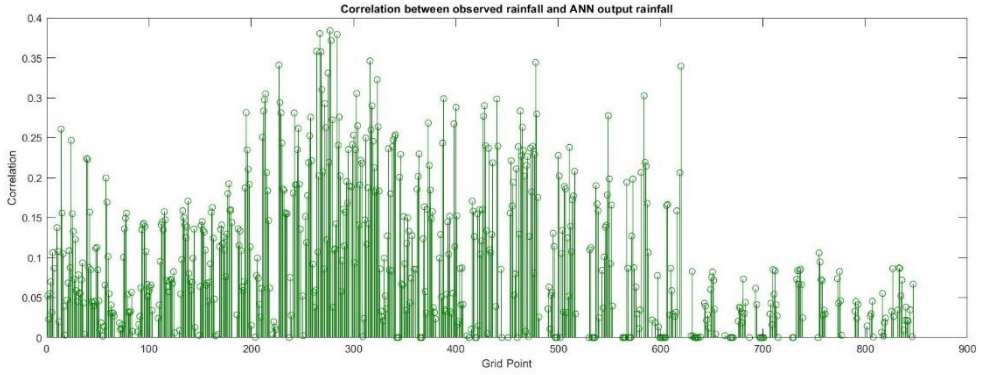


***Figure 4.10**: Correlation for ON-Testing for Northern Zone of India*

For ON, the correlation value for training period was more than 0.25 for most grid points while for the testing period the corresponding value oscillated between 0.05 and 0.1.

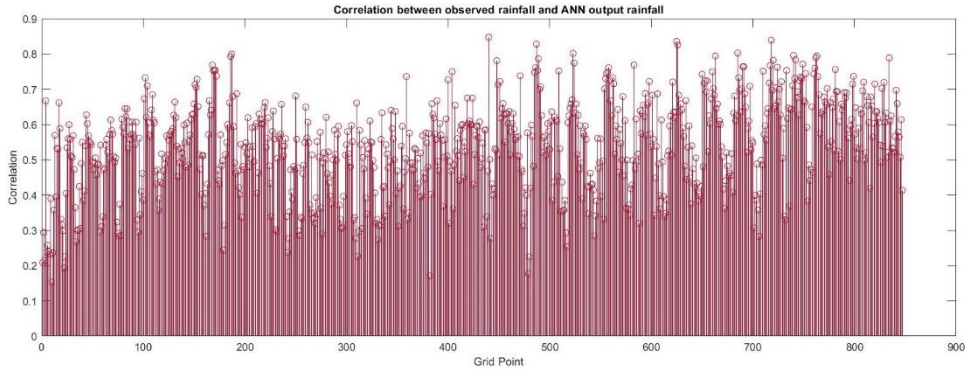


***Figure 4.11**: Correlation for DJF-Training for Northern Zone of India*

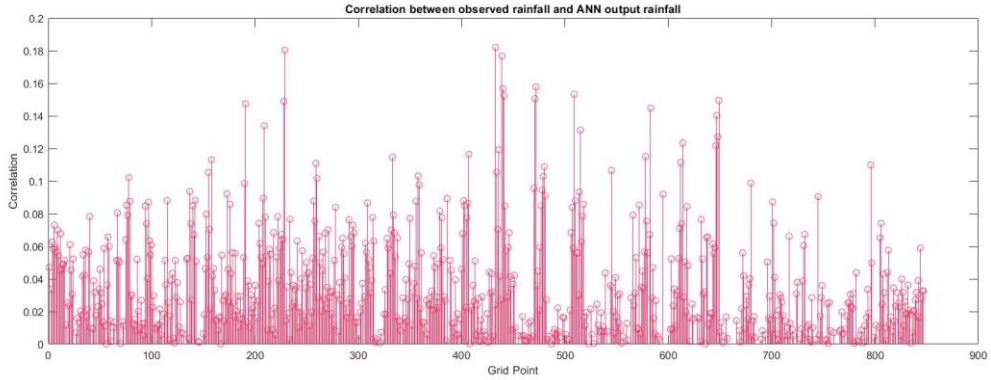


***Figure 4.12**: Correlation for DJF- Testing for Northern Zone of India*

For DJF, during the training period most grid points showed a correlation greater than 0.3 while for the testing period the corresponding value was 0.02. The plots between the average rainfall obtained from ANN and observed IMD rainfall for different seasons is presented in Figures 4.13 to Figure 4.16 which reveal that the ANN model captures the variation of observed average daily IMD rainfall efficiently for historic period.

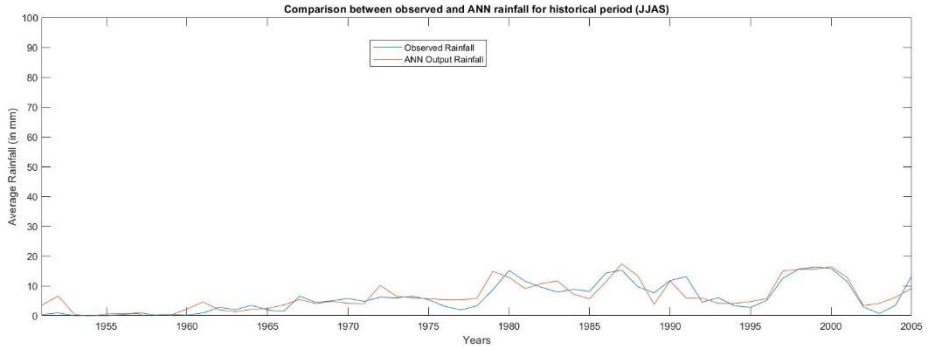


***Figure 4.13**: Calculated and Observed Average Rainfall for DJF for Northern Zone of India*

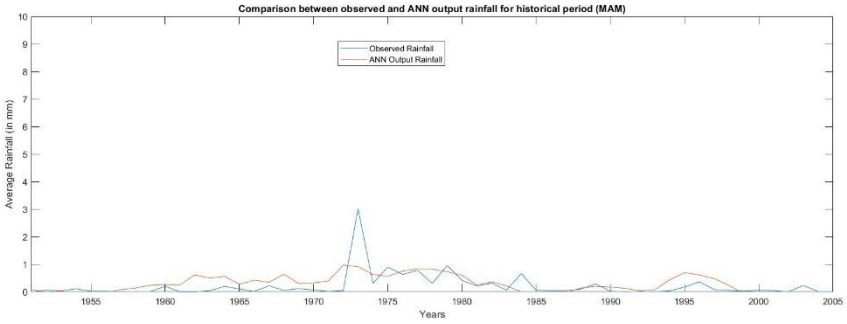


***Figure 4.14**: Calculated and Observed Average Rainfall for MAM for Northern Zone of India*

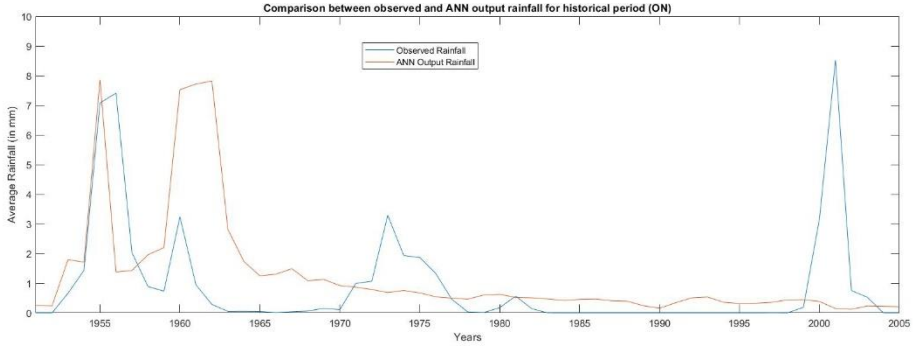


***Figure 4.15**: Calculated and Observed Average Rainfall for ON for Northern Zone of India*

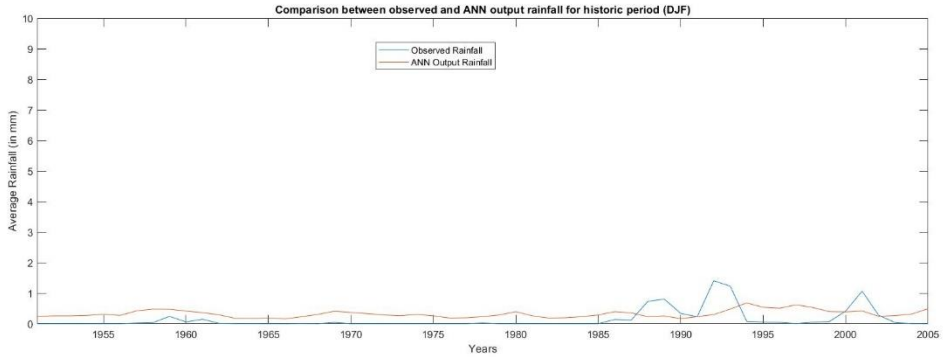


***Figure 4.16**: Calculated and Observed Average Rainfall for DJF for Northern Zone of India*

The plots between mean daily rainfall from ANN simulations and the observed IMD rainfall data for the historic period is shown in the Figure 4.17 to Figure 4.20 for different seasons of Western zone. The plot between mean daily rainfall from ANN simulations and the observed IMD rainfall data for the historic period is shown in the Figure 4.21 for a representative season JJAS of NEH zone.

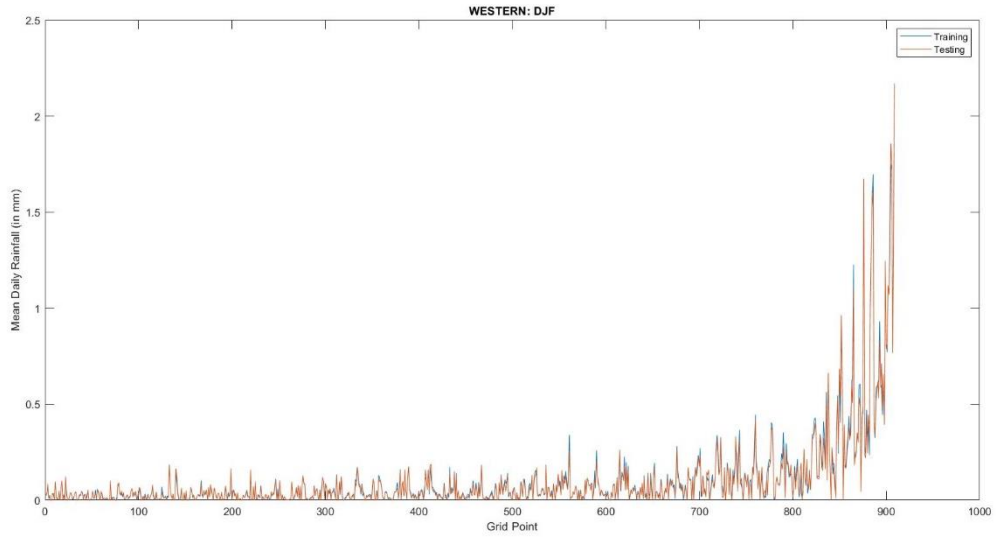


***Figure 4.17**: Calculated and Observed Average Rainfall for DJF for Western Zone of India for Historic Period*

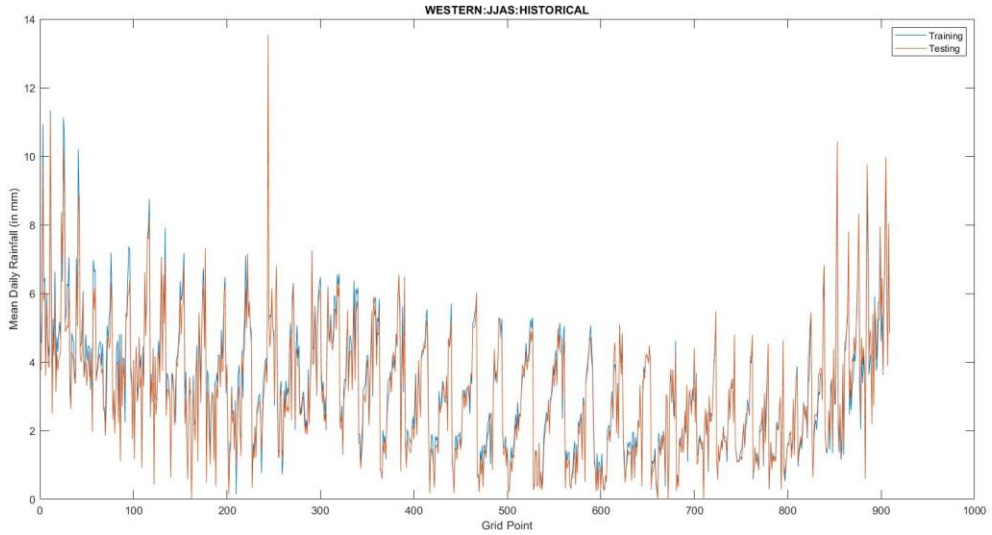


***Figure 4.18**: Calculated and Observed Average Rainfall for JJAS for Western Zone of India for Historic Period*

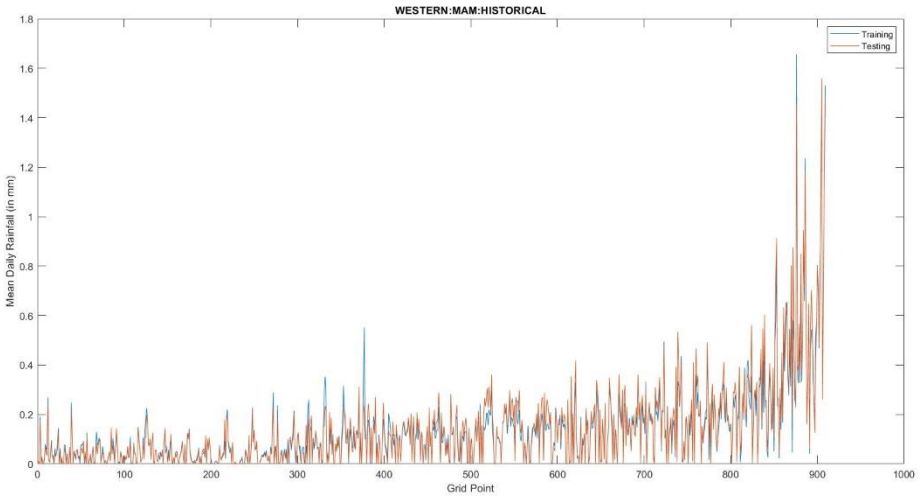


***Figure 4.19**: Calculated and Observed Average Rainfall for MAM for Western Zone of India for Historic Period*

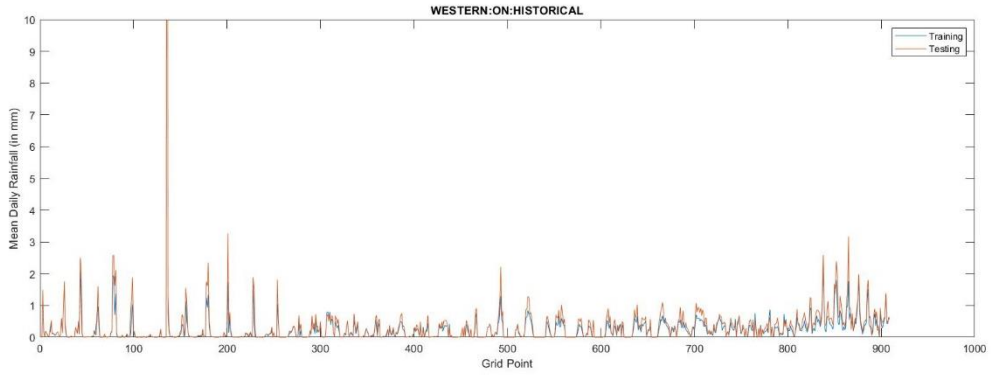


***Figure 4.20**: Calculated and Observed Average Rainfall for ON for Western Zone of India for Historic Period*

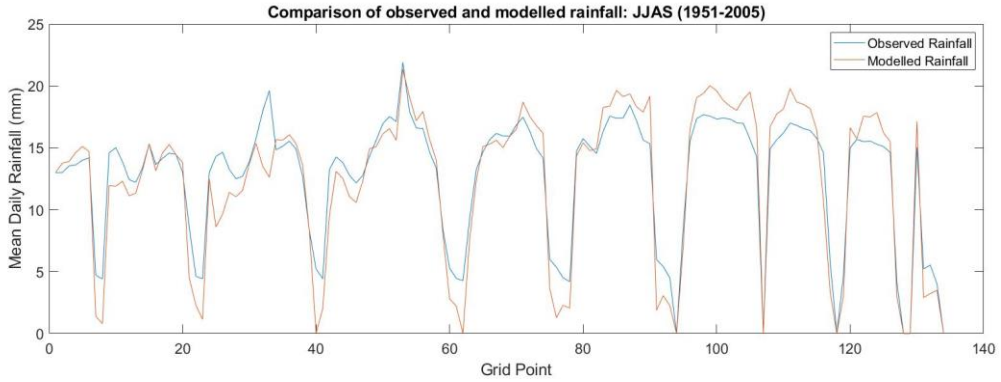


***Figure 4.21**: Calculated and Observed Average Rainfall for JJAS for NEH Zone of India for Historic Period*

As shown from the above graphs, the ANN is capable of simulating the relation between NCEP predictors and IMD rainfall. The trained models are then used for future projections of rainfall. Using the trained ANN models and the predictors from the GCM, the rainfall projections for future period was carried out for the period 2006 to 2100 and the average rainfall for few representative Zones of India are shown in Figures 4.22 to Figure 4.29.

