Development of Hydrometeorological Drought Index at River Basin Scale under Climate Change - A Case study of Krishna River Basin

Thesis submitted in partial fulfillment
of the requirements for the degree of

Master of Science
in
Computer Science and Engineering
by Research

by

Galla Sireesha Naidu
201302207
sireeshanaidu.galla@research.iiit.ac.in

International Institute of Information Technology
Hyderabad - 500 032, INDIA

June, 2020

International Institute of Information Technology
Hyderabad, India

CERTIFICATE

It is certified that the work contained in this thesis, titled "Development of Hydrometeorological Drought Index at River Basin Scale under Climate Change - A Case study of Krishna River Basin" by Galla Sireesha Naidu, has been carried out under my supervision and is not submitted elsewhere for a degree.

Date	Adviser: Dr. Shaik Rehana

To My Lt. Grand Father L.Chinnaswamy Naidu

Acknowledgments

I would like to take this opportunity to thank my research advisor Dr. Shaik Rehana, without whose vision this work would not have taken a form. I am grateful to Dr. Rehana for never doubting my ability to do things in different research areas and helping me in pursuing the right problem statements in areas, I was passionate about. I thank her for being constantly supportive in the most rough times over the past few years and motivating me to work hard through paper submissions and to meet conference deadlines. I would also like to thank Dr. K.S. Rajan for introducing and arousing interest in the domain of Spatial Informatics.

I would like to thank my friend and colleague N. Tinku Monish, for being a constant companion in this journey and helping me in solving the challenges in this work, and U.Sowjanya, for providing valuable inputs and the domain knowledge. My appreciation goes to the institute, IIIT-Hyderabad for providing an exposure to different research domains and the Lab for Spatial Informatics for focussing on real world problems like global environmental change, drought analysis and remote sensing impacting the society.

I specially thank my friends VenuMadhav Kattagoni and Manasa Lingamallu for patiently bearing me and getting me out of the darkest times in my life. I also thank them for being curious about my work, providing insights. I also thank my friends Divya Jyothi Gaddipati, Amitha Reddy, Lakshmi Gangumalla, Rohit SVK, Shubham Rathi, Akhil Batra and Sampreeth Reddy for their constant support in my life. I also thank Mohana Priya for her help in language correction.

Above everyone else, I thank my father Mr. Vijay Naidu Galla for inspiring and guiding me, my mother Mrs. Aruna Galla for believing in me and never giving up on me, and my brother Anil Krishna for caring and providing me with all the love and comfort. I thank my family members for the faith in me and celebrating my every achievement as their own. I also thank my cousins Pavan, Rahul and Gokul for their emotional support. Lastly, I thank my friend Vinisha, and my parents and brother for showing the importance of working hard to rise in life and giving to the society.

The research work presented in this thesis is funded by Science and Engineering Research Board (SERB), Department of Science and Technology, Government of India through Start-up Grant for Young Scientists (YSS) Project no. YSS/2015/002111 received by Dr. Shaik Rehana.

Abstract

Global warming, changes in precipitation levels have led us to keep an eye on water resources and to adopt new management strategies for their sustainability. Prediction of projected climatological variables accounting for greenhouse gases in the atmosphere, accurate projections of hydroclimatological variables under climate change is crucial for making adaptive measures and mitigation policies. An analysis of the changes in energy and water balance of a hydrologic system is required to establish these strategies and check their efficiency. Also an analysis of extreme variations in climate will give us a better picture as how a hydrologic system is interlinked with these events. Droughts are one such extreme events which can cause huge damage to ecosystems and society, if sufficient steps are not taken to assess and monitor them.

Various meteorological drought indices have evolved to characterize the drought at larger scales. Meteorological drought indicators with precipitation and Potential Evapotranspiration (PET) can incorporate atmospheric water demand based on energy available in drought characterization, which are ideal for energy-limited regions, however, for water-limited regions where Actual Evapotranspiration (AET) is the major hydrological variable, therefore drought is defined by water availability rather than energy. Furthermore, at river basin scales meteorological drought assessments may not be enough to understand the water availability to the crops and other environmental aspects of the river systems. The present thesis aimed to develop a hydrometeorological drought prediction index by considering precipitation, evapotranspiration and runoff. The study mainly emphasized to understand the effect of various meteorological and hydrological variables on drought and how the inclusion of AET can outperform a drought index based solely on PET and P at a river basin scale. In this context, the study proposed hydrologically calibrated Actual Evapotranspiration (AET) considering precipitation, potential evapotranspiration and runoff. The hydrologically calibrated AET was further used in the formulation of Standardized Precipitation Evapotranspiration Index (SPEI) to develop hydrometeorological drought index, Standardized Precipitation Actual Evapotranspiration Index (SPAEI). The proposed drought index SPAEI imposes the effect of precipitation, Potential Evapotranspiration, Actual Evapotranspiration using operational meteorological data sets of precipitation and temperatures and surface runoff.

The performance of PET and AET based drought indices was compared using historical droughts in terms of severity, areal extent, frequency and duration. The proposed AET based drought indices have effectively captured the historical drought years over the Krishna river basin. Inclusion of AET in the drought characterization along with precipitation and PET can drive the highly intensified drought events

determined by SPEI into moderate and less frequent droughts with short durations. As SPAEI is more reasonable in reflecting surface water-energy balance it enables better characterization of meteorological and hydrological droughts at regional scales.

Further, the study assessed the climate signals on drought assessment with Regional Circulation Model (RCM) outputs and Global Circulation Model(GCM) outputs. Meteorological climate projections for the future were obtained using statistical downscaling which involves establishing a relationship between predictands(local climate variables) and predictors(global coarse scale projections). This relationship is established using a set of statistical and machine learning techniques such as Quantile Bias Correction, K-means Clustering, Classification and Regression Trees(CART) and Support Vector Regression. The study observed that compared to RCM climate projections, GCM based statistical downscaling models have performed well in capturing the historical observations, indicating reliable predictions based on GCM based observations at river basin scales.

Although, statistical downscaling models has proved to be reliable projections for climate change impact assessment studies, the RCM data sets provides ease of quick implementation with less computational effort. Therefore, the study suggested the use of RCM projections for preliminary understanding and GCM based projections for a detailed climate change impact assessment study. The projected drought characteristics based on GCM analysis has observed that while the intensity and duration of the drought has been found to be slightly increasing, the frequency and the drought areal extent have been increasing alarmingly over the Krishna river basin. With the GCMs, it has been estimated that there is a net increase of 25%-31% in the drought areal extent, an increase of 50%-90% in drought frequencies in Krishna river basin from the current to future periods. The climate change impact assessment on drought characteristics based on RCM and GCM outputs can provide insights for the river basin water management and decision-making policies.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Climate change under anthropogenic global warming has aggravated the hydrological cycle and spurred to changes in hydrological variables such as precipitation, evapotranspiration, runoff etc. Prediction of the projected climatological variables accounting for green house gases in the atmosphere [33], accurate projections of hydro-climatological variables under climate change is crucial for making adaptive measures and mitigation policies [68, 67]. An analysis of changes in the water balance of a hydrologic system is required to establish these strategies and check their efficiency. Also, an analysis of extreme variations in climate will give us a better picture as how a hydrologic system is interlinked with these events. Among these extreme events, droughts are the most widespread and slowly developing atmospheric hazards which remain for a long duration affecting natural resources, environment, and millions of people globally every year [1, 18]. Furthermore, droughts corresponds to the failure of spatial and temporal precipitation (meteorological drought), decrease in available water and inadequate streamflows (hydrological drought), decrease in soil moisture and crop yields (agricultural drought), therefore consequent impact on ecosystem and socioeconomic activities of the human being (socio-economic drought) [104]. However, the river basin management and decision making can be based on the first two categories, namely meteorological and hydrological droughts [110].

1.2 Literature Review

1.2.1 Drought Analysis

The global land surface in extreme drought is predicted to increase from 1-3% for the present day to 30% by the 2090s [83]. More intense droughts and increased precipitation variability lead to increased stresses to water, agriculture and economic activities [64]. Due to the irregular summer monsoon and increase in air temperature, the frequency of extreme and widespread multi-year droughts has increased in India in recent decades, causing enormous damage to crops and society [61, 81]. India has experienced

23 large-scale droughts starting from 1891 to 2009 and the frequency of droughts is increasing [47]. The severity of droughts has been reported as increasing in many parts of the Indian sub-continental basins under climate change [61]. Given this, drought management studies at river basin scale including the variability of precipitation, evapotranspiration and runoff have gained much attention in the literature over Indian river basins [61].

The major factors for the persistence of droughts at river basin scale are changes in water balance due to the alterations in the water supply (precipitation), energy (potential evaporation) and land surface characteristics (vegetation and topography). A region experiencing meteorological drought may lead to lagged impact on the water availability shortages (hydrological drought) [97]. Here, a drought assessment solely based on meteorological aspects (precipitation, P) without considering water balance variable (Runoff, R) of hydrological cycle will not be sufficient for the regional water resources management decisions [49]. However, hydrological drought assessment entirely based on estimation of below normal streamflow may mislead due to the human influenced regulated flows, diversions, water transfers and instream abstractions [96]. An effective composite drought monitoring tool incorporating multivariate perspectives of hydrometeorology provides new directions to depict shortages of precipitation and streamflows combinedly [41].

In this context, several drought indices have been developed, which evaluate the deviation of climate variables in a given year from the normal conditions [18, 49]. These drought indices serve as monitoring tools and operational indicators for regional water resources management. The most widely used and tested worldwide drought index is Palmer Drought Severity Index (PDSI) developed by Palmer [67] that considers precipitation, evapotranspiration and soil water holding capacity. The applicability of PDSI is limited due to the computational complexity, requirement of significant meteorological data and applicability on different time-scales. The Standardized Precipitation Index (SPI) developed by McKee [55], which is considered as simple and most widely used universal drought index by the World Meteorological Organization (WMO) is based on precipitation. SPI measures the drought index on different time-scales and enable to detect different drought types and it is widely accepted in the research community for drought monitoring and early warning [39] . Much effort has been devoted to developing techniques for drought analysis considering precipitation as the prominent variable in the context of Indian drought analysis [6, 69, 16, 75, 37, 36, 32, 66, 62, 46].

However, drought indices based on precipitation only can estimate the drought under lack of precipitation but are not able to detect the drought conditions under higher than normal atmospheric evaporative demand [101]. While, precipitation and temperatures are the main regional surface variables which are affecting under climate change due to the emission of greenhouse gases in the atmosphere, the increase in temperatures will directly have an impact on the severity of droughts [2]. Furthermore, to study the climate change impacts on droughts for future scenarios, a drought indicator considering precipitation may not be sufficient. A drought indicator which can include both precipitation and temperature into account will be more suitable, particularly under extreme heat waves[60] . The Standardized Precipitation-Evapotranspiration Index (SPEI) has been proposed by Vicente-Serrano [99, 100], which considers the

Potential Evapotranspiration (PET) in addition to precipitation and it can be used at several time scales. Due to the consideration of PET, SPEI combines the sensitivity of the Palmer Drought Severity Index (PDSI) and the probabilistic and multi-temporal nature of SPI. In the recent years SPEI has been widely used to evaluate drought events worldwide [1, 3] as well as for Indian subcontinent [46, 52, 65].

To this end, the use of PET in the drought estimation can characterize the intensification of drying areas where precipitation is already under stress and also tries to drive the areas into a drought that would experience modest drying when the effect of precipitation is considered alone [17]. Although PET-based drought indices consider the climatic water demand, it is limited towards the inclusion of the effects of regional land surface changes and actual moisture availability in the drought estimation. The SPEI is based on the climatic water demand as it considers the difference between P and PET which is estimated based on empirical techniques such as Thornthwaite model [99]. However, the PET is the energy-driven ET and underestimated by empirical techniques whereas the drought indices estimated based on AET will consider both climatic water demand and actual available moisture. The efforts made in the literature to include AET in the drought indices are Drought Severity Index (DSI) [63] and U.S. Drought Monitor (USDM) [89]. However, these indices use AET estimated using remote sensing datasets and vegetation information from normalized difference vegetation index (NDVI) and tries to account for land surface changes implicitly. Recently, Kim and Rhee [45] developed the Standardized Evapotranspiration Deficit Index (SEDI) using the Actual Evapotranspiration estimated from Bouchet hypothesis and the structure of SPEI as a fully ET-based drought index without consideration of precipitation. Accurate Estimation of ET at river basin scales is necessary for drought management. The next section 1.2.2 explains in more detail the importance of Evapotranspiration(ET) in drought analysis.

1.2.2 Evapotranspiration in Climate Analysis

At regional scales, the ET flux is a more complex process influenced by the regional climate, land use changes due to human interventions in the landscape, water withdrawals from the rivers for agricultural practices etc. [7, 77]. ET can be studied in terms of Potential Evapotranspiration (PET), Actual Evapotranspiration (AET), Reference Evapotranspiration and Pan Evaporation. Among these four variables, PET and AET which represent the atmospheric evaporative demands based on energy and available water supply, are commonly used regional hydrological variables. PET can be estimated based on empirical techniques such as Thornthwaite model [99, 100]. Such empirical techniques used to estimate the PET based on the concept that the climatic moisture demand may exceed available moisture. Therefore, PET is the maximum possible moisture loss limited only by the energy endowment or it is the energy-driven ET [84]. Further, the application of empirical equations for the estimation of ET may lead to an underestimation of PET in humid and tropical regions [95]. Whereas, the Actual Evapotranspiration (AET) represents the transfer of moisture from the surface to the atmosphere in response to both the energy demand and available moisture supply.

Several parametric models have been developed for estimating Actual Evapotranspiration (AET) flux, based on the assumption that AET is limited by the water availability in terms of Precipitation (P) under very dry conditions and energy availability in terms of Potential Evapotranspiration (PET) under very wet conditions ([9, 28, 94]). Such empirical models, serve as basis for deriving long-term mean annual water balances and have evolved based on various climate, soil, vegetation conditions [87, 91]. These, parametric models are region-specific and are based on various hydro-climatic conditions which necessitates calibration in the hydrological partitioning of water-energy variables [5]. Various catchment processes can be conceptualised by introducing model parameters whose values can be determined through calibration (Biswal, 2016). In this context, application of calibration factors accounting for various process of net radiation [16], plant-available water [107], vegetation dynamics [24], water balance equation[42, 49] became widely applicable. Furthermore, the parametric ET formulations considering precipitation and temperatures can represent purely climate-determined estimates of ET [21]. To study the long-term hydro-climatic changes of ET, the basin averaged time-invariant calibration factors were introduced with the consideration of closure of water-balance by [42, 5] . However, use of such time invariant calibration factors estimated over a specified period in the hydrological partitioning of the river basin may limit to acquire the temporal variability of water-energy balance variables. Also, given the changes of global ET under increase of temperatures and changes in precipitation patterns under anthropogenic climate change [64, 44], implementation of such single constant calibration factors may not include the hydro-climatological variability of water balances at catchment scales. Therefore, the present study proposed data driven regional relationship relating annual P, ET and runoff (R) estimated using hydrological modelling(Section 1.2.3) at catchment scale accounting for the water balances. The obtained relationship can be further used with projected P, AET and R under climate change signals to estimate water-energy balance variables and drought for future scenarios under climate change.

1.2.3 Hydrological Modelling

A hydrological model is a simplification of a real world system(hydrological cycle) that aids in efficient estimation and management of water resources with the help of meteorological inputs like precipitation and temperature and other inputs like soils,vegetation,terrain etc to calculate surface runoff, evapotranspiration using the spatial variability of these inputs. It can be run in a sub daily,daily or monthly scale and mostly applied to river basins. According to [22], a rainfall-runoff hydrological model can be defined as a set of equations that helps in the estimation of runoff as a function of various parameters such as meteorological inputs, basin characteristics.

There are three different types of hydrological models.