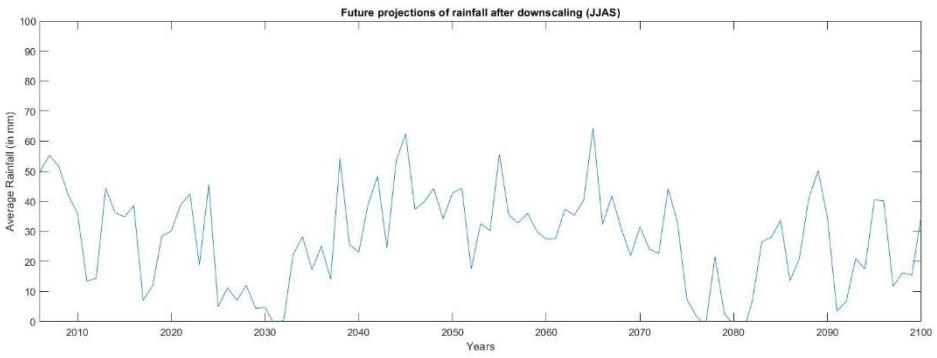


***Figure 4.22**: Projected Average Rainfall for JJAS for Northern Zone of India*

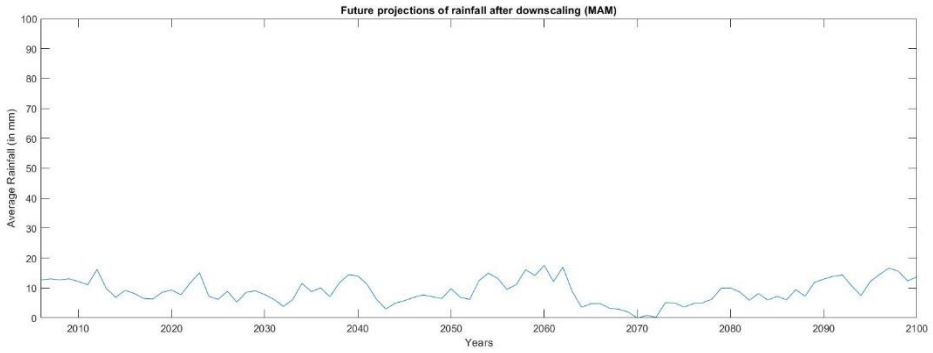


***Figure 4.23**: Projected Average Rainfall for MAM for Northern Zone of India*

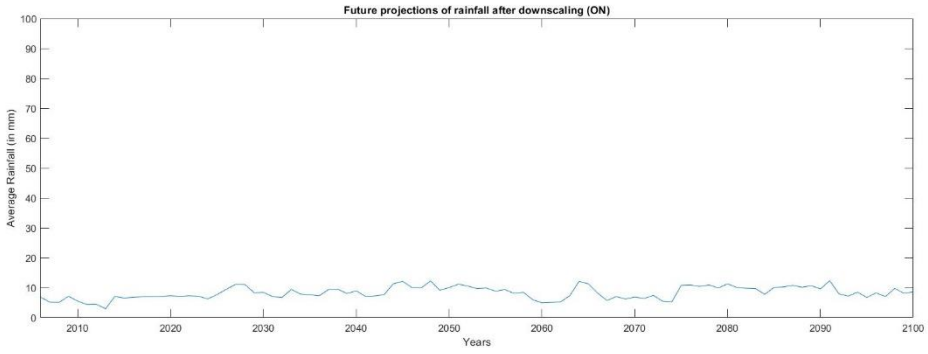


***Figure 4.24**: Projected Average Rainfall for ON for Northern Zone of India*

The projections show that the average rainfall in JJAS for non-historic period rose as high as 60 mm while for MAM and ON, it was relatively lower for Northern zone.

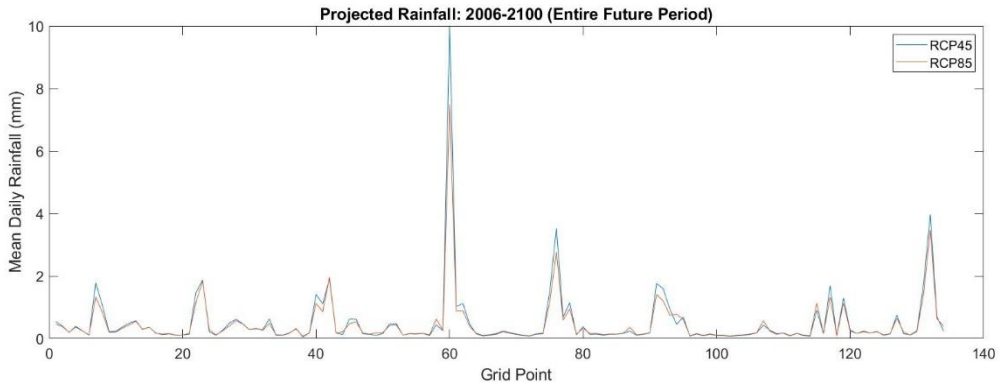


***Figure 4.25**: Projected Average Rainfall for both RCP45 and RCP85 ON for NEH Zone of India*

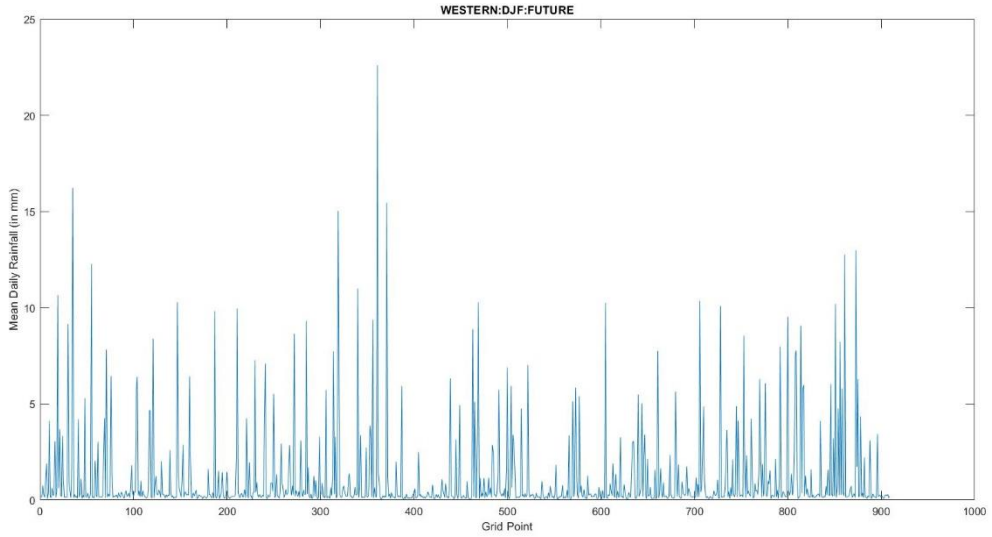


***Figure 4.26**: Projected Average Rainfall for DJF for Western Zone of India*

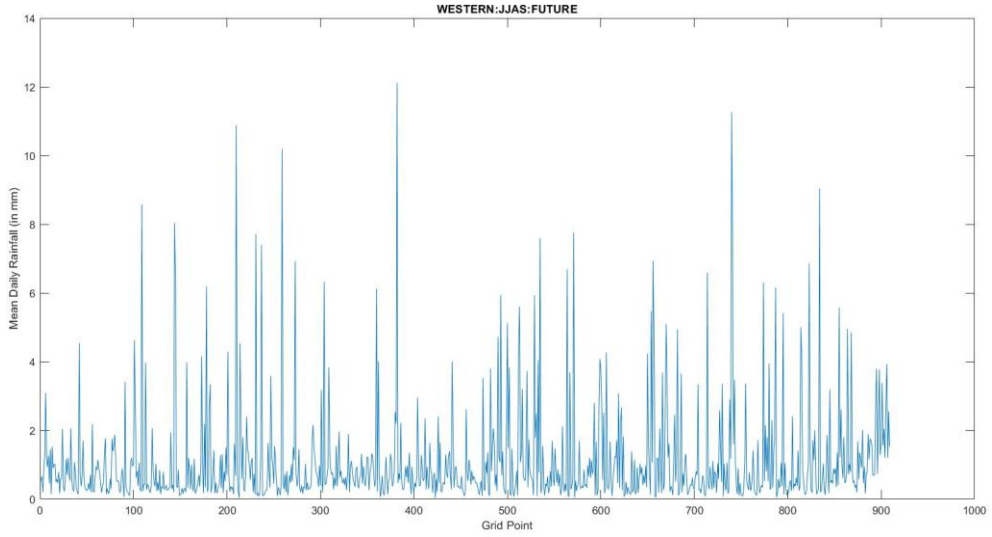


***Figure 4.27**: Projected Average Rainfall for JJAS for Western Zone of India*

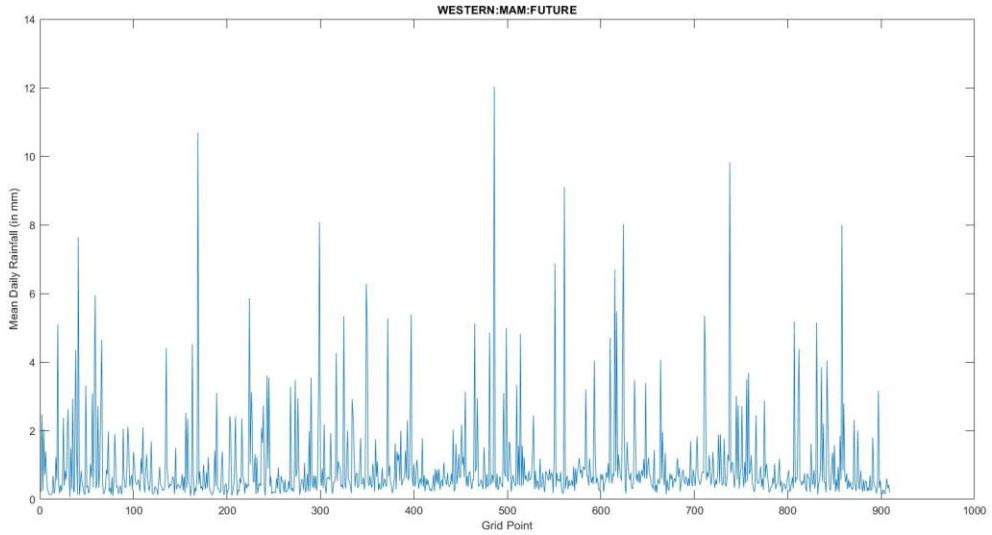


***Figure 4.28**: Projected Average Rainfall for MAM for Western Zone of India*

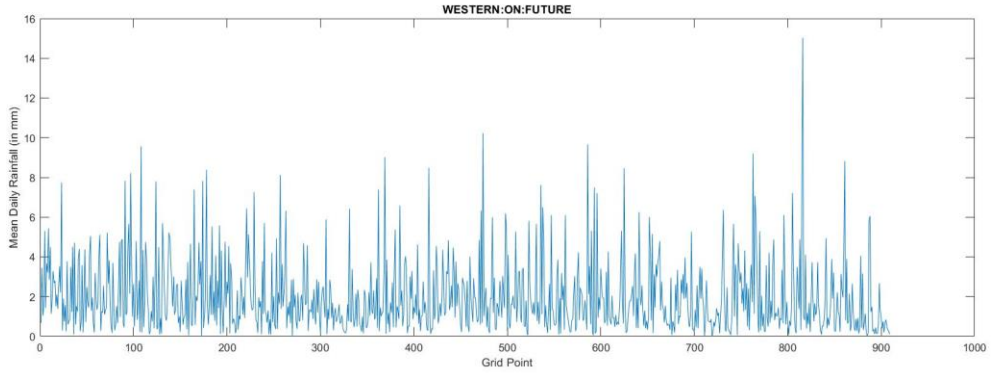


***Figure 4.29**: Projected Average Rainfall for ON for Western Zone of India*

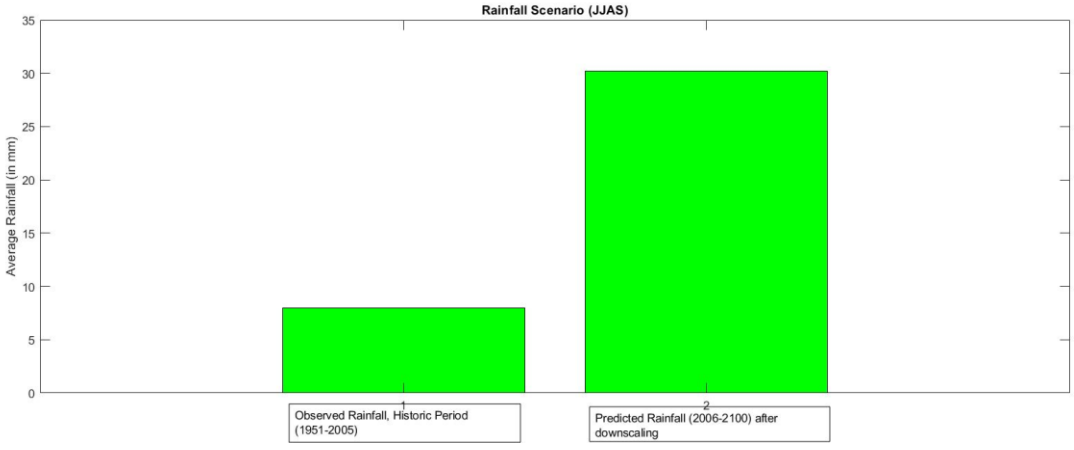


***Figure 4.30**: Average Rainfall for JJAS for Northern Zone of India*

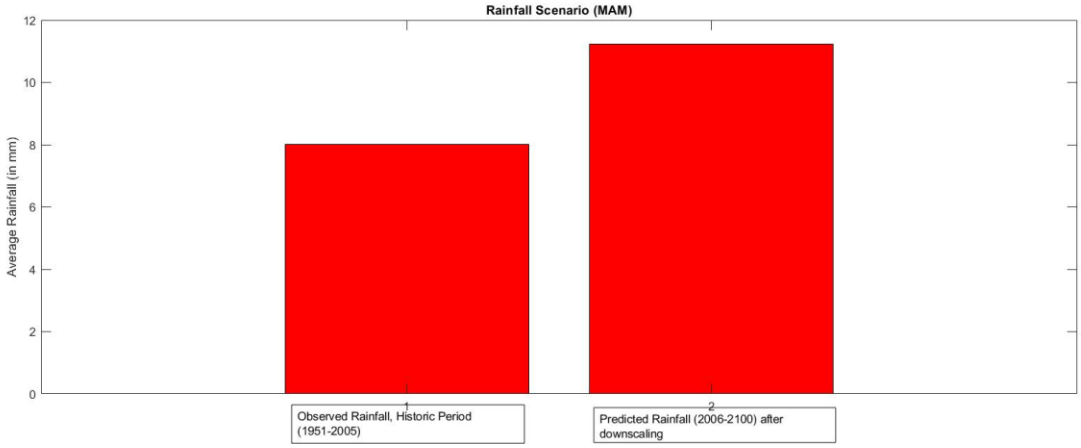


***Figure 4.31**: Average Rainfall for MAM for Northern Zone of India*

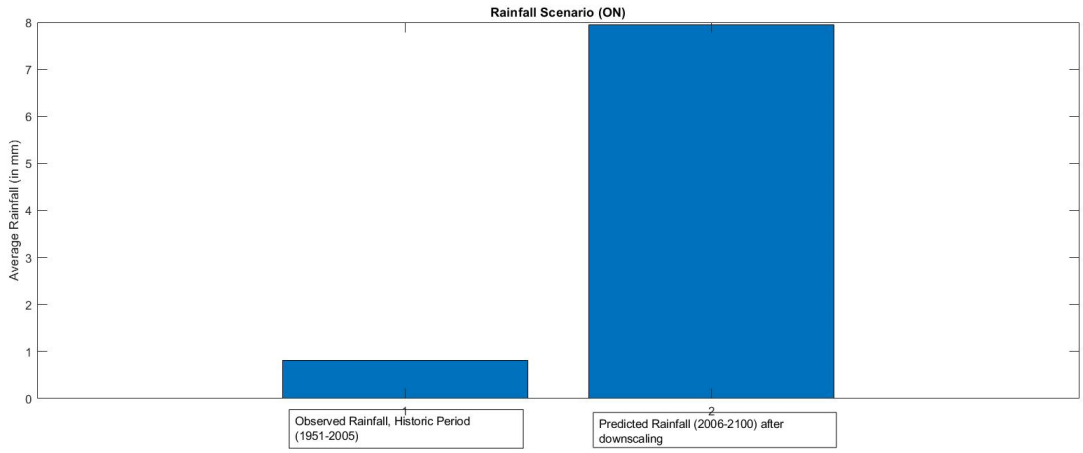


***Figure 4.32**: Average Rainfall for ON for Northern Zone of India*

The bar graphs show an increase in the mean rainfall for JJAS, MAM, ON for the future period in comparison to the historic period. The maximum increase was witnessed for JJAS (+21 mm) and the least was for MAM (+3 mm) for Northern zone of India.