1. **Empirical Models** such as Unit Hydrograph model are observation-oriented models and calculate surface runoff using a functional relationship from the existing meteorological data inputs and outputs without considering any characteristics of the basin [14].

2. **Conceptual Models** such as Stanford Watershed Model IV derives surface runoff mainly on physical properties of the basin, based on the observed or assumed empirical relationships among different hydrological variables.

3. **Physically based Models** such as VIC(Variable Infiltration Capacity) Model, SWAT(Soil Water Assessment Tool) are based on deriving surface runoff solving differential equations describing the physical laws of mass, energy, and momentum conservations. They are further divided into two different models based on the entity assumption.

   - **Lumped hydrological Model** in which the total study area is considered as a single hydrological entity and estimations are done as a whole.
   - **Distributed hydrological model** in which the basin is further divided into sub-basins where each sub basins functions as a hydrological response unit(HRUs) that are comprised of unique land cover, soil, and management combinations. To further simplify the analysis, gridded hydrological models are used in which rather than an entire sub-basin, grids are considered as HRUs. Here, runoff is individually calculated in each grid and then a routing model is used to calculate the surface runoff at specified outlets of the basin.

The following are the key characteristics of Water Balance Model

1. The model should be transferable between geographical locations, and model parameters should be physically relevant.

2. The model should be applied to Hydrologic response units (sub-basins) or a regular grid and

3. Runoff routing must be from the point of generation or the grid cell/sub-basin through the spatial domain along the river network

1.2.4 Climate projections

Climate change impact assessment studies have been advanced due to the availability of General Circulation Models (GCMs) as the most credible tools for investigating the physical processes of the earth surface-atmosphere system. The GCMs can simulate the projections of climatological variables for current as well as for future scenarios accounting for greenhouse gas emission scenarios. These are the numerical models, which analyze the atmosphere on an hourly basis in all three-dimensions based on the law of conservation of energy, mass, momentum and water vapor and ideal gas law [54]. These are complex computer simulations describing the circulation of air and ocean currents and how the energy is transported within a climate system. GCMs are classified as Atmospheric General Circulation Models (AGCM) or Oceanic General Circulation Models (OGCM) for modeling atmospheric and oceanic circulations [54]. Most of the climate change impact assessment studies mainly focus on the use of GCM outputs of various climatological variables and their integration with hydrological modelling [78, 90, 15].

The Intergovernmental Panel on Climate Change (IPCC) has been established by World Meteorological Organization (WMO) and the United Nations Environment Programme (UNEP) to provide scientific, technical and socio-economic information for understanding the climate change process. The

IPCC provides scientific information to the research community in terms of future possible climate change scenarios for policy and decision making [103, 68] . The IPCC has developed long term emission scenarios based on the radiative forcing, demographic, technical and socio-economic information, which are considered as a standard reference to be followed for the policymakers, scientists and other experts. Such emission scenarios enabled the scientific community to carry out climate change analysis, modelling, impact assessment, adaptation and mitigation studies. Based on the Assessment Report 4 (AR4), IPCC has defined Special Report on Emission Scenarios (SRES) of four storylines as A1, B1, A2 and B2 determined by driving forces such as demographic development, socio-economic development and technology change along with CO2 level changes [68] [1]. Whereas, the IPCC Assessment Report 5 (IPCC 2014) has replaced the SRES of AR4 with Representative Concentration Pathways (RCPs) RCP8.5, RCP6, RCP4.5 and RCP2.6. Here, the RCPs refer to time-dependent projections of atmospheric greenhouse gas concentrations[2] and the numbers 8.5, 6, 4.5, and 2 represent the radiative forcing, expressed as Watts/m2. For example, RCP 8.5 is high pathway for which the radiative forcing reaches greater than 8.5 Watts/m2 by the year 2100 and continues to rise. The RCP 6 and 4.5 are considered to be stabilization pathways, while for RCP 2 the radiative forcing peaks at approximately 3 Watts/m2 before the year 2100 and then declines. Integration of projected climatological variables under climate change scenarios with water resources decision and management models to study the impact assessment over water quantitative and qualitative availabilities and demands has been widely gaining much attention in the research community [33, 73, 78, 61].

Assessment of climate change impacts on water resources necessitates accurate projections of various hydroclimate variables, which involves downscaling the projections of climate variables to hydrological variables. It is of growing importance to create accurate projections of hydrometeorological variables by employing climate model outputs with General Circulation Models (GCMs) and Regional Circulation Models (RCMs) which can then be statistically or dynamically downscaled [26]. To obtain the projections of hydro-meteorological variables (precipitation, runoff, temperature etc.) at regional scales based on large-scale climate simulations (mean sea level pressure, wind speed, humidity etc.) obtained from GCMs, downscaling models have been advanced [40, 103, 93, 4, 76]. Coarse resolution of the Global Circulation Model(GCM) climate projections of the order of 2.5°-3.5 °(approximately 200km-500km) makes this data inefficient for meteorological and hydrological analysis. Downscaling is a technique by which these coarse resolution inputs can be re-estimated to a finer resolution of 0.25°(or) 0.5°(25-50km)(Spatial Downscaling) in monthly or lesser(sub monthly,daily) time scales(Temporal Downscaling).

The following are the important considerations and assumptions to keep in mind when performing downscaling [93].

[1] https://www.ipcc.ch/assessment-report/ar4/
[2] https://www.skepticalscience.com/rcp.php

- Local climate(predictand data) is a combination of large-scale climatic/atmospheric features(global, hemispheric, continental, regional - predictors) and local conditions (topography, water bodies, land surface).

- The downscaling technique should be based on a climate variable which does not exhibit large sub-grid scale variations. Variables such as precipitation, elevation change dynamically for a few kms, hence they can't be reliable as predictors.

- The variables used in the downscaling process should be a direct model output (e.g. sea level pressure) and not outputs based on parameters involving other model variables.

Broadly, the downscaling techniques are classified as dynamic and statistical downscaling models. The statistical downscaling model involves deriving empirical relationships between large-scale climate variable simulations (predictors) obtained from GCMs and regional scale hydro-climatological variables (predictands- Precipitation and Temperature) [103], to predict future local climate variables. It is computationally less intensive but relies heavily on historical climate observations and the assumption that relationships currently observed will hold for the future. The spatial resolution of statistical downscaling climate change projections depends on the scale of the regional hydrological variables and also the spatial resolution of GCMs, which are generally coarse ranging from 2.8°X2.8°to 1.1°X1.1°.

Dynamical Downscaling takes help of additional data and physical processes in models at high resolutions such as Regional climate models (RCMs) and embedding into an existing GCM. A regional climate model is a dynamic model similar to a GCM, but it can be thought of as being composed of three layers. One layer is largely driven by the GCM itself, another layer builds on available specific regional data, and the final layer based on its own physical processes to resolve the model using the data from other two layers. The RCMs work at finer resolution and provides better dynamic downscaling climate change projections for a particular region [12] with region specific parameterization [86]. The use of RCMs for impact assessment should be based on the evaluation of the climate projections with observed data given the debate on the use of RCM projections directly [72, 25]. Due to the fine resolution climate model projections of RCMs, which enables to synthesize the climate change projections and hydrological models to study the regional impact assessment studies, RCMs became widely applicable [19]. Further, RCMs provide dynamically downscaled GCM outputs at fine resolutions compared to coarser statistical downscaled outputs, which can be directly used in the impact assessment studies [88] by testing their performance with current climate [86]. Dynamic and Statistical downscaling can also be used together as a hybrid downscaling model. Dynamical Statistical downscaling is used to downscale GCM output with the help of an RCM and then use it to downscale RCMs to a finer resolution whereas Statistical-Dynamical downscaling model categorizes GCM outputs into a few characteristic states that are further used in RCM simulations.

The recent regional climate model outputs available through Coordinated Regional Climate Downscaling Experiment (CORDEX) are mainly associated with GCM projections from Coupled Model Intercomparison Project (CMIP5) [3], were downscaled with the RCMs run by various research institutes and are available for 14 domains covering the entire globe. The present study tried to incorporate such General Circulation Models (GCMs) and Regional Circulation Models (RCMs) CORDEX climate change projections of precipitation and temperatures for the assessment of regional hydrometeorological induced evapotranspiration fluxes at river basin scales. In this context, the present study will emphasize on the quantification of ET flux changes under hydro-climatological changes with respect to historical and future scenarios using GCM and RCM CORDEX outputs and attempts to analyze the regional climate induced changes of precipitation and temperature as well as observed variability of regional drought occurrence over Krishna River basin including precipitation, potential evapotranspiration and actual evapotranspiration at the basin scale. The Krishna river basin, India, is described in more detail in Section 1.3 whereas the various datasets used in the study are described in Section 1.4.

1.3 Study Area

The study was conducted on Krishna river basin, which is the fifth largest river system in India. Krishna River basin occupies an area of 2, 58, 948 km^2 which is 8% of the total geographical area of the country. Nearly 44% of the basin lies in Karnataka, 26% of the basin falls in Maharashtra, about 15% in Telangana and another 15% in Andhra Pradesh within the range 73°17′-81°9′E and 13°10′-19°22′ N as shown in the figure 1.1.

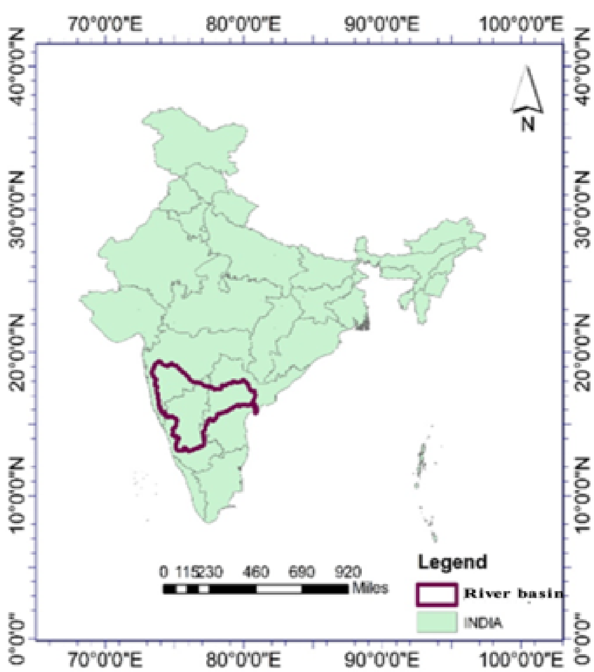


Figure 1.1: Krishna river basin in the Indian Subcontinent

[3]http://cmip-pcmdi.llnl.gov/cmip5/

The river originates in the Western Ghats and flows for about 1400 km before reaching to the Bay of Bengal. The major tributaries of the river are Ghataprabha, Malaprabha, Tunga-Bhadra, Bhima, Vedavathi, and Musi. There are two major cropping seasons: Kharif occurs from June to November and Rabi from December to March [30, 31].

Most of the Krishna river basin is covered by arid climate (Figure 1.2) with annual average precipitation in the basin as 784 mm, of which approximately 90% occurs during the South-West Monsoon from June to October. Some parts of the Krishna River basin, especially the Rayalaseema area of Andhra Pradesh, Bellary, Raichur, Dharwar, Chitradurga, Belgaum and Bijapur districts of Karnataka and Pune, Sholapur, Osmanabad and Ahmednagar districts of Maharashtra are drought-prone (Source: [4].

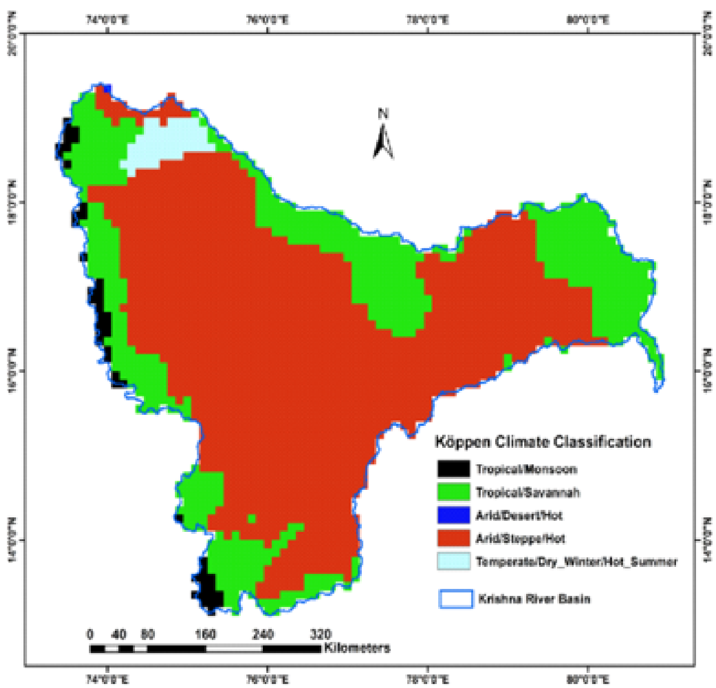


Figure 1.2: Climate classification in Krishna river basin

Figure 1.3 shows the number of drought-affected districts in each state of Andhra Pradesh, Maharashtra, Karnataka, and Telangana for the period of 2000 to 2016 based on the Farmer's portal, Department of Agriculture & Cooperation and Farmers Welfare, Ministry of Agriculture and Farmers Welfare, Government of India[5]. The number of drought-affected districts is reported more in Maharashtra(which is in the northern part of the basin) and few number of drought-affected districts in the undivided state of Andhra Pradesh. Most number of drought-affected districts are observed for the years 2002, 2009, 2014 and 2015 (Figure 1.3) over Krishna River basin.

Recently, severe drought has been experienced in the Krishna river basin during 2001 to 2004 where surface water resources were almost entirely committed to human consumptive uses, groundwater was

[4]http://india-wris.nrsc.gov.in/wrpinfo/index.php?title=Krishna
[5]http://farmer.gov.in/Drought/Droughtreport.aspx

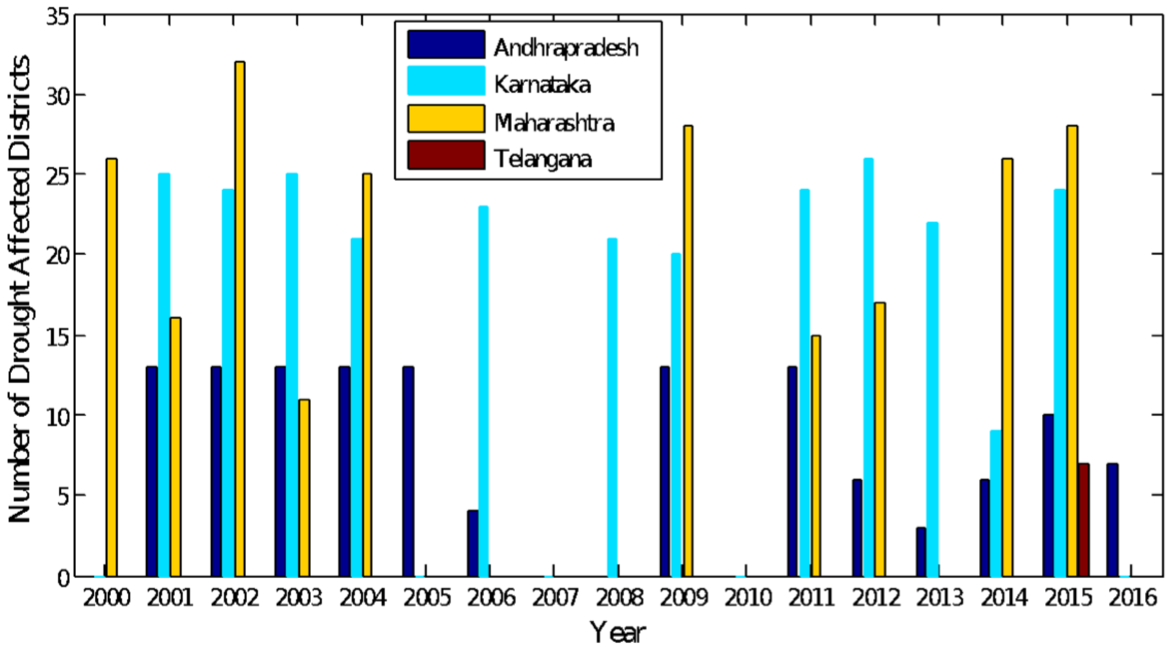


Figure 1.3: Drought affected districts by state from 2000 to 2016

over-abstracted and the discharge to the ocean almost nil [98]. A significant increase in the severity of droughts was reported over Krishna river basin during the period of 1948-2012 based on the study of Shah and Mishra [81]. Therefore, the present thesis has aimed to study the drought variability over Krishna river basin under current and future projections under climate change.