4.5. Summary and Conclusion

In this study, the MLP-ANN based statistical downscaling is performed using PCA analysed, bias corrected NCEP predictors and IMD rainfall data for historical period 1951 to 2005. The projection of rainfall for future period (2006-2100) is also carried out using the developed MLP-ANN model and five different GCMs outputs.

The findings of this study suggest that the MLP-ANNs are good choice as a downscaling technique to capture the variations of rainfall in India. The correlation coefficient between

calculated and observed rainfall for most of the grid points were under acceptable limit. The calculated average rainfall were quite similar to those of the observed average rainfall for most of the grid points of the study area.

Further research can focus on devising more accurate models for projection. The degree of impact of local factors on rainfall could be included to improve the accuracy further. An analysis could also be made focused on determining the impact of the rising CO_2 concentration on rainfall.

Chapter 5

Bias-corrected climate projections from Coupled Model Intercomparison Project-5

5.1. Introduction

India is one of the most densely populated regions of the world. A majority of the population in India depends on agriculture for their livelihood. India is among the global hot spots that are likely to face the detrimental impacts of climate change (Diffenbaugh & Giorgi, 2012; Suarez-Gutierrez et al., 2020). Considerable changes in precipitation and temperature are projected in India that will have implications for water resources and agriculture (Immerzeel et al., 2010; Knox et al., 2012; Lobell & Burke, 2008; Turner & Annamalai, 2012). The risk of floods and droughts are likely to increase in India under the warming climate (Aadhar & Mishra, 2019, 2020; Ali et al., 2019; Arnell & Gosling, 2016; Hirabayashi et al., 2013; Winsemius et al., 2016). Both recent droughts and floods have affected a large population and caused enormous damage to agriculture and infrastructure in India (Hunt & Menon, 2020; Mishra & Shah, 2018; Webster et al., 2011). Similarly, the frequency and intensity of severe heatwaves have increased in India and projected to increase in the future (Im et al., 2017; JA et al., 2020; Mazdiyasni et al., 2017; Mishra et al., 2017; Mukherjee & Mishra, 2018). Overall, the frequency of both precipitation and temperature extremes has considerably increased in the past decades and likely to rise further under the warming climate (JA et al., 2020; Mukherjee et al., 2018).

Projections from the General Circulation Models (GCMs) play a vital role in understanding the future changes in climate. However, spatial resolution at which GCMs are run is often too coarse to get reliable projections at the regional and local scale (Christensen et al., 2008). Precipitation and temperature projections at higher spatial resolution are required for the climate impact assessments (Barbero et al., 2017; Cayan et al., 2010; Maurer et al., 2010). Moreover, precipitation and temperature from the GCMs have a bias due to their coarse resolution or model parameterizations (Ashfaq et al., 2017; Mishra et al., 2014). Therefore, for the assessment of the climate change and its impacts on different sectors (e.g., water resources, agriculture), bias-correction is required (Christensen et al., 2008; Eisner et al., 2012; Maraun et al., 2017; Piani et al., 2010; Pierce et al., 2015; Thrasher et al., 2012; A. W. Wood et al., 2004). Both statistical and dynamical approaches are used for downscaling and bias correction of climate change projections from GCMs. Statistical approaches are based on the distribution and relationship between the observed and projected data for the historical period (Pierce et al., 2015; Thrasher et al., 2012). On the other hand, dynamical downscaling approaches are based on regional climate model forced with the boundary conditions from the coarse resolution GCMs (Giorgi & Gutowski, 2015; Mearns et al., 2013). Both statistical and dynamical downscaling approaches have limitations (Abatzoglou & Brown, 2012; Maurer & Hidalgo, 2008). The primary limitation of the dynamical downscaling is related to the requirement of computational efforts to run the regional climate models at higher spatial and temporal resolution (Mishra et al., 2014; White & Toumi, 2013). Moreover, dynamical downscaling may not remove the bias in climate variables, which might require corrections based on the statistical approaches (White & Toumi, 2013). Given these limitations, statistical bias correction approaches are widely used in climate change impact assessments (Gutmann et al., 2014; Xu & Wang, 2019).

Considering the climate change impacts in India, we develop a bias-corrected dataset of daily precipitation, maximum and minimum temperatures using output from five GCMs that participated in the Coupled Model Intercomparison Project-5 (CMIP5). The five GCMs were selected based on the availability of daily precipitation, maximum and minimum temperatures for the historical and three scenarios (RCP 2.6, RCP 4.5, RCP 8.5). We used empirical quantile mapping (EQM) to develop bias-corrected data at daily temporal and 0.25° spatial resolution for India. The bias-corrected projections from 13 CMIP-GCMs can be used for estimating the projected changes in mean and extreme climate in India.

5.2. Methods

We obtained observed daily gridded precipitation, minimum and maximum temperatures for India for the 1951-2005 period. Daily precipitation at 0.25° was obtained from the India Meteorological Department (IMD) (Pai et al., 2014). Pai et al. (2014) developed gridded daily precipitation for India using station observations from more than 6000 stations located across India. The precipitation captures critical features of the Indian summer monsoon, including higher rainfall in the Western Ghats and northeastern India and lower rainfall in the semi-arid and arid regions of western India. Besides, gridded precipitation captures the orographic rain in the Western Ghats and foothills of Himalaya. The gridded precipitation data from IMD has been used for various hydroclimatic applications (Mishra et al., 2016; D. Shah & Mishra, 2019). Gridded daily maximum and minimum temperatures from IMD were developed using station-based observations from more than 350 stations located across India (Srivastava et al., 2009). There is bias in temperature observations from IMD in the Himalayan region, which can be attributed to sparse station density (R. Shah & Mishra, 2014). We used gridded observations for bias correction as station data are not available.

We obtained daily precipitation, maximum and minimum temperatures from five CMIP5-GCMs from https://esgf-node.llnl.gov/search/cmip5/. Precipitation, maximum and minimum temperatures from CMIP6-GCMs are available at different spatial resolutions (Table 1). For instance, the spatial resolution of the CMIP5 projection varies from 1.4° (CNRM) to more than 2.8° (CanESM5). All the three variables were selected for the historical (1850-2005) and three representative concentration pathways (RCPs) scenarios under r1i1p1f1 initial condition at daily time scale (Taylor et al., 2012). The three RCPs assume an increase of 2.6 W/m^2 (low forcing sustainability pathway; RCP-2.6), 4.5 W/m^2 (medium forcing middle of the road pathway; RCP-4.5), and 8.5 W/m^2 (high end forcing pathway; RCP-8.5) radiative forcing by the end of the 21st century. Further details on the CMIP5 can be obtained from Taylor et al. (2012). We regridded all the variables from CMIP5 to 1° spatial resolution to make them consistent. However, the effect of regridding using bilinear interpolation was checked by comparing the gridded datasets against the raw data for all-India mean of precipitation, maximum and minimum temperatures. We did not find any considerable differences in the all-India averaged precipitation and temperature using regridded and raw output from the GCMs.

Outputs of the various atmospheric (e.g., maximum and minimum temperatures, and precipitation) variables obtained from GCMs are known to exhibit systematic biases. Hence,

these outputs need to be bias-corrected to produce reliable estimates at regional and local scales for climate impact assessment. To achieve this, statistical transformations that attempt to find a function that maps the model output to a new distribution such that the resulting distribution matches that of observations. In general, this transformation can be formulated as (Piani et al., 2010):

$$x_m^o = f(x_m) \quad \text{..... (5.1)}$$

Where x_m^o is the bias-corrected model output.

If the statistical distribution of x_m and x_0 are known, the transformation can be written as:

$$x_m^o = F_0^{-1}(F_m(x_m)) \quad \text{..... (5.2)}$$

Where F_m and F_o are the Cumulative Distribution Functions (CDFs) of x_m and x_o respectively.

In Empirical Quantile Mapping [EQM, (Aadhar & Mishra, 2017; Andrew W. Wood et al., 2002)), instead of assuming parametric distributions, empirical CDFs (Jakob Themeßl et al., 2011; Julien et al., 2007; Thrasher et al., 2012) are estimated from the percentiles calculated from x_m and x_0. As a result, EQM and its variants can be applied to both temperature and precipitation even if their underlying distributions are different and hence recommended for statistical bias correction (Cannon, 2011).

In the context of statistical downscaling, since the observations are at a higher resolution than models, EQM on bilinearly interpolated model outputs at observation resolution is often used to address the scale mismatch and generate post-processed model outputs (Aadhar & Mishra, 2017). We choose non-parametric transformation approaches over the parametric approaches as has shown better skills in the comparison to parametric methods in reducing biases from GCM as well as Regional Climate Model (RCM) outputs (Gudmundsson et al., 2012).

We used EQM to statistically downscale the daily maximum and minimum temperatures, and precipitation. We use the outputs (x_m) from five CMIP5-GCMs (Table 1), which are available at different resolutions (Table 1). Observations for the three variables at the resolution of 0.25-degree are obtained from the IMD. We used the 1951-2005 period to obtain the transformation function to map the distribution of x_m to x_o. For precipitation, the drizzle effect is corrected by using a wet day threshold of 1 mm/day (Gudmundsson et al., 2012; Piani et al., 2010). If the values from model projections are larger (smaller) than the training values used to estimate the empirical CDF, the correction found for the highest (lowest) quantile of the training period is

used. We used mapped transformation to bias correct the outputs for the historical period and the RCP 2.6, RCP 4.5, and RCP 8.5 scenarios for the 2006-2100 period for all the three variables. Quantile mapping based statistical bias correction has been widely used, and its performance was found to be satisfactory in comparison to the other methods (Bürger et al., 2012; Maurer et al., 2010; Thrasher et al., 2012).

5.3. Results

The bias in mean annual precipitation, maximum and minimum temperatures was corrected against the observations from IMD. The CMIP5-GCMs show a dry bias (15-20%) in precipitation in the majority of India (Aadhar & Mishra, 2019). A high cold and warm bias in both maximum and minimum temperatures were found in the CMIP5-GCMs (Aadhar & Mishra, 2019). We applied the EQM approach to correct the bias in the CMIP5-GCM output at daily timescale. The bias in precipitation and temperatures extremes and mean in the raw output was corrected separately. The 90^{th} percentiles of precipitation of rainy days (precipitation more than 1mm), maximum and minimum temperatures and monthly mean data were corrected for the historical period (1951-2006) from CMIP5-GCMs against the observed dataset. We find that the EQM based bias correction has successfully removed the bias in precipitation, maximum and minimum temperatures across India (Fig. 1). We compared the season cycle of bias-corrected precipitation, maximum and minimum temperatures from the CMIP5-GCMs against the observed dataset for the 1951-2005 period. We find that the seasonal cycle of the bias-corrected precipitation, maximum, and minimum temperatures compare well against the observations (Fig. 1). The bias was substantially reduced after the bias correction in all the three variables for the historical (1951-2005) period (Fig. 1a-c). The reduction in bias in precipitation, maximum and minimum temperatures shows the effectiveness of our bias correction approach based on EQM.

Overall, our results show that the EQM approach successfully corrects the bias in the CMIP5-GCMs, which can be used for climate impacts studies in India. Also, the bias-corrected dataset can be used for hydrological studies in the Indian sub-continental river basins.

The bias corrected precipitation and minimum and maximum temperature datasets of CMIP5 are submitted to IIT Bombay.

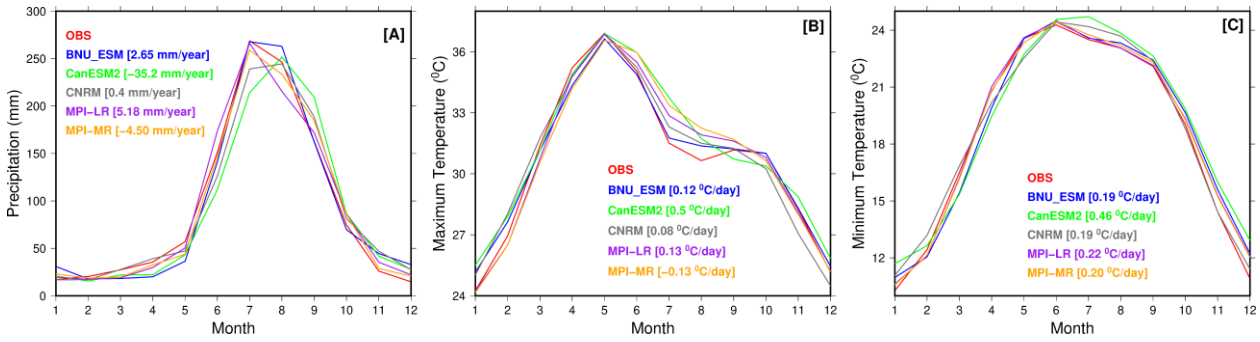


Figure 5.1*: Seasonal cycle of bias-corrected precipitation, maximum and minimum temperatures. Comparison of the each CMIP5-GCM mean seasonal cycle of bias-corrected (a) precipitation, (b) maximum temperature, and (c) minimum temperature against the observations for the 1951-2005 period (red).*

Table 5.1: List of the General Circulation Models (GCMs) used in the study. RCP 2.6, RCP 4.5 and RCP 8.5 are used for all the GCMs.

Model	**Modeling Group**	**Native resolution**	
		Longitude	**Latitude**
BNU-ESM	College of Global Change and Earth System Science, Beijing Normal University (GCESS)	2.8	2.8
CanESM2	Canadian Centre for Climate Modelling and Analysis (CCCMA)	2.8	2.8
CNRM	Centre National de Recherches Météorologiques Centre Européen de Recherche et Formation Avancée en Calcul Scientifique	1.4	1.4
MPI-ESM-LR	Max-Planck-Institut für Meteorologie (MPI-M)	1.875	1.875
MPI-ESM-MR		1.875	1.875

Chapter 6

Uncertainty Analysis

No climate model and downscaling technique is perfect in simulating various climate features. Uncertainty in prediction stems from varied factors such as parameterization, model structural flaws, choice of downscaling technique, selection of predictor variable etc. To address such uncertainty, climate modelers employ model ensembles i.e. to understand the average climate behavior from different downscaled outputs. Every downscaled output in these ensembles has slightly different features producing different outputs. Details on ensembles employed under this particular exercise are given in the Table 6.1.

Table 6.1 Ensembles used in preparing the annual precipitation time-series at each grid point.

RCP Scenario	**Downscaling method**	**Contributing institute**	**Driving AOGCM**	**Ensemble Size**
RCP 4.5	Kernel Regression	IIT Bombay	CCCma CanESM2	14 members
			CNRM CM5	
			MPI ESM MR	
			MPI ESM LR	
			BNU ESM	
	Artificial Neural Network（ANN）	IIT Guwahati	CCCma CanESM2	
			CNRM CM5	
			MPI ESM MR	
			MPI ESM LR	
			BNU ESM	
	Bias Correction Spatial	IIT Gandhinagar	CNRM CM5	
			MPI ESM MR	

	Disaggregation（BCSD）		MPI ESM LR	
			BNU ESM	
RCP 8.5	Kernel Regression	IIT Bombay	CCCma CanESM2	15 members
			CNRM CM5	
			MPI ESM MR	
			MPI ESM LR	
			BNU ESM	
	Artificial Neural Network（ANN）	IIT Guwahati	CCCma CanESM2	
			CNRM CM5	
			MPI ESM MR	
			MPI ESM LR	
			BNU ESM	
	Bias Correction Spatial Disaggregation（BCSD）	IIT Gandhinagar	CCCma CanESM2	
			CNRM CM5	
			MPI ESM MR	
			MPI ESM LR	
			BNU ESM	

In this project, to characterize the uncertainty in the downscaled outputs, for each grid point, various percentile values have been worked out for the ensemble-averaged annual precipitation time-series. Future period has been divided into different windows such as 2021-40, 2041-60, 2061-80, and 2081-2100.