The observed India Meteorological Department(IMD) data estimates for precipitation and temperature are available for a common period of 1951 to 2014, whereas observed discharge estimates from Central Water Commission(CWC) are available from mid 1965 to 2015. Hence, a period of 1966-2014 is chosen for the current period analysis. The RCM and GCM estimates are available from 2006-2100 for the future periods. To get a better picture of the impacts of climate change, the future period is divided into three time slices of 2021-2040, 2041-2060 and 2061-2080. The detail description of the data is presented in the Section 1.4

1.4 Datasets

1.4.1 Meteorological Data

1.4.1.1 Precipitation Dataset

The gridded daily precipitation data from the India Meteorological Department (IMD) available for the period of 1901-2015 at 0.25°X 0.25°resolution was used as precipitation observational dataset [74].

1.4.1.2 Temperature Dataset

The gridded daily mean temperature data from the India Meteorological Department (IMD) available for the period of 1951-2014 at 1°X 1°resolution was used as temperature observational dataset. The temperature was interpolated to 0.25°X 0.25°resolution using the inverse distance weighting method.

The daily precipitation and temperatures data sets obtained from IMD present in .grd format at 0.25°* 0.25°resolution available over the Indian land mass were cropped for Krishna River Basin. They were aggregated over monthly time scale for a common time period of 1951-2014 to serve as primary inputs to calculate surface runoff and drought indices at 0.25°X 0.25°resolution.

1.4.2 Hydrological Data

1.4.2.1 Discharge Dataset

The discharge data in m^3/s was obtained from Central Water Commission (CWC) from Krishna Godavari Board, India for about 25 discharge locations for the period of 1965 June to 2015 June on a daily scale. This data after baseflow separation was also aggregated over monthly time scale in order to compare with calculated surface runoff and for validation of the hydrological model developed.

1.4.2.2 Digital Elevation Model(DEM) Data

The Digital Elevation Model (DEM) data with a resolution of 30-arc second (approximately 1km) was collected from Global 30 Arc-Second Elevation (GTOPO30) dataset provided by USGS (U.S Geological Survey). Using raster extraction in Quantum Geographic Information System (QGIS), the KRB basin was delineated using the DEM data and a binary mask file was created(Appendix A). Elevation varies from 50m to 900m in the basin.

1.4.3 Observational Data

1.4.3.1 ET Reference Data

The satellite-based land surface global ET product derived from the Numerical Terradynamic Simulation Group was adopted to study the strength of the proposed ET models [6]. This continuous satellite-derived global land surface ET was developed based on Moderate Resolution Imaging Spectroradiometer (MODIS) data, meteorological observations and satellite-based vegetation parameters. The ET data accounts for the canopy transpiration and soil evaporation with modified Penman-Monteith approach, biome-specific canopy conductance from Normalized Difference Vegetation Index (NDVI) and open water evaporation from Priestley-Taylor approach [106], which was found in general agreement with most of the global basins [50]. The ET data was extracted for KRB at 0.5°X0.5°resolution from 1983 to

[6] http://files.ntsg.umt.edu/data/ET_global_monthly_ORIG/Global_HalfDegResolution

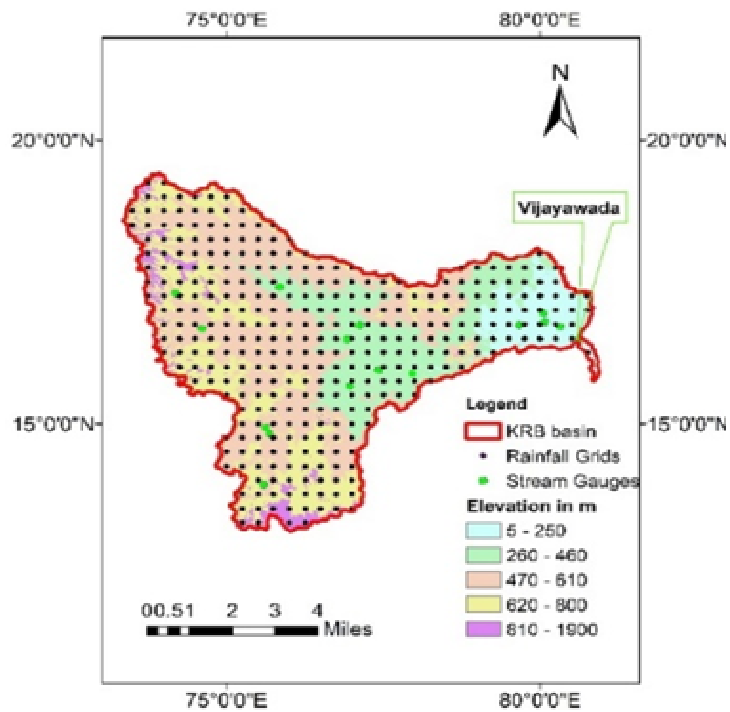


Figure 1.4: Various hydro-meteorological stations and elevation map superimposed on the Krishna river basin

2006 monthly data in HDF5 format using the 'hdf5read' command in MATLAB. The original ET data at 0.5 °X0.5°resolution was rescaled to 0.25°X0.25°resolution by bilinear spatial interpolation method.

Another satellite-based ET dataset, Global Land Evaporation Amsterdam Model (GLEAM) used in the study has estimates of the different components of Evapotranspiration: Potential Evapotranspiration, Actual Evapotranspiration, Soil evaporation etc. Potential Evapotranspiration (Ep) is estimated using the Priestly-Taylor approach [106] using surface net radiation and near surface air temperature and then converted into Actual Evapotranspiration by applying an evaporative stress factor S derived based on observations of microwave Vegetation Optical Depth (VOD) and estimates of root-zone soil moisture [53, 58]. This dataset at 0.25°X0.25°resolution updated for the time period 1980 to 2018 was interpolated to the required grids. Daily estimates are accumulated at monthly and annual scales and compared to model and water balance estimates of ET.

1.4.3.2 NCEP/NCAR Reanalysis Datasets

The NCEP/NCAR Reanalysis data set is a continually updated globally gridded dataset that represents the state of the Earth's atmosphere, incorporating observations and numerical weather prediction (NWP) model output from 1948 to the present time. It is a joint product from the National Centers for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR) [7]. This

[7]https://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.html

data is available at a spatial resolution of 2.5 °x 2.5 °in sub-daily,daily and monthly timescales. Monthly data of Specific Humidity(shum), Air Temperature(air), Zonal Wind Velocity(uwnd), Meridional Wind Velocity(vwnd) and Mean Sea Level Pressure(slp) are considered from the period of 1951 to 2014 for calibration and validation of the statistical downscaling model.

1.4.4 Climate projections Data

1.4.4.1 RCM Cordex Data

CORDEX (Coordinated Regional Downscaling Experiment) is mainly associated with GCM projections from Coupled Model Intercomparison Project (CMIP5) [8] and were downscaled with the RCMs run by various research institutes. CORDEX data sets are available for 14 domains covering entire globe, and present study selected South-Asian domain of the CORDEX project from Centre for Climate Change Research, Indian Institute of Tropical Meteorology, Pune, India [9].