For one grid point (23N latitude, and 77E longitude), the uncertainty in the ensemble-averaged annual precipitation time-series, have been explained using the box-plots as shown in the Figure 6.1. For the same grid point (23N, 77E), Figure 6.2 and 6.3 shows the time-series plots for annual rainfall projections that are derived from each downscaled output.

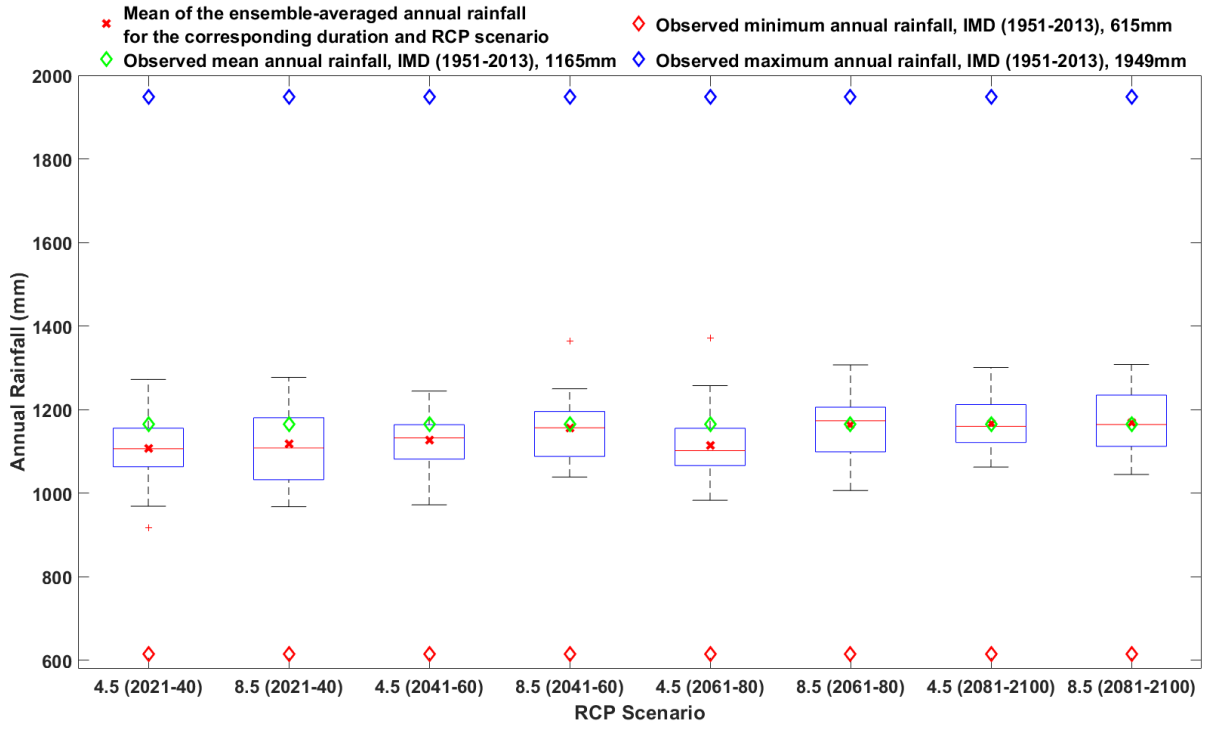


Figure 6.1: *Uncertainty in the ensemble-averaged annual rainfall at one grid point (23N, 77E).*

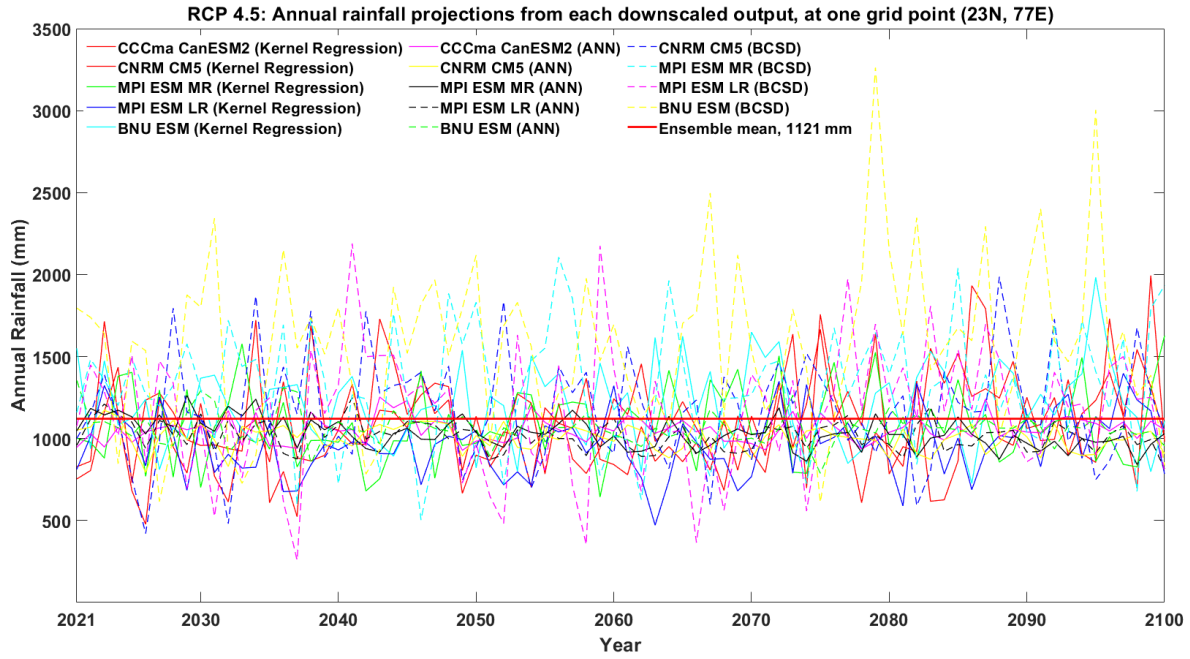


Figure 6.2: *Time-series plot for projected annual rainfall (RCP 4.5), corresponding to each downscaled output, at one grid point (23N, 77E).*

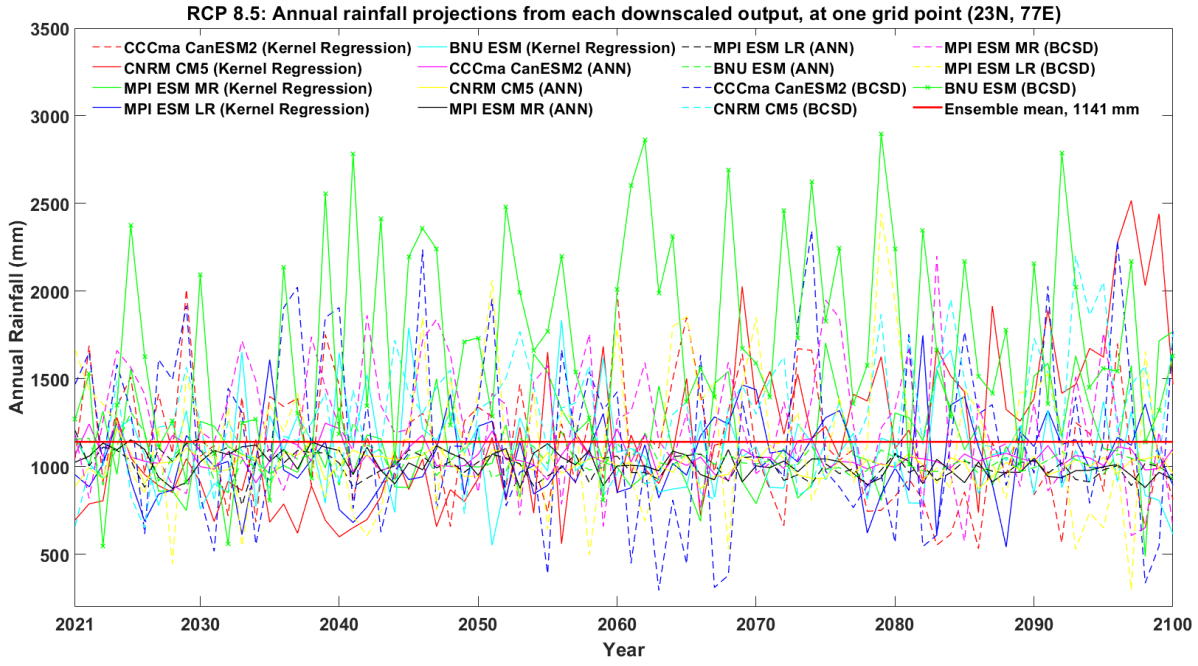


Figure 6.3: *Time-series plot for projected annual rainfall (RCP 8.5), corresponding to each downscaled output at one grid point (23N, 77E).*

For each ensemble/RCP scenario, this entire exercise was done using the MATLAB® software in following steps:

1. Downscaled datasets for projected daily precipitation is available at http://www.regclimindia.in/ in NetCDF (.nc file) format, these were converted into the .mat extension. For each downscaling method and corresponding to all AOGCM outputs, one .mat file was created to store the projected daily precipitation time-series into the cell array data type.

2. For all grid points, the daily time-series from these .mat files were converted into the annual time-series. For each downscaling method and corresponding to all AOGCM outputs, these annual time-series data were stored into one separate .mat file.

3. From this annual precipitation datasets, ensemble-averaged annual time-series were prepared at each grid point.

4. At each grid point, the ensemble-averaged annual time-series was segregated according to four different durations such as 2021-2040, 2041-2060, 2061-2080, and 2081-2100. Various percentile values were worked out for each of these segregated time-series.

For all grid points, the final outputs of this particular exercise have been arranged in a way as shown in the Figure 6.4. These are uploaded on the website for two RCP scenario (i.e. RCP 4.5 and RCP 8.5), and for each future time window (http://www.regclimindia.in/uncen.html).

[Variable: Average annual rainfall (mm/year), 2021-2040, RCP4.5]

Lat	Lon	Min	Max	10th	25th	50th	75th	90th percentile
16.25	73.25	2399	2805	2442	2521	2613	2682	2792
16.25	73.5	3139	3584	3224	3288	3454	3517	3555
16.25	73.75	2960	3313	3020	3060	3188	3267	3297
16.25	74	1418	1663	1447	1468	1540	1589	1616
16.25	74.25	858	980	885	905	937	964	973
16.25	74.5	658	782	675	697	715	742	748
16.25	74.75	504	575	512	530	538	563	571
16.25	75	464	538	466	477	489	519	530
16.25	75.25	476	554	480	488	501	533	543
16.25	75.5	537	641	549	579	590	612	627
16.25	75.75	533	637	544	569	589	603	615
16.25	76	511	609	525	538	567	586	593
16.25	76.25	476	568	491	505	525	540	558
16.25	76.5	502	611	510	536	552	577	593
16.25	76.75	553	669	560	585	605	629	637
16.25	77	498	615	504	519	539	574	580
16.25	77.25	500	635	504	522	552	580	595
16.25	77.5	524	671	532	544	572	593	618
16.25	77.75	530	684	541	560	583	607	632
16.25	78	560	694	565	576	595	641	671
16.25	78.25	521	653	531	538	559	594	632
16.25	78.5	483	621	510	521	537	568	602
16.25	78.75	572	713	587	609	621	648	688
16.25	79	578	741	604	620	648	676	705
16.25	79.25	591	777	610	636	662	712	752
16.25	79.5	682	957	700	758	804	866	905
16.25	79.75	702	891	723	761	801	839	881
16.25	80	712	954	745	778	829	871	911
16.25	80.25	793	993	811	866	895	929	974
16.25	80.5	768	957	797	826	861	907	951
16.25	80.75	787	1028	836	853	904	956	1021

***Figure 6.4**: For future time periods, the end output of this exercise have been arranged in the above format at all grid points, and for both RCPs (RCP4.5 and RCP8.5).*

Online Portal details:

The data from this project can be access through online portal. The online portal can be accessed from http://www.regclimindia.in .The data can be freely accessed and used for research purposes with some terms and conditions specified at the portal itself.

The website allows you to choose a particular method, GCM and variable to download. Choose model name under “Model and Data” tab on the website to select method and then model. This opens up the list of variables available for download. Click on desired variable and choose a

Historical or Future Scenarios. Source lists will also be available on the website for downloading multiple files.

List of variables available:

1. Downscaled Precipitation
2. Near-surface Air Temperature (TAS);
3. Air Temperature at 850hpa pressure level (T850);
4. Air Temperature at 500hpa pressure level (T500);
5. Eastward Near-surface Wind Velocity (UAS);
6. Eastward Wind Velocity at 850hpa pressure level (U850);
7. Northward Near-surface Wind Velocity (VAS);
8. Northward Wind Velocity at 850hpa pressure level (V850);
9. Specific Humidity at 850hpa pressure level (Q850);
10. Sea level Air Pressure (PSL); and
11. Geo-potential Height at 500hpa pressure level
12. Maximum Near-surface Air Temperature (T_{max})
13. Minimum Near-surface Air Temperature (T_{min})

Variables 2 to 11 are bias-corrected with respect to NCEP/NCAR data and are used as predictors for statistical downscaling. Variables 12 and 13 are bias-corrected with respect to IMD dataset.

Chapter 7

Conclusions

The research work, reported in this report contributes towards obtaining climate projections at fine resolutions, generating future patterns of hydro-meteorological variables. Focus of the present research is on temperature and rainfall, which are the most important 'impact relevant' climate variables. The major conclusions derived are

- Daily maximum, mean, and minimum temperature show positive trend indicating rise in temperature in entire India.
- The Rainfall projections for future show non-uniform changes in mean rainfall as compared to the historic period simulations. Effect of orography is clearly seen even in future changes
- There is a wide variation in projections across models and hence uncertainty modelling is a must before using the projections for planning.
- The entire outputs are made available in http://www.regclimindia.in

References

Aadhar, S., & Mishra, V. (2017). High-resolution near real-time drought monitoring in South Asia. *Scientific Data*, *4*, 170145. Retrieved from http://dx.doi.org/10.1038/sdata.2017.145

Aadhar, S., & Mishra, V. (2019). A Substantial rise in the area and population affected by dryness in South Asia under 1.5, 2.0 and 2.5□C warmer worlds. *Environmental Research Letters*, *14*, 114021. https://doi.org/10.1088/1748-9326/ab4862

Aadhar, S., & Mishra, V. (2020). Increased drought risk in South Asia under warming climate: Implications of uncertainty in potential evapotranspiration estimates. *Journal of Hydrometeorology*. https://doi.org/10.1175/jhm-d-19-0224.1

Abatzoglou, J. T., & Brown, T. J. (2012). A comparison of statistical downscaling methods suited for wildfire applications. *International Journal of Climatology*, *32*(5), 772–780. https://doi.org/10.1002/joc.2312

Ahmed, K., Shahid, S., Haroon, S.B. (2015), Multilayer perceptron neural network for downscaling rainfall in arid region: A case study of Baluchistan, Pakistan; J Earth Syst Sci (2015) 124: 1325. https://doi.org/10.1007/s12040-015-0602-9

Ali, H., Modi, P., & Mishra, V. (2019). Increased flood risk in Indian sub-continent under the warming climate. *Weather and Climate Extremes*, *25*, 100212. https://doi.org/10.1016/j.wace.2019.100212

Arnell, N. W., & Gosling, S. N. (2016). The impacts of climate change on river flood risk at the global scale. *Climatic Change*, *134*(3), 387–401. https://doi.org/10.1007/s10584-014-1084-5

Ashfaq, M., Rastogi, D., Mei, R., Touma, D., & Ruby Leung, L. (2017). Sources of errors in the simulation of south Asian summer monsoon in the CMIP5 GCMs. *Climate Dynamics*, *49*(1–2), 193–223. https://doi.org/10.1007/s00382-016-3337-7

Barbero, R., Fowler, H. J., Lenderink, G., & Blenkinsop, S. (2017). Is the intensification of precipitation extremes with global warming better detected at hourly than daily resolutions? *Geophysical Research Letters*, *44*(2), 974–983. https://doi.org/10.1002/2016GL071917

Bishop CM (1995) Neural networks for pattern recognition. Clarendon Press, Oxford. ISBN-13: 978–0198538646.