Daily precipitation and temperature data simulated by 3 RCMs, driven by various GCMs, were obtained from the CORDEX [10]. The three CORDEX experiments considered are :

1. RegCM4(LMDZ), The Abdus Salam International Centre for Theoretical Physics (ICTP) Regional Climatic Model version 4 (RegCM4; [35]), with deriving GCM as IPSL LMDZ4, from Centre for Climate Change Research (CCCR), Indian Institute of Tropical Meteorology (IITM), India;

2. CCLM4(MPI), Consortium for Small-scale MOdelling (COSMO) model in CLimate Mode version 4.8 (CCLM; [23]) , with deriving GCM as Max Planck Institute for Meteorology, Germany, Earth System Mode (MPI-ESM-LR; [34]), from Institute for Atmospheric and Environmental Sciences (IAES), Goethe University, Frankfurt am Main (GUF), Germany;

3. REMO2009(MPI) Regional model, with deriving GCM as MPI-ESM-LR [34], from Climate Service Center, Hamburg, Germany.

The projections for the period of 2006 to 2080 were analysed under the Representative Concentration Pathway (RCP) 4.5 representing atmospheric radiation at 4.5 $Wm^{-}2$ at the end of 2100.

1.4.4.2 GCM Data

GCM projections from the fifth phase of the Coupled Model Intercomparison Project (CMIP5, [11], are considered to study climate change in the IPCC AR5 assessment process. The CMIP5 includes two types of climate experiments such as the

[8] http://cmip-pcmdi.llnl.gov/cmip5/
[9] http://cccr.tropmet.res.in/home/index.jsp
[10] www.cordex.org
[11] http://cmip-pcmdi.llnl.gov/cmip5/

1. long-term projections (century time scale and longer scenarios)

2. near-term predictions (10–30 yr), also called decadal prediction experiments.

Monthly data of Specific Humidity(shum), Air Temperature(ta), Zonal Wind Velocity(uwnd), Meridional Wind Velocity(vwnd) and Mean Sea Level Pressure(slp) are extracted for the historical period of 1950 to 2005 and for the future period of 2006 to 2100 from the IPCC Data Archive.

The GCM experiments considered to assess the climate change in Krishna river basin are

1. CanESM2 (The Second Generation Earth Model) derived from the Canadian Centre for Climate Modelling and Analysis, Canada

2. MIROC-ESM derived from the Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology Japan.

3. BCC-CSM1.1(m) (The First Generation fully coupled Climate System Model including atmosphere, ocean, land, and sea-ice components and incorporating global carbon cycle and dynamic vegetation cover) derived from the Beijing Climate Center,China

1.5 Data preprocessing

Most of the data preprocessing was done in MATLAB and QGIS. A meshgrid following the shapefile of the basin was created in MATLAB. The precipitation and temperature datasets present in .grd format were read using the 'fopen' and 'fread' commands in MATLAB. A common time period of 1951 to 2014 for the precipitation and temperature data was chosen to estimate the PET and AET. To estimate the observed surface runoff, baseflow has been separated from the observed discharge using the BaseflowSeparation function in CRAN EcoHydrology Package in R. Then taking into the account the time period of observed discharge, hydrological analysis has been carried out from 1966 to 2014 and comparison with the ET dataset has been carried out from 1983 to 2006 as well as from 1980 to 2012 considering the other observational ET dataset. PCRaster was used to estimate the surface runoff at each cell and for routing the runoff at Vijayawada. PCRaster clone maps were created based on the shapefile. The overview of the different events in a drought prediction model including the software and applications used is given in the Figure 1.5.

The thesis aims to develop a hydrometeorological drought prediction model by considering precipitation, evapotranspiration and surface runoff, the objectives of the thesis are explained in more detail in Section 1.6.

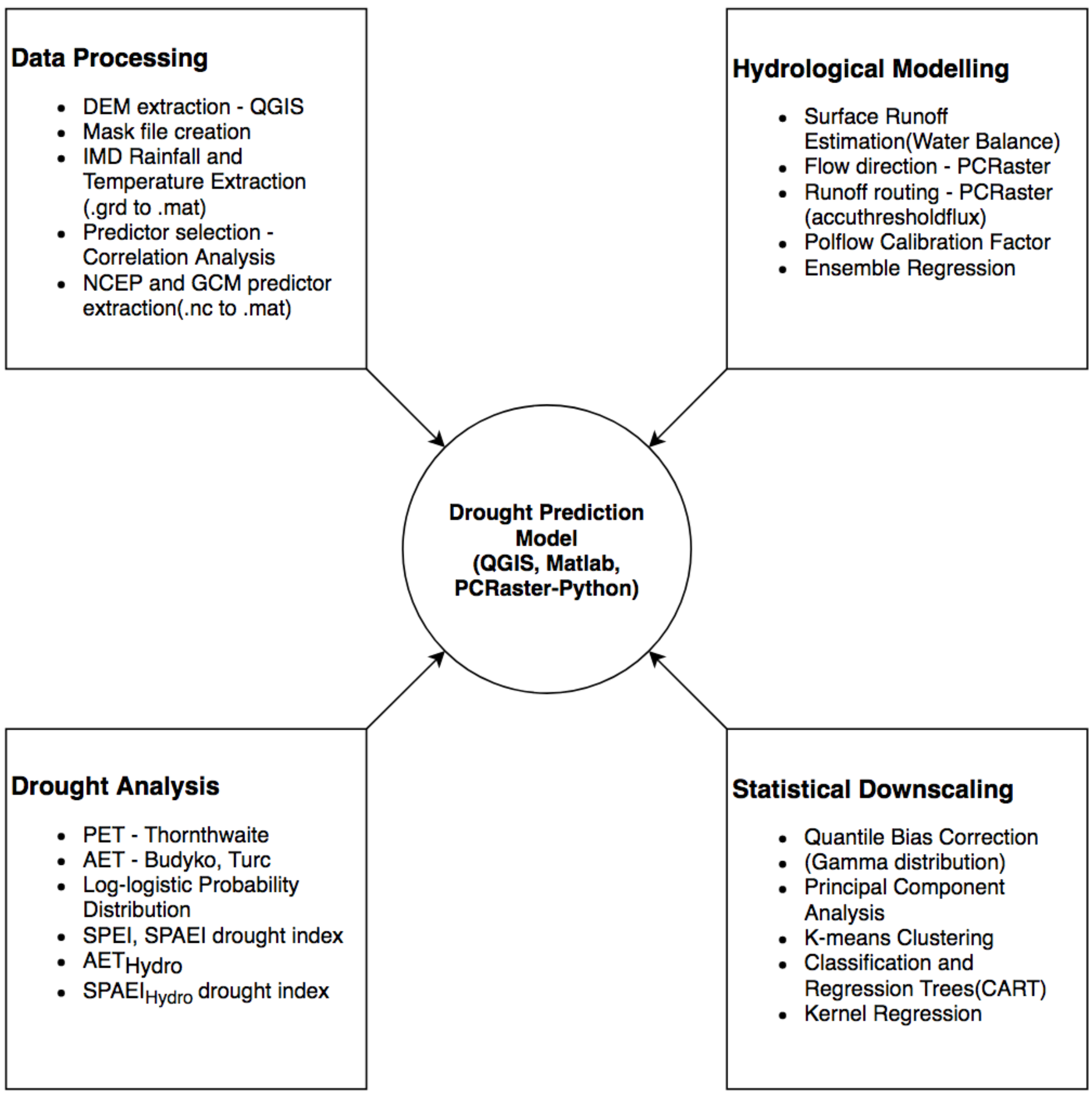


Figure 1.5: Events in a Drought Prediction Model

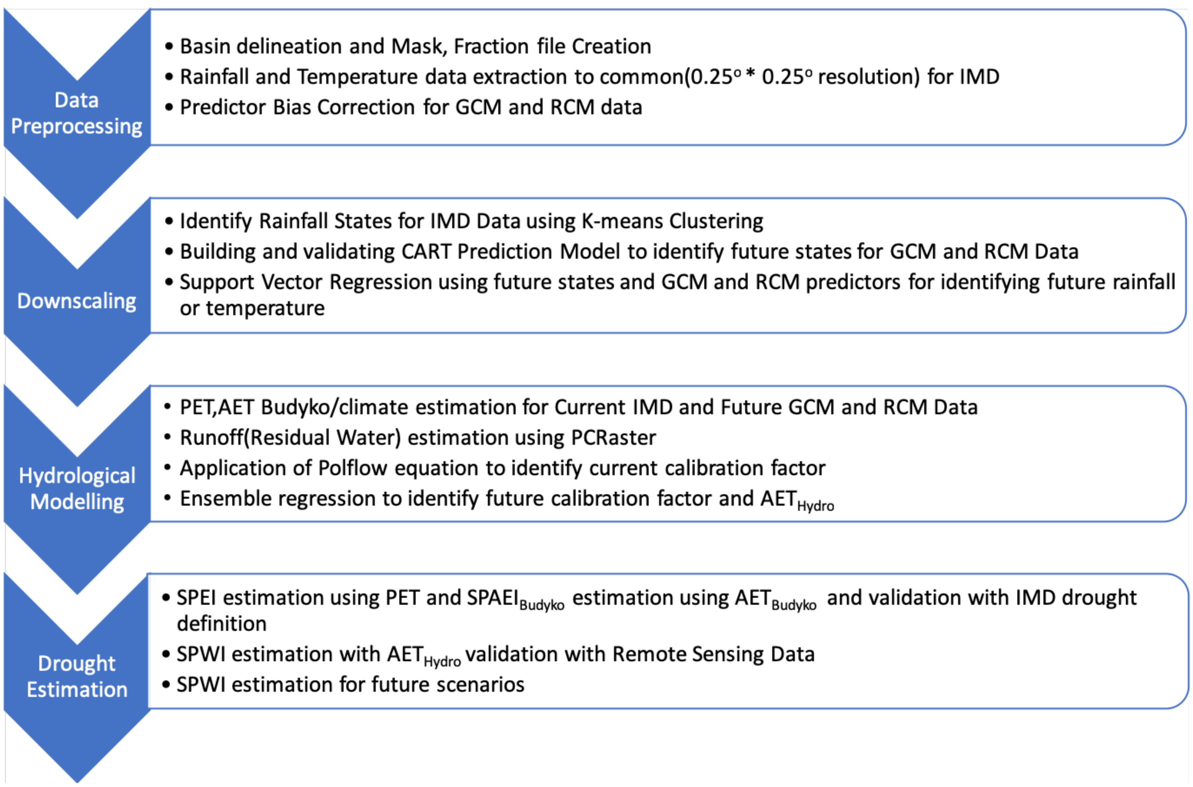


Figure 1.6: Drought Estimation Methodology for Current and Future Scenarios

1.6 Objective of the Thesis

- To develop hydrologically induced AET which can account for the precipitation, potential evapotranspiration and runoff into account
- To develop a hydrometeorological drought prediction index by considering the modelled hydrologically induced AET
- To study the climate change impact on hydrometeorological drought index using Regional Climate Models (RCMs) and General Circulation Model (GCM) outputs.

1.7 Thesis Organisation

This thesis is organised into 5 chapters as follows.

- Chapter 1 provides an introduction to the drought assessment problem and its motivation and objective are discussed as well as a detailed literature review, and an introduction to the study area and the datasets used are mentioned. A flowchart 1.6 containing the methodology is also presented.
- Chapter 2 contains the different methods used to estimate the Potential and Actual Evapotranspiration as well as the description of a regression model to calibrate the AET Hydrologically and an analysis of surface water and energy balances.
- Chapter 3 contains the comparison of various meteorological drought indices SPEI, $SPAEI_{Budyko}$, $SPAEI_{Turc}$ and hydrological drought index SRI, as well as the hydrometeorological drought index $SPAEI_{Hydro}$ for the current period using IMD Data.
- Chapter 4 contains the development of a statistical downscaling model for the GCM and RCM datasets and the analysis of water and energy balances and drought intensity, frequency and duration assessment in the future.
- Chapter 5 provides a summary of the concepts used and the findings of this study.

Chapter 2

Modelling of Hydrological Induced Regional Evapotranspiration

Introduction

Evapotranspiration (ET) is the largest flux next to rainfall and it is the most complex variable of the hydrological cycle affecting drought. ET can be studied in terms of Potential Evapotranspiration (PET) and Actual Evapotranspiration (AET), which are most important in the water balance partition. Also, the PET and AET represent the atmospheric evaporative demands based on energy and available water supply and are commonly used regional hydrological variables. PET is estimated based on the energy available for vaporization of water without considering the effect of moisture availability on the landscape. Most of the models to estimate PET are empirical and are temperature dependent [92, 38]. Whereas, AET is a complex variable to quantify since it is influenced by climate, vegetation, soil moisture and amount of water available on the landscape etc.

The conventional methods of estimation of AET include water balance, estimated as the difference between precipitation and runoff at the basin outlet. In this context, empirical models are also developed based on the assumption that AET is limited by the water availability in terms of precipitation under very dry conditions and energy availability in terms of PET under very wet conditions. However, these formulations are region specific and provide limitation over accounting for the effect of water balances and represent purely climate-determined estimates of AET. Introduction of basin averaged calibration factors considering the overall water-balance perspectives is valuable to study the long-term hydro-climatic changes of AET at catchment scales. The study proposed a modelling framework to estimate hydrologically calibrated AET at catchment scale. This chapter presents the analysis of hydrologically calibrated AET and study of its predictability with remote sensing data and water balance based AET data with Krishna river basin as a case study. The proposed hydrologically calibrated AET as presented in this Chapter will be further used in the drought estimation in addition to precipitation to develop hydrological drought monitoring tool at basin scale (Chapter 3).

2.1 Modelling of Actual Evapotranspiration using Empirical Methods

Evapotranspiration is the sum of evaporation (movement of water from earth's surface such as land, water bodies to atmosphere) and transpiration (movement of water from plants into the atmosphere). Methods to estimate the Potential Evapotranspiration (PET) are described in the subsection 2.1.1 and climate induced Actual Evapotranspiration (AET_{Clim}) are described in the subsection 2.1.2.

2.1.1 Potential Evapotranspiration

Potential Evapotranspiration is referred to as the maximum amount of Evapotranspiration that can occur when there is unlimited water supply. It is solely dependent on temperature and does not consider the effect of precipitation.

1. **Thornthwaite method**

   PET can be estimated based on Thornthwaite [92] model, which considers the monthly average air temperature and geographical location of the region of interest as input variables as follows:

$$PET = 16k\left(\frac{10T}{I}\right)^a \tag{2.1}$$

   where

   - T is the mean monthly temperature (°C)
   - I = Heat index

$$I = \sum_{j=1}^{12}\left[\frac{T_j}{5}\right]^{1.5} \tag{2.2}$$

     – T_j is the mean monthly temperature (°C) during the month 'j' for the location of interest

   - a = Location dependent coefficient

$$a = 6.75 * 10^{-7}I^3 - 7.7 * 10^{-5}I^2 + 1.8 * 10^{-2}I + 0.49 \tag{2.3}$$

     I (heat Index) and a (Location dependent coefficient) have the same units as Temperature°C

   - K is the correction coefficient depending on the latitude and month, given as follows:

$$k = \left(\frac{N}{12}\right)\left(\frac{NDM}{30}\right) \tag{2.4}$$

     – where NDM is the number of days of the month and N is the maximum number of sun hours as follows

$$N = \left(\frac{24}{\phi}\right)w_s \tag{2.5}$$

     – w_s is the hourly angle of sun rising, which can be calculated as follows

$$w_s = arccos(-tan\psi tan\delta) \tag{2.6}$$

– ψ is the latitude in radians. If δ is the solar declination, in radians and J is the average Julian day of the month, then δ can be estimated as follows:

$$\delta = 0.4093sen\left(\frac{2\pi J}{365} - 1.405\right) \tag{2.7}$$

2. **Hargreaves method**

   Hargreaves method is also one of the well accepted PET estimation methods in the literature and considers the maximum, minimum and mean temperature data and the geographical location of the region using the following equation

$$PET = 0.0023 * (T_{max} - T_{min})^{\frac{1}{2}} * (T_{mean} + 17.8) * R_a \tag{2.8}$$

   where T_{mean}, T_{min}, T_{max} are the mean, minimum and maximum temperature data respectively and R_a is the extra-terrestrial radiation expressed in equivalent evaporation units calculated from the latitude and time of the year.