Bürger, G., Murdock, T. Q., Werner, A. T., Sobie, S. R., & Cannon, A. J. (2012). Downscaling extremes-an intercomparison of multiple statistical methods for present climate. *Journal of Climate*, *25*(12), 4366–4388. https://doi.org/10.1175/JCLI-D-11-00408.1

Cannon, A. J. (2011). Quantile regression neural networks: Implementation in R and application to precipitation downscaling. *Computers and Geosciences*, *37*(9), 1277–1284. https://doi.org/10.1016/j.cageo.2010.07.005

Cayan, D. R., Das, T., Pierce, D. W., Barnett, T. P., Tyree, M., & Gershunova, A. (2010). Future dryness in the Southwest US and the hydrology of the early 21st century drought. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(50), 21271–21276. https://doi.org/10.1073/pnas.0912391107

Christensen, J. H., Boberg, F., Christensen, O. B., & Lucas-Picher, P. (2008). On the need for bias correction of regional climate change projections of temperature and precipitation. *Geophysical Research Letters*, *35*(20). https://doi.org/10.1029/2008GL035694

Climate Change in Australia: Downscaling https://www.climatechangeinaustralia.gov.au/en/

Climate Models, https://www.climate.gov/

Coulibaly P, Anctil F, Bobée B (2000), Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. J Hydrol 230:244–257. doi:10.1016/S0022-1694(00)00214-6.

C.R. Mechoso, A. Arakawa, in Encyclopedia of Atmospheric Sciences (Second Edition), 2015.

Diffenbaugh, N. S., & Giorgi, F. (2012). Climate change hotspots in the CMIP5 global climate model ensemble. *Climatic Change*, *114*(3–4), 813–822. https://doi.org/10.1007/s10584-012-0570-x

Eisner, S., Voss, F., & Kynast, E. (2012). Statistical bias correction of global climate projections - Consequences for large scale modeling of flood flows. *Advances in Geosciences*, *31*, 75–82. https://doi.org/10.5194/adgeo-31-75-2012

Flato, G., J. Marotzke, B. Abiodun, P. Braconnot, S.C. Chou, W. Collins, P. Cox, F. Driouech, S. Emori, V. Eyring, C. Forest, P. Gleckler, E. Guilyardi, C. Jakob, V. Kattsov, C. Reason and M. Rummukainen, 2013: Evaluation of Climate Models. In: Climate Change (2013), The Physical Science Basis. Contribution of Working Group I to the Fifth Assess-ment 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

Giorgi, F., & Gutowski, W. J. (2015). Regional Dynamical Downscaling and the CORDEX Initiative. *Annual Review of Environment and Resources*, *40*(1), 467–490. https://doi.org/10.1146/annurev-environ-102014-021217

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., & Engen-Skaugen, T. (2012). Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations – A comparison of methods. *Hydrology and Earth System Sciences*, *16*(9), 3383–3390. https://doi.org/10.5194/hess-16-3383-2012

Gutmann, E., Pruitt, T., Clark, M. P., Brekke, L., Arnold, J. R., Raff, D. A., & Rasmussen, R. M. (2014). An intercomparison of statistical downscaling methods used for water resource assessments in the United States. *Water Resources Research Research*, *50*, 7167–7186. https://doi.org/10.1002/2014WR015559.Received

Haykin S (1994) Neural networks: a comprehensive foundation.Macmillan College. ISBN 13: 9780023527616

Hewitson BC, Crane RG (1996) Climate downscaling: techniques and application. Clim Res 7:85–96.

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., et al. (2013). Global flood risk under climate change. *Nature Climate Change*, *3*(9), 816–821. https://doi.org/10.1038/nclimate1911

Hunt, K. M. R., & Menon, A. (2020). The 2018 Kerala floods: a climate change perspective. *Climate Dynamics*, *54*(3–4), 2433–2446. https://doi.org/10.1007/s00382-020-05123-7

Im, E. S., Pal, J. S., & Eltahir, E. A. B. (2017). Deadly heat waves projected in the densely populated agricultural regions of South Asia. *Science Advances*. https://doi.org/10.1126/sciadv.1603322

Immerzeel, W. W., van Beek, L. P. H., & Bierkens, M. F. P. (2010). Climate change will affect the Asian water towers. *Science (New York, N.Y.)*, *328*(5984), 1382–5. https://doi.org/10.1126/science.1183188

JA, N., van der Wiel, K., Bhatia, U., Stone, D., Selten, F. M., & Mishra, V. (2020). A seven-fold rise in the probability of exceeding the observed hottest summer in India in a 2°C warmer world. *Environmental Research Letters*. https://doi.org/10.1088/1748-9326/ab7555

Jain, A. and Singla, S. (2014), Modeling monsoon rainfall using single and multiple hidden layer artificial neural network models, Intl. J. Wat. Resour. & Env. Mgmt., 5(1-2), 127-140.

Jakob Themeßl, M., Gobiet, A., & Leuprecht, A. (2011). Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. *International Journal of Climatology*, *31*(10), 1530–1544. https://doi.org/10.1002/joc.2168

Julien, B., L., T., F., H., & E., M. (2007). Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies. *International Journal of Climatology*, *27*(August), 1643–1655. https://doi.org/10.1002/joc

Knox, J., Hess, T., Daccache, A., & Wheeler, T. (2012). Climate change impacts on crop productivity in Africa and South Asia. *Environmental Research Letters*, *7*(3). https://doi.org/10.1088/1748-9326/7/3/034032

Lee Hannah, Climate Change Biology (Second Edition), 2015.

Lobell, D. B., & Burke, M. B. (2008). Why are agricultural impacts of climate change so uncertain? the importance of temperature relative to precipitation. *Environmental Research Letters*, *3*(3). https://doi.org/10.1088/1748-9326/3/3/034007

Maraun, D., Shepherd, T. G., Widmann, M., Zappa, G., Walton, D., Gutiérrez, J. M., et al. (2017). Towards process-informed bias correction of climate change simulations. *Nature Climate Change*, *7*(11), 764–773. https://doi.org/10.1038/nclimate3418

Martin TH, Howard BD, Mark B, Orlando DJ (1996) Neural networks design. PWS Publishing Company. ISBN-13: 978–0–9717321–1–7.

Maurer, E. P., & Hidalgo, H. G. (2008). Utility of daily vs. monthly large-scale climate data: an intercomparison of two statistical downscaling methods. *Hydrology and Earth System Sciences*, *14*(6), 1125–1138. https://doi.org/10.5194/hess-14-1125-2010

Maurer, E. P., Hidalgo, H. G., Das, T., Dettinger, M. D., & Cayan, D. R. (2010). The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. *Hydrology and Earth System Sciences*, *14*(6), 1125–1138. https://doi.org/10.5194/hess-14-1125-2010

Mazdiyasni, O., AghaKouchak, A., Davis, S. J., Madadgar, S., Mehran, A., Ragno, E., et al. (2017). Increasing probability of mortality during Indian heat waves. *Science Advances*, *3*(6), 1–6. https://doi.org/10.1126/sciadv.1700066

Mearns, L. O., Sain, S., Leung, L. R., Bukovsky, M. S., McGinnis, S., Biner, S., et al. (2013). Climate change projections of the North American Regional Climate Change Assessment

Mishra, V., & Shah, H. L. (2018). Hydroclimatological Perspective of the Kerala Flood of 2018. *JOURNAL GEOLOGICAL SOCIETY OF INDIA*, *92*, 645–650. https://doi.org/10.1007/s12594-018-1079-3

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

Mishra, V., Aadhar, S., Asoka, A., Pai, S., & Kumar, R. (2016). On the frequency of the 2015 monsoon season drought in the Indo-Gangetic Plain. *Geophysical Research Letters*, *43*(23), 12,102-12,112. https://doi.org/10.1002/2016GL071407

Mishra, V., Mukherjee, S., Kumar, R., & Stone, D. A. (2017). Heat wave exposure in India in current, 1.5 °C, and 2.0 °C worlds. *Environmental Research Letters*, *12*(12), 124012. https://doi.org/10.1088/1748-9326/aa9388

Mukherjee, S., & Mishra, V. (2018). A sixfold rise in concurrent day and night-time heatwaves in India under 2 °C warming. *Scientific Reports*, *8*(1), 16922. https://doi.org/10.1038/s41598-018-35348-w

Mukherjee, S., Aadhar, S., Stone, D., & Mishra, V. (2018). Increase in extreme precipitation events under anthropogenic warming in India. *Weather and Climate Extremes*. https://doi.org/10.1016/j.wace.2018.03.005

N. Murata, S. Yoshizawa, and S. Amari, (1993) Learning curves, model selection and complexity of neural networks,‖ in Advances in Neural Information Processing Systems 5, S. Jose Hanson, [9] J. D. Cowan, and C. Lee Giles, ed. San Mateo, CA: Morgan Kaufmann, pp. 607-614.

Pai, D. S., Sridhar, L., Rajeevan, M., Sreejith, O. P., Satbhai, N. S., & Mukhopadyay, B. (2014). Development of a new high spatial resolution (0 . 25 ° × 0 . 25 °) Long Period (1901-2010)

daily gridded rainfall data set over India and its comparison with existing data sets over the region. *Mausam*, *65*(1), 1–18.

Piani, C., Haerter, J. O., & Coppola, E. (2010). Statistical bias correction for daily precipitation in regional climate models over Europe. *Theoretical and Applied Climatology*, *99*(1–2), 187–192. https://doi.org/10.1007/s00704-009-0134-9

Pierce, D. W., Cayan, D. R., Maurer, E. P., Abatzoglou, J. T., & Hegewisch, K. C. (2015). Improved bias correction techniques for hydrological simulations of climate change. *Journal of Hydrometeorology*, *16*(6), 2421–2442. https://doi.org/10.1175/JHM-D-14-0236.1

Program (NARCCAP). *Climatic Change*, *120*(4), 965–975. https://doi.org/10.1007/s10584-013-0831-3

Rajashekhar S. Laddimath,Nagraj S. Patil, (2019), MAPAN-Journal of Metrology Society of India 34(9) DOI: 10.1007/s12647-018-00299-0.

Salvi, K., Kannan, S., and Ghosh, S. (2013), High-resolution multisite daily rainfall projections in India with statistical downscaling for climate change impacts assessment, Journal of Geophysical Research Atmospheres,118, 3557–3578.

Shah, D., & Mishra, V. (2019). Integrated Drought Index (IDI) for drought monitoring and assessment in India. *Water Resources Research*. https://doi.org/10.1029/2019wr026284

Shah, R., & Mishra, V. (2014). Evaluation of the Reanalysis Products for the Monsoon Season Droughts in India. *Journal of Hydrometeorology*, *15*(4), 1575–1591. https://doi.org/10.1175/JHM-D-13-0103.1

Sonali B. Maind, Priyanka Wankar, (2014), International Journal on Recent and Innovation Trends in Computing and Communication Volume: 2 Issue: 1.

Srivastava, A. K., Rajeevan, M., & Kshirsagar, S. R. (2009). Development of a high resolution daily gridded temperature data set (1969 – 2005) for the Indian region. *Atmospheric Science Letters*, *10*(October), 249–254. https://doi.org/10.1002/asl

Suarez-Gutierrez, L., Müller, W. A., Li, C., & Marotzke, J. (2020). Dynamical and thermodynamical drivers of variability in European summer heat extremes. *Climate Dynamics*, *54*(9), 4351–4366. https://doi.org/10.1007/s00382-020-05233-2

Subramanya, K. (2013), Engineering Hydrology, 4e, Tata McGraw-Hill Education.

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*. https://doi.org/10.1175/BAMS-D-11-00094.1

Thrasher, B., Maurer, E. P., McKellar, C., & Duffy, P. B. (2012). Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping. *Hydrology and Earth System Sciences*, *16*(9), 3309–3314. https://doi.org/10.5194/hess-16-3309-2012

Turner, A. G., & Annamalai, H. (2012). Climate change and the South Asian summer monsoon. *Nature Climate Change*. https://doi.org/10.1038/nclimate1495

Vu, M. T., Aribarg, T., Supratid, S., Raghavan, S. V., and Liong, S-Y. (2016), Statistical downscaling rainfall using artificial neural network: significantly wetter Bangkok?, Theor Appl Climatol, 126:453–467.

Vidyarthi, V. K., and Jain, A. (2020), Knowledge Extraction from Trained ANN Drought Classification Model, Journal of Hydrology 585:124804 DOI: 10.1016/j.jhydrol.2020.124804.

Webster, P. J., Toma, V. E., & Kim, H. M. (2011). Were the 2010 Pakistan floods predictable? *Geophysical Research Letters*, *38*(4), 1–5. https://doi.org/10.1029/2010GL046346

White, R. H., & Toumi, R. (2013). The limitations of bias correcting regional climate model inputs. *Geophysical Research Letters*, *40*(12), 2907–2912. https://doi.org/10.1002/grl.50612

Winsemius, H. C., Aerts, J. C. J. H., Van Beek, L. P. H., Bierkens, M. F. P., Bouwman, A., Jongman, B., et al. (2016). Global drivers of future river flood risk. *Nature Climate Change*. https://doi.org/10.1038/nclimate2893

Wood, A. W., Leung, L. R., Sridhar, V., & Lettenmaier, D. P. (2004). Hydrologic Implications of Dynamical and Statistical Approaches to Downscaling Climate Model Outputs. *Climatic Change*, *62*(1–3), 189–216. https://doi.org/10.1023/B:CLIM.0000013685.99609.9e

Wood, Andrew W., Maurer, E. P., Kumar, A., & Lettenmaier, D. P. (2002). Long-range experimental hydrologic forecasting for the eastern United States. *Journal of Geophysical Research D: Atmospheres*, *107*(20), 1–15. https://doi.org/10.1029/2001JD000659

Xu, L., & Wang, A. (2019). Application of the Bias Correction and Spatial Downscaling Algorithm on the Temperature Extremes From CMIP5 Multimodel Ensembles in China. *Earth and Space Science*, *6*(12), 2508–2524. https://doi.org/10.1029/2019EA000995

Appendix 5: Annual Progress Report and Service Requests

For the financial year: 2019-2020
Name of the Institute: Indian Institute of Technology, Bombay
Title of the scheme: Statistical Downscaling for hydro-climatic Projections with CMIP5 simulations to assess Impacts of Climate Change
Financial Progress:

Year	Opening Balance	Grant Received	Total	Expenditure	End Balance
2019-20	1,03,13,182.00	Nil	1,03,13,182.00	95,37,129	7,76,053

Head wise Expenditure:

Head of Expenditure	Expenditure till the end of previous year	Expenditure during the current year	Total
Remuneration/Emoluments of Manpower	274677	400383	675060
TE	30546	0	30546
Infrastructure	281455	9136746	9418201
Experimental Charges	0	0	0
Total	586678	9537129	10123807

Man months utilization: For each of the research staff, indicate following:

Designation	Man-months sanctioned	Cumulative Man-months utilised	Is the work done proportionate to man-months utilised? (Yes/No)
Pushpendra (Junior Research Fellow)	2 years	18-07-2018 to 10-10-2018	YES
		11-10-2018 to 07-01-2019	
		08-01-2019 to 03-04-2019	
		04-04-2019 01-07-2019	
		02-07-2019 to 25-07-2019	
VikramSingh Chandel (Junior Research Fellow)	2 years	24-07-2019 to 16-10-2019	YES
		17-10-2019 to 07-01-2020	

Work Progress

Indicate the progress against the agreed Work Schedule (As per the item 16 of the sanctioned proposals)

Funds to be released: Amount Rs.17,67,977/-

Date of Completion: (a) Scheduled Date 31/03/2020

(b) Expected Date –31/03/2020

If there is a delay then state reasons for delay.

Any **Revision of Estimate**? If yes, then enclose details with justifications.

Enclosures (tick): UC Appendix-7 Narrative Progress Report

Subimal Ghosh.

Signature of Principal Investigator

Name Subimal Ghosh

Date 28th September 2020

Appendix 6: Utilization Certificate

FORM GFR 19- A

Utilization Certificate 2019-2020

[See Rule 212 (1)]

Sl. No.	Letter No. and date	Amount
1.	Nil	Nil
	Total	Nil

Certified that out of **Rs.Nil** of grants-in-aid sanctioned during the year 2019-2020 in favour of Registrar IIT Bombay Under this Ministry / Department Letter No. given in the margin and **Rs.1,03,13,182.00/-(Rupees One Crore Three Lakh Thirteen Thousand One Hundred Eighty Two only)** on account of unspent balance of the previous year, a sum of Rs.95,37,129/- (Rupees Ninety Five Lakhs Thirty Seven Thousand One Hundred Twnety Nine Only) has been utilized for the purpose of Project for which it was sanctioned and that the balance of **Rs.7,76,053/- (Rupees Seven Lakh Seventy Six Thousand Fifty Three Only)**

2. Certified that I have satisfied myself that the conditions on which the grants-in-aid was sanctioned have been duly fulfilled / are being fulfilled and that I have exercised the following checks to see that the money was actually utilised for the purpose for which it was sanctioned.

Kinds of checks exercised

1. Vouchers
2. Bank Book
3. Cash Book

28/9/2020

Signature .....
Name .....
Designation .....
Date .....

सह.संकायाध्यक्ष, अनुसंधान एवं विकास
कार्य. निदेशक, आय आय टी मुंबई
Associate Dean, Research and Development
Offg. Director, IIT Bombay
पवई, मुंबई / Powai, Mumbai - 76

25/9

Salary details of Project "RD/0116-MWR0004-001"
Staff Pushpendra(30002839)

Month	Consolidated	HRA	Total	
07/18	11290	3387	14677	14 days
08/18	25000	7500	32500	
09/18	25000	7500	32500	
10/18	25000	7500	32500	
11/18	25000	7500	32500	
12/18	25000	7500	32500	
01/19	25000	7500	32500	
02/19	25000	7500	32500	
03/19	25000	7500	32500	
04/19	25000	7500	32500	
05/19	25000	7500	32500	
06/19	25000	7500	32500	
07/19	25000	7500	32500	
08/19	25000	7500	32500	
09/19	20833	6250	27083	25 days
	357123	107137	464260	

D.O.J - 18/07/2018
D.O.R - 25/09/2019

Breif
Pushpendra 464260
VikramSingh 210800
Total sal paid 675060

D.O.J - 24/07/19
Staff VikramSingh Chandel(30003435)

Month	Consolidated	HRA	Total	
07/19	8000	1920	9920	8 days
08/19	31000	7440	38440	
09/19	31000	7440	38440	
10/19	31000	7440	38440	
11/19	31000	7440	38440	
12/19	31000	7440	38440	
01/20	7000	1680	8680	7 days
	170000	40800	210800	

D.O.J - 24/07/2019
D.O.R - 07/01/2020

Appendix-7: Statement of Equipment Purchased

Assets acquired wholly or substantially out of Government grants Register maintained by grantee Institution
Name of Sanctioning Authority **Ministry of Water Resource, RD & GR**

Serial No	
Name of Grantee Institution	Indian Institute of Technology, Bombay
No. and date of sanction	16/22/2016-R&D/3003-3027 Dated: November 7, 2016
Amount of the sanctioned grant	1,04,00,000.00
Brief purpose of the grant	Statistical Downscaling for hydro-climatic Projections with CMIP5 simulations to assess Impacts of Climate Change
Whether any condition regarding the right of ownership of Government in the property or other assets acquired out of the grant was incorporated in the grant-in-aid sanction.	No.
Particulars of assets actually Credited or acquired	As per list attached
Value of the Assets as on (Date)	As per list attached
Purpose for which utilised at Present	Research
Encumbered or not	No
Reasons if encumbered	--
Disposed of or not	--
Reasons and authority, if any, for disposal	--
Amount realised on disposal	--
Remarks	--

Subimal Ghosh.

Principal Investigator

Subimal Ghosh
28th September 2020

28/9/2020

Head of the Department

कृते निदेशक, आय आय टी मुंबई
Associate Dean, Research and Development
For Director, IIT Bombay
पवई, मुंबई / Powai, Mumbai - 76

Details of Equipment Procured During the Period

S. No.	Name of Party and Equipment	Sanctioned Cost	Actual Cost (factual figure, not in lakh)	No. Date of Invoice	No. and Date of Purchase Order
1	2	3	4	5	6
1	M/s.Unicomp InfoSolutions Pvt Ltd. (External hard disk for storing data)		12980	4845/17-18 dt. 28.03.2018	Nil
2	Memory Cards Sandisk		3300	650/18-19 14.06.2018	Nil
3	portable External Hard Drive		8999	BOM-1 325425 20.07.2018	Nil
4	Backup Plus Hub USB 3.0 Desktop 3.5 External inch External Hard Drive for PC and Mac		8999	BOM-4 7754745 20.07.2018	Nil
5	M/s.Unicomp InfoSolutions Pvt Ltd. (Printer)		5100	1542/18-19 dt. 24.08.2018	Nil
6	M/s.Unicomp InfoSolutions Pvt Ltd. HP Pavilion All-in-One-27-qa180in		139240	2067/18-19 dt.29.09.2018	4500005635 dt.10.09.2018
7	M/s.Unicomp InfoSolutions Pvt Ltd. Dlink 8 port 10/100 DES 1008A		708	2107-18-19 dt.05.10.2018	Nil
8	M/s.Unicomp InfoSolutions Pvt Ltd. Harddisk 4 Tb Seagate Sata 2.5" Backup plus		9617	1395/18-19 dt.10.08.2018	Nil
9	Unicomp InfoSolutions Pvt Ltd Dell Desktop 7060MT Core i7 DELL 22" MONITOR S2218h	Rs.1,04,00,000/-	92512	4004/18-19 dt.07.03.2019	4500008709 dt.22.02.2019
10	"Unicomp InfoSolutions Pvt Ltd" Harddisk 4 Tb Seagate sata		9727	0198/19-20 dt. 24.04.2019	Nil
11	"Unicomp InfoSolutions Pvt Ltd" DDR4 Kingston Hyperx 8 GB Ram		13010	0323/19-20 dt.	Nil
12	Locuz Enterprise Solutions Ltd Sing Upgradation of Existing HPC cluster DAP Charges		4037758	0505019TS0000906 dt. 19.07.2019	4500008926
13	Locuz Enterprise Solutions Ltd Singapore Branch Computer Node Storage Node		4354406	0505019OS000734 dt. 11.03.2019	4500005985
14	Locuz Enterprise Solutions Limited, Pune Installation and Commision Charges		460200	EP/SER/112/19-20 dt. 08.07.2019	4500008953
15	Locuz Enterprise Solutions Limited PDU cable and Power Cable for workstation		11800	EP/048/19-20 dt.118.11.2019	Nil
16	Locuz Enterprise Solutions Limited, Pune Domain Cloud Storage		249845	EP/SER/152/19-20 dt. 21.08.2019	4500010321
Total			9418201		

Subimal Ghosh.

Signature of PI

Subimal Ghosh
28th September 2020

Signature of Finance Officer /Chartered Accountant

सहायक कुलसचिव/Assistant Registrar
परियोजना एवं परामर्श लेखा
Project & Consultation Accounts
संकायाध्यक्ष, अनुसंधान एवं विकास कार्यालय
Dean (R & D) Office, IRCC
भारतीय प्रौद्योगिकी संस्थान मुंबई
Indian Institute of Technology Bombay,
पवई, मुंबई/Powai, Mumbai - 400 076

28/9/2020

Signature of Head of Institute

सह.संकायाध्यक्ष, अनुसंधान एवं विकास
कृते निदेशक, आय आय टी मुंबई
Associate Dean, Research and Development
For Director, IIT Bombay
पवई, मुंबई / Powai, Mumbai - 76

Annexure-II

Request for Annual Installment with up to date Statement of Expenditure

(Year Means Financial Year i.e 01.04.2020 to 31.03.2021)

1	Sanction Order No. & Date :	16/22/2016-R&D/3003-3027 Dated 07.11.2016
2	Total Project Cost :	Rs. 1,01,16,464/-
3	Sanctioned/Revised Project Cost :	Nil
4	Data of Commencement :	07.03.2018
5	Statement of Expenditure Month wise expenditure Incurred during (01.04.2020 to 31.03.2021)	

Month & Year	Expenditure Incurred/ Committed
Apr-20	-
May-20	-
Jun-20	-
Jul-20	-
Aug-20	848,323
Sep-20	-
Oct-20	-
Nov-20	-
Dec-20	-
Jan-21	-
Feb-21	-
Mar-21	-
Total	**848,323**

6. Grant received in each year

a	1st Year	8,910,800
b	2nd Year	-
c	3rd Year	-
d	Interest, 2017-2018	20,642
	Interest, 2018-2019	290,975
	Interest, 2019-2020	255,647
	Interest, 2020-21	22,459
	Total (a+b+c+d)	**9,500,523**

Note:

1. Expenditure under the sanctioned heads at any point of time, should not exceed funds allocated under the head without prior approval of MOWR i.e. figures in column (VII) should not exceed corresponding figures in column (iii)]

2. Utilisation Certificate for each financial year ending 31st March has to be enclosed, along with request for carry forward permission to next year

Signature

Annexure IInd Continued

Statement of Expenditure

Statement showing the grant received and the Expenditure incurred during the period 01.04.2019 to 31.03.2021

r. No.	Sanctioned Heads	Fund Allocated (indicate sanctioned or revised)	Expenditure Incurred				Total Expenditure IV+V+VI+VII	Balance as on 31.03.2021	Requirement of Funds upto 31st March next year	Remarks (if any)
			1st Year 07/03/2018 to 31/03/2018	2nd Year 01/04/2018 to 31/03/2019	3rd Year 01/04/2019 to 31/03/2020	4th Year 01/04/2020 to 31/03/2021				
(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)=(III)-(VIII)		
1	Salary	936,000	-	243,300	77,419	-	320,719	**615,281**		
2	Travel Expenditure	100,000	-				-	**100,000**		
3	Infrastructure/Equipment	8,600,000	-	579,000	6,786,588	848,323	8,213,911	**386,089**		
4	Experimental Charges	-	-				-	-		
5	Contingencies	30,029	-				-	**30,029**		
6	Overhead	450,435	300,290	150,145	-	-	450,435	-		
	Total Expenditure	**10,116,464**	**300,290**	**972,445**	**6,864,007**	**848,323**	**8,985,065**	**1,131,399**	-	

Vimal Mishra,
Name and Signature of Principal Investigator
Date:08-05-2021

Signature of Competent Financial Authority (with Seal):
Date:

10/05/2021
उप कुलसचिव (वित्त एवं लेखा)
भारतीय प्रौद्योगिकी संस्थान गांधीनगर
पालज, गांधीनगर

GFR-12 A

FORM OF UTILIZATION CERTIFICATE

UTILIZATION CERTIFICATE FOR THE YEAR 2020 - 2021 in respect of recurring

GRANTS-IN-AID/SALARIES/CREATION OF CAPITAL ASSETS

1 Name of the Scheme : Statistical Downscaling for hydro-climatic projections with CMIP5 simulations to assess Impacts of Climate Change

2 Whether recurring or non recurring grants **Non Recurring & Recurring Grant**

3 Grants position of the beginning of the Financial year

i) Cash in Hand/Bank Rs. 13,41,322/-

ii) Unadjusted advances Rs. 0/-

iii) Total Rs. 13,41,322/-

4 Details of grants received, expenditure incurred and closing balances : (Actuals)

Unspent Balances of Grants received years {figure as at sl. No. 3(iii)}	Interest Earned thereon	Interest deposited back to the Government	Grant received during the year			Total available funds(1+2-3+4)	Expenditure incurred	Closing Balances (5-6)
1	2	3	4			5	6	7
			Sanction no. (i)	Date (ii)	Amount (iii)			
1,341,322	22,459	-				1,363,781	848,323	515,458

Component wise utilization of grants :

	Grant-in-aid-General		Grant-in-aid-salary	Grant-in-aid-Creation of Capital	Total
				848,323	848,323
			-		-
1.	Travel Expenditure	-			-
2.	Experimental Charges	-			-
3.	Contingencies	-			-
4.	Institutional Overhead	-			-
	TOTAL	-	-	**848,323**	**848,323**

Details of grants position at the end of the year

(i) Cash in Hand/Bank 515,458

(ii) Unadjusted Advances -

(iii) Total 515,458

Certified that I have satisfied myself that the conditions on which grants were sanctioned have been duly fulfilled/are being fulfilled and that I have exercised following checks to see that the money has been actually utilized for the purpose which it was sanctioned :

i The main accounts and other subsidiary accounts and registers (including assets registers) are maintained as prescribed in the relevant Act/Rules/Standing instructions (mention the act/Rules) and have been duly audited by designated auditors. The figures depicted above tally with the audited figures mentioned in financial statements/accounts.

ii There exist internal controls for safeguarding public funds/assets, watching outcomes and achievements of physical targets against the financial inputs, ensuring quality in asset creation etc. & the periodic evaluation of internal controls is exercised to ensure their effectiveness.

iii To the best of our knowledge and belief, no transactions have been entered that are in violation of relevant Act/Rules/standing instructions and scheme guidelines.

iv The responsibilities among the key functionaries for execution of the scheme have been assigned in clear terms and are not general in nature.

v The benefits were extended to the intended beneficiaries and only such areas/districts were covered where the scheme was intended to operate.

vi The expenditure on various components of the scheme was in the proportions authorized as per the scheme guidelines and terms and conditions of the grants-in-aid.

vii It has been ensured that the physical and financial performance under "**Statistical Downscaling for hydro-climatic projections with CMIP5 simulations to assess Impacts of Climate Change**"(name of the scheme) has been according to the requirements, as prescribed in the guidelines issued by Govt. of India and the performance/targets achieved statement for the year to which the utilization of the fund resulted in outcomes given at Annexure-I duly enclosed.

viii The utilization of the fund resulted in outcomes given at Annexure-II duly enclosed (to be formulated by the Ministry/Department concerned as per their requirements/specifications)

ix Details of various schemes executed by the agency through grants-in-aid received from the same Ministry or from other Ministries is enclosed at Annexure-II(to be formulated by the Ministry/Department concerned as per their requirements/specifications)

Date:
Place:

Signature 10/05/2021
उप कुलसचिव (वित्त एवं लेखा)
भारतीय प्रौद्योगिकी संस्थान गांधीनगर
Name.....भारतीय प्रौद्योगिकी संस्थान गांधीनगर

Deputy Registrar (F&A)
(Head of the Finance)

Signature 11.05.2021
Dean R&D
IIT Gandhinagar
Name.....संकायाध्यक्ष,आरएंडडी
आईआईटीगांधीनगर

Dean R&D

Signature

Name Vimal Mishra
Principal Investigator
(Strike out inapplicable terms)

IIT Gandhinagar
Indian Institute of Technology Gandhinagar

Final Report

Project Title

Bias-corrected climate projections from Coupled Model Intercomparison Project-5

Submitted by

Vimal Mishra
Associate Professor,
Department of Civil Engineering,
Indian Institute of Technology Gandhinagar

March 2021

Table of Content

List of Table

List of Figure

Abstract

Climate change is expected to pose enormous challenges on agriculture, water supplies, housing, and the livelihoods of millions of people in South Asia. Here, we developed the daily bias-corrected maximum and minimum temperatures and precipitation using output from five GCMs that participated in the Coupled Model Intercomparison Project-5 (CMIP5) at 0.25° spatial resolution over India. For the historical and three future scenarios (RCP 2.6, RCP 4.5, RCP 8.5), five CMIP5-GCMs were selected based on the availability of daily precipitation, maximum and minimum temperatures. The daily precipitation, maximum and minimum temperatures datasets were bias-corrected using Empirical Quantile Mapping (EQM) for the historical (1951–2005) and future (2006–2100) climate. We compared the bias-corrected precipitation, maximum, and minimum temperatures, as well as the seasonal period, to observations and find that it fits well. After bias correction, the bias in all three variables for the historical data was significantly decreased (1951-2005). The efficacy of EQM-based bias correction method is shown by the decrease in bias in precipitation, maximum and minimum temperatures.

Chapter I
Introduction

Introduction

India is one of the most densely populated regions of the world. A majority of the population in India depends on agriculture for their livelihood. India is among the global hot spots that are likely to face the detrimental impacts of climate change (Diffenbaugh & Giorgi, 2012; Suarez-Gutierrez et al., 2020). Considerable changes in precipitation and temperature are projected in India that will have implications for water resources and agriculture (Immerzeel et al., 2010; Knox et al., 2012; Lobell & Burke, 2008; Turner & Annamalai, 2012). The risk of floods and droughts are likely to increase in India under the warming climate (Aadhar & Mishra, 2019, 2020; Ali et al., 2019; Arnell & Gosling, 2016; Hirabayashi et al., 2013; Winsemius et al., 2016). Both recent droughts and floods have affected a large population and caused enormous damage to agriculture and infrastructure in India (Hunt & Menon, 2020; Mishra & Shah, 2018; Webster et al., 2011). Similarly, the frequency and intensity of severe heatwaves have increased in India and projected to increase in the future (Im et al., 2017; JA et al., 2020; Mazdiyasni et al., 2017; Mishra et al., 2017; Mukherjee & Mishra, 2018). Overall, the frequency of both precipitation and temperature extremes has considerably increased in the past decades and likely to rise further under the warming climate (JA et al., 2020; Mukherjee et al., 2018).

Projections from the General Circulation Models (GCMs) play a vital role in understanding the future changes in climate. However, spatial resolution at which GCMs are run is often too coarse to get reliable projections at the regional and local scale (Christensen et al., 2008). Precipitation

and temperature projections at higher spatial resolution are required for the climate impact assessments (Barbero et al., 2017; Cayan et al., 2010; Maurer et al., 2010). Moreover, precipitation and temperature from the GCMs have a bias due to their coarse resolution or model parameterizations (Ashfaq et al., 2017; Mishra et al., 2014). Therefore, for the assessment of the climate change and its impacts on different sectors (e.g., water resources, agriculture), bias-correction is required (Christensen et al., 2008; Eisner et al., 2012; Maraun et al., 2017; Piani et al., 2010; Pierce et al., 2015; Thrasher et al., 2012; A. W. Wood et al., 2004). Both statistical and dynamical approaches are used for downscaling and bias correction of climate change projections from GCMs. Statistical approaches are based on the distribution and relationship between the observed and projected data for the historical period (Pierce et al., 2015; Thrasher et al., 2012). On the other hand, dynamical downscaling approaches are based on regional climate model forced with the boundary conditions from the coarse resolution GCMs (Giorgi & Gutowski, 2015; Mearns et al., 2013). Both statistical and dynamical downscaling approaches have limitations (Abatzoglou & Brown, 2012; Maurer & Hidalgo, 2008). The primary limitation of the dynamical downscaling is related to the requirement of computational efforts to run the regional climate models at higher spatial and temporal resolution (Mishra et al., 2014; White & Toumi, 2013). Moreover, dynamical downscaling may not remove the bias in climate variables, which might require corrections based on the statistical approaches (White & Toumi, 2013). Given these limitations, statistical bias correction approaches are widely used in climate change impact assessments (Gutmann et al., 2014; Xu & Wang, 2019).

Considering the climate change impacts in India, we develop a bias-corrected dataset for daily precipitation, maximum and minimum temperatures using output from five GCMs that participated in the Coupled Model Intercomparison Project-5 (CMIP5). The five GCMs were

selected based on the availability of daily precipitation, maximum and minimum temperatures for the historical and three scenarios (RCP 2.6, RCP 4.5, RCP 8.5). We used empirical quantile mapping (EQM) to develop bias-corrected data at daily temporal and 0.25° spatial resolution for India. The bias-corrected projections from 13 CMIP-GCMs can be used for estimating the projected changes in mean and extreme climate in India.

Chapter II
Data and Methods

Methods

We obtained observed daily gridded precipitation, minimum and maximum temperatures for India for the 1951-2005 period. Daily precipitation at 0.25° was obtained from the India Meteorological Department (IMD) (Pai et al., 2014). Pai et al. (2014) developed gridded daily precipitation for India using station observations from more than 6000 stations located across India. The precipitation captures critical features of the Indian summer monsoon, including higher rainfall in the Western Ghats and northeastern India and lower rainfall in the semi-arid and arid regions of western India. Besides, gridded precipitation captures the orographic rain in the Western Ghats and foothills of Himalaya. The gridded precipitation data from IMD has been used for various hydroclimatic applications (Mishra et al., 2016; D. Shah & Mishra, 2019). Gridded daily maximum and minimum temperatures from IMD were developed using station-based observations from more than 350 stations located across India (Srivastava et al., 2009). There is bias in temperature observations from IMD in the Himalayan region, which can be attributed to sparse station density (R. Shah & Mishra, 2014). We used gridded observations for bias correction as station data are not available.

We obtained daily precipitation, maximum and minimum temperatures from five CMIP5-GCMs from https://esgf-node.llnl.gov/search/cmip5/. Precipitation, maximum and minimum temperatures from CMIP6-GCMs are available at different spatial resolutions (Table 1). For instance, the spatial resolution of the CMIP5 projection varies from 1.4° (CNRM) to more than

2.8° (CanESM5). All the three variables were selected for the historical (1850-2005) and three representative concentration pathways (RCPs) scenarios under r1i1p1f1 initial condition at daily time scale (Taylor et al., 2012). The three RCPs assume an increase of 2.6 W/m^2 (low forcing sustainability pathway; RCP-2.6), 4.5 W/m^2 (medium forcing middle of the road pathway; RCP-4.5), and 8.5 W/m^2 (high end forcing pathway; RCP-8.5) radiative forcing by the end of the 21st century. Further details on the CMIP5 can be obtained from Taylor et al. (2012). We regridded all the variables from CMIP5 to 1° spatial resolution to make them consistent. However, the effect of regridding using bilinear interpolation was checked by comparing the gridded datasets against the raw data for all-India mean of precipitation, maximum and minimum temperatures. We did not find any considerable differences in the all-India averaged precipitation and temperature using regridded and raw output from the GCMs.

Outputs of the various atmospheric (e.g., maximum and minimum temperatures, and precipitation) variables obtained from GCMs are known to exhibit systematic biases. Hence, these outputs need to be bias-corrected to produce reliable estimates at regional and local scales for climate impact assessment. To achieve this, statistical transformations that attempt to find a function that maps the model output to a new distribution such that the resulting distribution matches that of observations. In general, this transformation can be formulated as (Piani et al., 2010):

$$x_m^o = f(x_m) \qquad (1)$$

Where x_m^o is the bias-corrected model output. If the statistical distribution of x_m and x_0 are known, the transformation can be written as:

$$x_m^o = F_0^{-1}(F_m(x_m)) \quad (2)$$

Where F_m and F_o are the Cumulative Distribution Functions (CDFs) of x_m and x_o respectively. In Empirical Quantile Mapping [EQM, (Aadhar & Mishra, 2017; Andrew W. Wood et al., 2002)), instead of assuming parametric distributions, empirical CDFs (Jakob Themeßl et al., 2011; Julien et al., 2007; Thrasher et al., 2012) are estimated from the percentiles calculated from x_m and x_0. As a result, EQM and its variants can be applied to both temperature and precipitation even if their underlying distributions are different and hence recommended for statistical bias correction (Cannon, 2011).

In the context of statistical downscaling, since the observations are at a higher resolution than models, EQM on bilinearly interpolated model outputs at observation resolution is often used to address the scale mismatch and generate post-processed model outputs (Aadhar & Mishra, 2017). We choose non-parametric transformation approaches over the parametric approaches as has shown better skills in the comparison to parametric methods in reducing biases from GCM as well as Regional Climate Model (RCM) outputs (Gudmundsson et al., 2012).

We used EQM to statistically downscale the daily maximum and minimum temperatures, and precipitation. We use the outputs (x_m) from five CMIP5-GCMs (Table 1), which are available at different resolutions (Table 1). Observations for the three variables at the resolution of 0.25-degree are obtained from the IMD. We used the 1951-2005 period to obtain the transformation function to map the distribution of x_m to x_o. For precipitation, the drizzle effect is corrected by using a wet day threshold of 1 mm/day (Gudmundsson et al., 2012; Piani et al., 2010). If the values from model projections are larger (smaller) than the training values used to estimate the empirical CDF, the correction found for the highest (lowest) quantile of the training period is used. We used mapped

transformation to bias correct the outputs for the historical period and the RCP 2.6, RCP 4.5, and RCP 8.5 scenarios for the 2006-2100 period for all the three variables. Quantile mapping based statistical bias correction has been widely used, and its performance was found to be satisfactory in comparison to the other methods (Bürger et al., 2012; Maurer et al., 2010; Thrasher et al., 2012).

Table 1. List of the General Circulation Models (GCMs) used in the study. RCP 2.6, RCP 4.5 and RCP 8.5 are used for all the GCMs.

Model	**Modeling Group**	**Native resolution**	
		Longitude	**Latitude**
BNU-ESM	College of Global Change and Earth System Science, Beijing Normal University (GCESS)	2.8	2.8
CanESM2	Canadian Centre for Climate Modelling and Analysis (CCCMA)	2.8	2.8
CNRM	Centre National de Recherches Météorologiques Centre Européen de Recherche et Formation Avancée en Calcul Scientifique	1.4	1.4
MPI-ESM-LR	Max-Planck-Institut für Meteorologie (MPI-M)	1.875	1.875
MPI-ESM-MR		1.875	1.875

Chapter III
Results

Results

The bias in mean annual precipitation, maximum and minimum temperatures was corrected against the observations from IMD. The CMIP5-GCMs show a dry bias (15-20%) in precipitation in the majority of India (Aadhar & Mishra, 2019). A high cold and warm bias in both maximum and minimum temperatures were found in the CMIP5-GCMs (Aadhar & Mishra, 2019). We applied the EQM approach to correct the bias in the CMIP5-GCM output at daily timescale. The bias in precipitation and temperatures extremes and mean in the raw output was corrected separately. The 90^{th} percentiles of precipitation of rainy days (precipitation more than 1mm), maximum and minimum temperatures and monthly mean data were corrected for the historical period (1951-2006) from CMIP5-GCMs against the observed dataset. We find that the EQM based bias correction has successfully removed the bias in precipitation, maximum and minimum temperatures across India (Fig. 1). We compared the season cycle of bias-corrected precipitation, maximum and minimum temperatures from the CMIP5-GCMs against the observed dataset for the 1951-2005 period. We find that the seasonal cycle of the bias-corrected precipitation, maximum, and minimum temperatures compare well against the observations (Fig. 1). The bias was substantially reduced after the bias correction in all the three variables for the historical (1951-2005) period (Fig. 1a-c). The reduction in bias in precipitation, maximum and minimum temperatures shows the effectiveness of our bias correction approach based on EQM.

Overall, our results show that the EQM approach successfully corrects the bias in the CMIP5-GCMs, which can be used for climate impacts studies in India. Also, the bias-corrected dataset can be used for hydrological studies in the Indian sub-continental river basins.

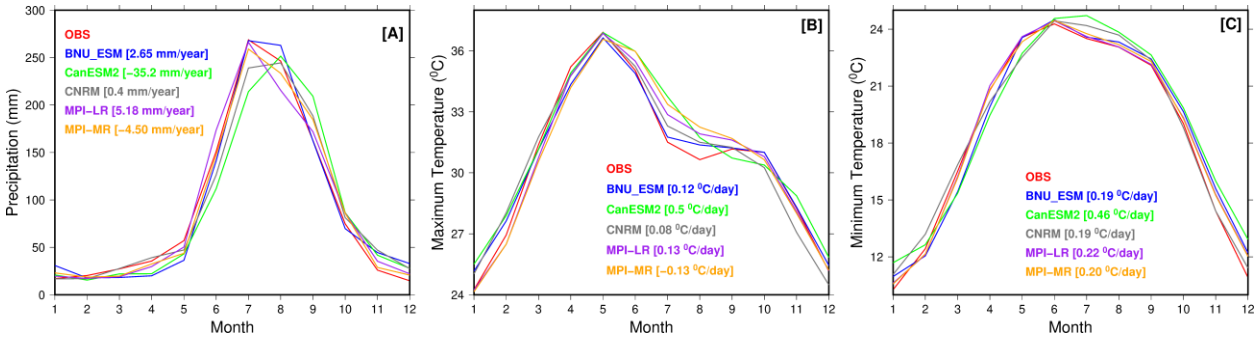


Figure 1. Seasonal cycle of bias-corrected precipitation, maximum and minimum temperatures. Comparison of the each CMIP5-GCM mean seasonal cycle of bias-corrected (a) precipitation, (b) maximum temperature, and (c) minimum temperature against the observations for the 1951-2005 period (red).

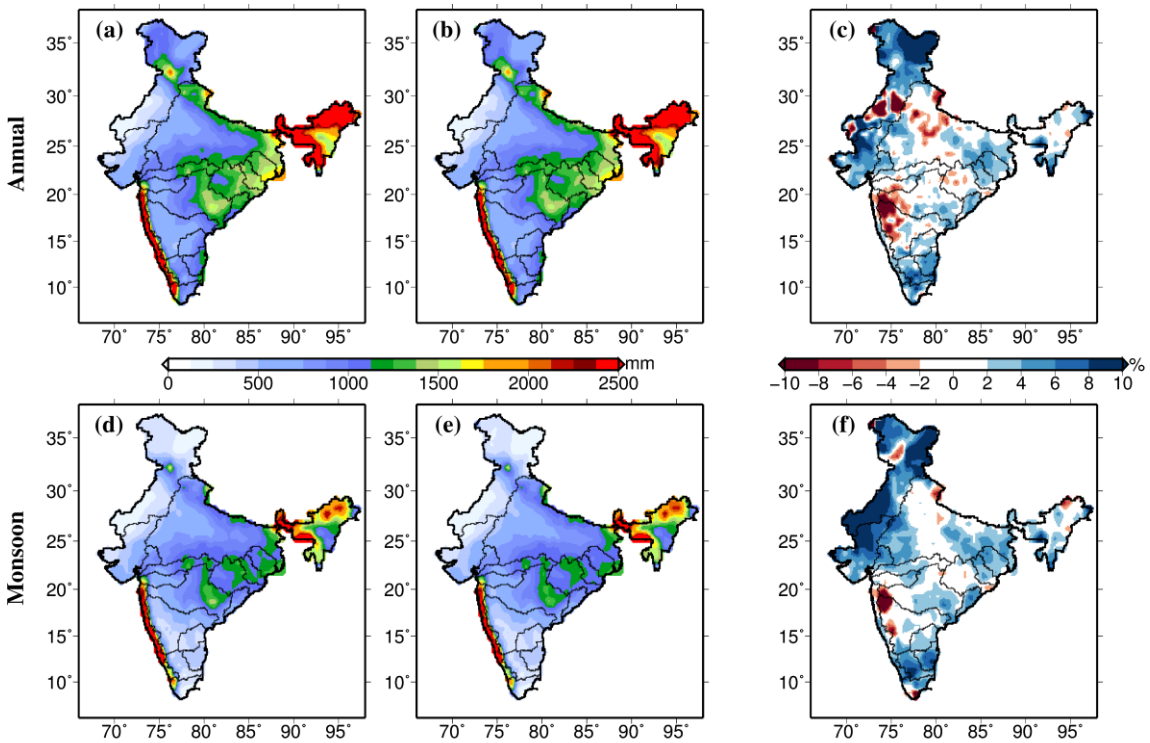


Figure 2. (a, b) Annual mean (1970-2000) precipitation of observed and bias-corrected five ensembles mean GCMs. (d, e) Monsoon mean (1970-1999) precipitation of observed and bias-corrected five ensembles mean GCMs. (c,f) Bias in precipitation.

We also bias corrected GCMs precipitation and air temperature data against IMD observation from 1951 to 2100 for RCP2.6 and RCP8.5 (Figure 2 and Figure 3). Annual and monsoon precipitation bias were less than 10%. Bias corrected precipitation successfully captures seasonality and distribution. However, we find more than 10% bias in precipitation over Rajasthan region, which could be attributed due to arid region. Air temperature bias was less than 0.2 ºC (Figure 3).

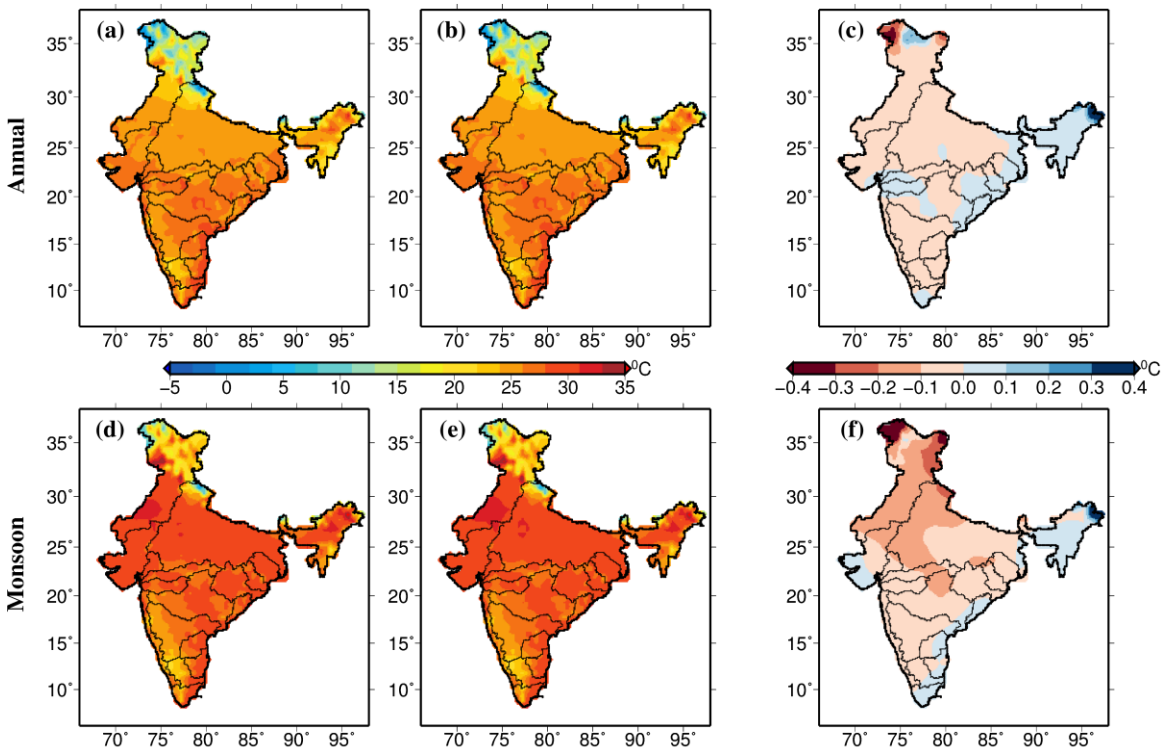


Figure 3. (a, b) Annual mean (1970-2000) air temperature of observed and bias-corrected five ensembles mean GCMs. (d, e) Monsoon mean (1970-1999) precipitation of observed and bias-corrected five ensembles mean GCMs. (c,f) Bias in precipitation. (this figure is same as Figure 2 but for air temperature).

The bias corrected precipitation and minimum and maximum temperature datasets of CMIP5 are submitted to IIT Bombay.

Chapter IV
Conclusion

Conclusion

The risk of floods and droughts are likely to increase in India under the warming climate (Aadhar & Mishra, 2019, 2020; Arnell & Gosling, 2016). Bias in data leads to inaccurate metrological variable forecasts, which results the poor hydroloical forcating. To study the extremes, bias corrected data is needed. Dynamical downscaling may not remove the bias in climate variables, which might require corrections based on the statistical approaches (White & Toumi, 2013). Statistical bias correction techniques are commonly used in climate change impact assessments (Gutmann et al., 2014; Xu & Wang, 2019). As a result, we used statistical bias correction techniques in this analysis. We showed the bias-correction of precipitation, maximum and minimum temperatures from five CMIP5-GCMs using a statistical (EQM approach) method in India. Moreover, the outputs were validated by comparing bias-corrected datasets to observed datasets from the IMD, which showed good results.

Based on our analysis, the following conclusion can be listed below:

- ✓ The EQM method works well for bias correction.
- ✓ Over the majority of India, the CMIP5-GCMs show a dry bias (15-20%) in precipitation.
- ✓ Bias corrected precipitation successfully captures seasonality and distribution.
- ✓ The bias-corrected precipitation, maximum, and minimum temperatures have a good seasonal pattern that matches observed results.
- ✓ The bias-corrected precipitation and temperature daatsets can be used to study for projected changes in mean and extreme climate in India.

Reference

Reference

Aadhar, S., & Mishra, V. (2017). High-resolution near real-time drought monitoring in South Asia. *Scientific Data*, *4*, 170145. Retrieved from http://dx.doi.org/10.1038/sdata.2017.145

Aadhar, S., & Mishra, V. (2019). A Substantial rise in the area and population affected by dryness in South Asia under 1.5, 2.0 and 2.5 C warmer worlds. *Environmental Research Letters*, *14*, 114021. https://doi.org/10.1088/1748-9326/ab4862

Aadhar, S., & Mishra, V. (2020). Increased drought risk in South Asia under warming climate: Implications of uncertainty in potential evapotranspiration estimates. *Journal of Hydrometeorology*. https://doi.org/10.1175/jhm-d-19-0224.1

Abatzoglou, J. T., & Brown, T. J. (2012). A comparison of statistical downscaling methods suited for wildfire applications. *International Journal of Climatology*, *32*(5), 772–780. https://doi.org/10.1002/joc.2312

Ali, H., Modi, P., & Mishra, V. (2019). Increased flood risk in Indian sub-continent under the warming climate. *Weather and Climate Extremes*, *25*, 100212. https://doi.org/10.1016/j.wace.2019.100212

Arnell, N. W., & Gosling, S. N. (2016). The impacts of climate change on river flood risk at the global scale. *Climatic Change*, *134*(3), 387–401. https://doi.org/10.1007/s10584-014-1084-5

Ashfaq, M., Rastogi, D., Mei, R., Touma, D., & Ruby Leung, L. (2017). Sources of errors in the simulation of south Asian summer monsoon in the CMIP5 GCMs. *Climate Dynamics*, *49*(1–2), 193–223. https://doi.org/10.1007/s00382-016-3337-7

Barbero, R., Fowler, H. J., Lenderink, G., & Blenkinsop, S. (2017). Is the intensification of precipitation extremes with global warming better detected at hourly than daily resolutions? *Geophysical Research Letters*, *44*(2), 974–983. https://doi.org/10.1002/2016GL071917

Bürger, G., Murdock, T. Q., Werner, A. T., Sobie, S. R., & Cannon, A. J. (2012). Downscaling extremes-

an intercomparison of multiple statistical methods for present climate. *Journal of Climate*, *25*(12), 4366–4388. https://doi.org/10.1175/JCLI-D-11-00408.1

Cannon, A. J. (2011). Quantile regression neural networks: Implementation in R and application to precipitation downscaling. *Computers and Geosciences*, *37*(9), 1277–1284. https://doi.org/10.1016/j.cageo.2010.07.005

Cayan, D. R., Das, T., Pierce, D. W., Barnett, T. P., Tyree, M., & Gershunova, A. (2010). Future dryness in the Southwest US and the hydrology of the early 21st century drought. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(50), 21271–21276. https://doi.org/10.1073/pnas.0912391107

Christensen, J. H., Boberg, F., Christensen, O. B., & Lucas-Picher, P. (2008). On the need for bias correction of regional climate change projections of temperature and precipitation. *Geophysical Research Letters*, *35*(20). https://doi.org/10.1029/2008GL035694

Diffenbaugh, N. S., & Giorgi, F. (2012). Climate change hotspots in the CMIP5 global climate model ensemble. *Climatic Change*, *114*(3–4), 813–822. https://doi.org/10.1007/s10584-012-0570-x

Eisner, S., Voss, F., & Kynast, E. (2012). Statistical bias correction of global climate projections - Consequences for large scale modeling of flood flows. *Advances in Geosciences*, *31*, 75–82. https://doi.org/10.5194/adgeo-31-75-2012

Giorgi, F., & Gutowski, W. J. (2015). Regional Dynamical Downscaling and the CORDEX Initiative. *Annual Review of Environment and Resources*, *40*(1), 467–490. https://doi.org/10.1146/annurev-environ-102014-021217

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., & Engen-Skaugen, T. (2012). Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations – A comparison of methods. *Hydrology and Earth System Sciences*, *16*(9), 3383–3390. https://doi.org/10.5194/hess-16-3383-2012

Gutmann, E., Pruitt, T., Clark, M. P., Brekke, L., Arnold, J. R., Raff, D. A., & Rasmussen, R. M. (2014). An intercomparison of statistical downscaling methods used for water resource assessments in the

United States. *Water Resources Research Research*, *50*, 7167–7186. https://doi.org/10.1002/2014WR015559.Received

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., et al. (2013). Global flood risk under climate change. *Nature Climate Change*, *3*(9), 816–821. https://doi.org/10.1038/nclimate1911

Hunt, K. M. R., & Menon, A. (2020). The 2018 Kerala floods: a climate change perspective. *Climate Dynamics*, *54*(3–4), 2433–2446. https://doi.org/10.1007/s00382-020-05123-7

Im, E. S., Pal, J. S., & Eltahir, E. A. B. (2017). Deadly heat waves projected in the densely populated agricultural regions of South Asia. *Science Advances*. https://doi.org/10.1126/sciadv.1603322

Immerzeel, W. W., van Beek, L. P. H., & Bierkens, M. F. P. (2010). Climate change will affect the Asian water towers. *Science (New York, N.Y.)*, *328*(5984), 1382–5. https://doi.org/10.1126/science.1183188

JA, N., van der Wiel, K., Bhatia, U., Stone, D., Selten, F. M., & Mishra, V. (2020). A seven-fold rise in the probability of exceeding the observed hottest summer in India in a 2°C warmer world. *Environmental Research Letters*. https://doi.org/10.1088/1748-9326/ab7555

Jakob Themeßl, M., Gobiet, A., & Leuprecht, A. (2011). Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. *International Journal of Climatology*, *31*(10), 1530–1544. https://doi.org/10.1002/joc.2168

Julien, B., L., T., F., H., & E., M. (2007). Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies. *International Journal of Climatology*, *27*(August), 1643–1655. https://doi.org/10.1002/joc

Knox, J., Hess, T., Daccache, A., & Wheeler, T. (2012). Climate change impacts on crop productivity in Africa and South Asia. *Environmental Research Letters*, *7*(3). https://doi.org/10.1088/1748-9326/7/3/034032

Lobell, D. B., & Burke, M. B. (2008). Why are agricultural impacts of climate change so uncertain? the importance of temperature relative to precipitation. *Environmental Research Letters*, *3*(3). https://doi.org/10.1088/1748-9326/3/3/034007

Maraun, D., Shepherd, T. G., Widmann, M., Zappa, G., Walton, D., Gutiérrez, J. M., et al. (2017). Towards process-informed bias correction of climate change simulations. *Nature Climate Change*, *7*(11), 764–773. https://doi.org/10.1038/nclimate3418

Maurer, E. P., & Hidalgo, H. G. (2008). Utility of daily vs. monthly large-scale climate data: an intercomparison of two statistical downscaling methods. *Hydrology and Earth System Sciences*, *14*(6), 1125–1138. https://doi.org/10.5194/hess-14-1125-2010

Maurer, E. P., Hidalgo, H. G., Das, T., Dettinger, M. D., & Cayan, D. R. (2010). The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. *Hydrology and Earth System Sciences*, *14*(6), 1125–1138. https://doi.org/10.5194/hess-14-1125-2010

Mazdiyasni, O., AghaKouchak, A., Davis, S. J., Madadgar, S., Mehran, A., Ragno, E., et al. (2017). Increasing probability of mortality during Indian heat waves. *Science Advances*, *3*(6), 1–6. https://doi.org/10.1126/sciadv.1700066

Mearns, L. O., Sain, S., Leung, L. R., Bukovsky, M. S., McGinnis, S., Biner, S., et al. (2013). Climate change projections of the North American Regional Climate Change Assessment Program (NARCCAP). *Climatic Change*, *120*(4), 965–975. https://doi.org/10.1007/s10584-013-0831-3

Mishra, V., & Shah, H. L. (2018). Hydroclimatological Perspective of the Kerala Flood of 2018. *JOURNAL GEOLOGICAL SOCIETY OF INDIA*, *92*, 645–650. https://doi.org/10.1007/s12594-018-1079-3

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

Mishra, V., Aadhar, S., Asoka, A., Pai, S., & Kumar, R. (2016). On the frequency of the 2015 monsoon season drought in the Indo-Gangetic Plain. *Geophysical Research Letters*, *43*(23), 12,102-12,112. https://doi.org/10.1002/2016GL071407

Mishra, V., Mukherjee, S., Kumar, R., & Stone, D. A. (2017). Heat wave exposure in India in current, 1.5 °C, and 2.0 °C worlds. *Environmental Research Letters*, *12*(12), 124012.

https://doi.org/10.1088/1748-9326/aa9388

Mukherjee, S., & Mishra, V. (2018). A sixfold rise in concurrent day and night-time heatwaves in India under 2 °C warming. *Scientific Reports*, *8*(1), 16922. https://doi.org/10.1038/s41598-018-35348-w

Mukherjee, S., Aadhar, S., Stone, D., & Mishra, V. (2018). Increase in extreme precipitation events under anthropogenic warming in India. *Weather and Climate Extremes*. https://doi.org/10.1016/j.wace.2018.03.005

Pai, D. S., Sridhar, L., Rajeevan, M., Sreejith, O. P., Satbhai, N. S., & Mukhopadyay, B. (2014). Development of a new high spatial resolution (0 . 25 ° × 0 . 25 °) Long Period (1901-2010) daily gridded rainfall data set over India and its comparison with existing data sets over the region. *Mausam*, *65*(1), 1–18.

Piani, C., Haerter, J. O., & Coppola, E. (2010). Statistical bias correction for daily precipitation in regional climate models over Europe. *Theoretical and Applied Climatology*, *99*(1–2), 187–192. https://doi.org/10.1007/s00704-009-0134-9

Pierce, D. W., Cayan, D. R., Maurer, E. P., Abatzoglou, J. T., & Hegewisch, K. C. (2015). Improved bias correction techniques for hydrological simulations of climate change. *Journal of Hydrometeorology*, *16*(6), 2421–2442. https://doi.org/10.1175/JHM-D-14-0236.1

Shah, D., & Mishra, V. (2019). Integrated Drought Index (IDI) for drought monitoring and assessment in India. *Water Resources Research*. https://doi.org/10.1029/2019wr026284

Shah, R., & Mishra, V. (2014). Evaluation of the Reanalysis Products for the Monsoon Season Droughts in India. *Journal of Hydrometeorology*, *15*(4), 1575–1591. https://doi.org/10.1175/JHM-D-13-0103.1

Srivastava, A. K., Rajeevan, M., & Kshirsagar, S. R. (2009). Development of a high resolution daily gridded temperature data set (1969 – 2005) for the Indian region. *Atmospheric Science Letters*, *10*(October), 249–254. https://doi.org/10.1002/asl

Suarez-Gutierrez, L., Müller, W. A., Li, C., & Marotzke, J. (2020). Dynamical and thermodynamical drivers of variability in European summer heat extremes. *Climate Dynamics*, *54*(9), 4351–4366. https://doi.org/10.1007/s00382-020-05233-2

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*. https://doi.org/10.1175/BAMS-D-11-00094.1

Thrasher, B., Maurer, E. P., McKellar, C., & Duffy, P. B. (2012). Technical Note: Bias correcting climate model simulated daily temperature extremes with quantile mapping. *Hydrology and Earth System Sciences*, *16*(9), 3309–3314. https://doi.org/10.5194/hess-16-3309-2012

Turner, A. G., & Annamalai, H. (2012). Climate change and the South Asian summer monsoon. *Nature Climate Change*. https://doi.org/10.1038/nclimate1495

Webster, P. J., Toma, V. E., & Kim, H. M. (2011). Were the 2010 Pakistan floods predictable? *Geophysical Research Letters*, *38*(4), 1–5. https://doi.org/10.1029/2010GL046346

White, R. H., & Toumi, R. (2013). The limitations of bias correcting regional climate model inputs. *Geophysical Research Letters*, *40*(12), 2907–2912. https://doi.org/10.1002/grl.50612

Winsemius, H. C., Aerts, J. C. J. H., Van Beek, L. P. H., Bierkens, M. F. P., Bouwman, A., Jongman, B., et al. (2016). Global drivers of future river flood risk. *Nature Climate Change*. https://doi.org/10.1038/nclimate2893

Wood, A. W., Leung, L. R., Sridhar, V., & Lettenmaier, D. P. (2004). Hydrologic Implications of Dynamical and Statistical Approaches to Downscaling Climate Model Outputs. *Climatic Change*, *62*(1–3), 189–216. https://doi.org/10.1023/B:CLIM.0000013685.99609.9e

Wood, Andrew W., Maurer, E. P., Kumar, A., & Lettenmaier, D. P. (2002). Long-range experimental hydrologic forecasting for the eastern United States. *Journal of Geophysical Research D: Atmospheres*, *107*(20), 1–15. https://doi.org/10.1029/2001JD000659

Xu, L., & Wang, A. (2019). Application of the Bias Correction and Spatial Downscaling Algorithm on the Temperature Extremes From CMIP5 Multimodel Ensembles in China. *Earth and Space Science*, *6*(12), 2508–2524. https://doi.org/10.1029/2019EA000995