Thornthwaite method is considered for the study based on the fact that it requires only average temperature data and geographical location of the region. Wwhereas PET estimation methods such as Penman-Monteith (PM) approach, which is generally considered as a standard method suggested by Food and Agriculture Organization (FAO), requires additional information such as regarding wind speed, maximum and minimum temperatures, relative humidity and solar radiation. Therefore, the study used Thornthwaite model for the estimation of PET and whereas Langbein method does not account for location based characteristics. Hargreaves method is is used to compare the estimates of PET based on Thornthwaite model.

2.1.2 Actual Evapotranspiration

Actual Evapotranspiration is the amount of water that is actually removed from a surface due to the processes of evaporation and transpiration [71]. Various empirical models have been developed for estimating AET which is based on the assumption that AET is limited by the water availability in terms of precipitation under very dry conditions and available energy under very wet conditions in terms of potential evapotranspiration [9, 28, 57, 107].

1. **Budyko method for estimating Actual Evapotranspiration(AET_{Clim}/AET_{Budyko})**

   One of the classical model to estimate AET relating long-term-average water and energy balances at catchment scales using precipitation and PET is Budyko Equation [9]. Budyko has developed a relationship between three hydro-climatic variables for a basin: Precipitation (P), Potential Evapotranspiration (PET), and Actual Evapotranspiration (AET). The Budyko hypothesis states that the ratio of the AET over precipitation (AET/P) is fundamentally related to the ratio of the PET over precipitation (PET/P) ([9, 28]) as follows:

$$\frac{AET}{P} = 1 + \frac{PET}{P} - \left(1 + \left(\frac{PET}{P}\right)^{w}\right)^{\frac{1}{w}} \tag{2.9}$$

The parameter 'w' accounts for the effects of climate variability, basin characteristics such as soil, vegetation, terrain, etc [24]. The present study used Budyko equation as implemented by [107] for estimating the AET, given as follows:

$$AET_{Budyko} = \left[P\left(1 - exp\left(-\frac{PET}{P}\right)\right) PETtanh\left(\frac{P}{PET}\right)\right]^{0.5} \tag{2.10}$$

The original Budyko equation(2.9) has been developed for a long-time scale (e.g. [10, 109]). However, the Budyko framework can be applied over short periods of monthly and annual scales (e.g. [108, 13, 49]), if the parameter 'w', which represents the joint effect of climate and land surface is estimated. For a reasonable application of the Budyko equation as developed by [107] (Eq. 2.10), we used a 12-month scale for the estimation of annual AET and corresponding drought indices at annual scale. Any timescale lower than 12 months might result in the accumulated precipitation tending to zero, thus resulting (PET/P) value tending to infinity, which is not suitable for the implementation of Budyko(Eq. 2.10). Therefore, the monthly precipitation and PET estimated based on Thornthwaite model at monthly scale were aggregated as given as follows:

$$P_i^k = \sum_{i-k+1}^{i} P_i \tag{2.11}$$

$$PET_i^k = \sum_{i-k+1}^{i} PET_i \tag{2.12}$$

Where P_i^k and PET_i^k are the accumulated precipitation(in mm) and PET(in mm) in month i. Next, the accumulated AET(in mm) to a k-month scale is estimated as:

$$AET_i^k = \sum_{i-k+1}^{i} AET_i \tag{2.13}$$

where k=6,12,18,24

2. **Turc method for estimating Actual Evapotranspiration(**AET_{Clim}**/**AET_{Turc}**)**

   Another AET model which also considers precipitation and PET and accounting for the soil and vegetative characteristics implicitly is Turc model [94]. It is also one of the widely used AET model in the hydrological applications (e.g. [85, 5, 42]). The Turc model estimates the annual AET (mm) by using accumulated precipitation 2.11 and accumulated PET 2.12 in mm as follows:

$$AET_{Turc} = \frac{P}{\sqrt{0.9 + \frac{P^2}{PET^2}}} \tag{2.14}$$

To understand hydrological induced AET accounting for the water balance of the river basin with runoff, R, and precipitation, the AET estimated from 2.10 should be calibrated hydrologically.

2.2 Hydrologically Calibrated ET flux at Catchment Scale using PCRaster

The accumulated precipitation surplus (P-AET), which is considered as the amount of water available is calculated. The amount of available water from each cell will be transported to the basin outlet from the upstream cells by following the elevation and local drainage map. At a grid cell, i, which can be considered as a basic hydrologic unit of a distributed hydrological model, the precipitation surplus or Residual Available Water, $RAW_{calculated,i}$ can be calculated using annual total Precipitation (P_i) and climate induced AET ($AET_{Clim,i}$) (from Budyko or Turc empirical models) as follows:

$$RAW_{calculated,i} = P_i - AET_{Clim,i} \tag{2.15}$$

Where the P_i and $AET_{Clim,i}$ are at annual scale in mm/year.

The discharge at the basin outlet $R_{Calculated,outlet}$ was estimated by accumulating the flow at grid cell, i, ($RAW_{calculated,i}$) and from all upstream grid cells ($RAW_{calculated,\Omega}$) according to the flow direction of the river and corresponding to the area of each grid cell (A_{Cell}) as follows:

$$R_{Calculated,outlet} = (RAW_{calculated,i} + RAW_{calculated,\Omega}) * A_{Cell} \tag{2.16}$$

$R_{Calculated,outlet}$ is the uncalibrated total runoff from the basin outlet at steady state and it is generally not consistent with observed runoff. Here, calibration factors can be introduced on AET, with consideration of closure of water balance using the outputs from the hydrological model to represent the observed runoff at the basin outlet [43, 5]. Basin averaged calibration factor X_{Cal}, was introduced to correct the uncalibrated AET over the river basin. By introducing a calibration factor on $AET_{Clim,i}$, accounting for the closure of water balance at the catchment outlet, a hydro-meteorological induced AET can be estimated. Therefore, the precipitation surplus or RAW accounting for the calibration factor, X_{Cal}, can be written as follows:

$$RAW_{actual,i)} = P_i - X_{Cal} * AET_{Clim,i} \tag{2.17}$$

$$R_{Observed,outlet} = (RAW_{actual,i} + RAW_{actual,\Omega}) * A_{Cell} \tag{2.18}$$

The calibration factors [43] for the entire river basin for each annual time scale can be estimated by comparing the observed and uncalibrated runoff estimated from the hydrological model at the basin outlet as follows:

$$\frac{R_{Calculated,outlet}}{R_{Observed,outlet}} = \frac{P - AET_{Clim}}{(P - (X_{Cal} * AET_{Clim})} \tag{2.19}$$

$$X_{Cal} = \frac{R_{Observed,outlet}}{R_{Calculated,outlet}} + (1 - \frac{R_{Observed,outlet}}{R_{Calculated,outlet}})\frac{\Sigma P}{\Sigma AET_{Clim}} \tag{2.20}$$

Where, $R_{Observed,outlet}$ and $R_{Calculated,outlet}$, are the long-term annual average observed and simulated runoff at the basin outlet in m^3/s respectively and ΣP and ΣAET_{Clim} are the long-term accumulated annual average observed precipitation and AET (Budyko or Turc empirical models) over the basin in mm/year. From the Equation 2.20, it is clear that calibration factor is a function of precipitation (P), AET, observed runoff($R_{Observed,outlet}$) and simulated runoff ($R_{Calculated,outlet}$). The annual scale basin averaged calibration factors estimated based on Eq. 2.20 can be applied on the AET_{Clim} (Budyko or turc models) to study the changes of AET under hydrometeorological or water-use over the river basin as given as follows:

$$AET_{Hydro} = X_{Cal} * AET_{Clim} \tag{2.21}$$

Therefore, the AET_{Hydro} represents the evaporative demand of the atmosphere accounting for energy available in terms of PET and water supply in terms of runoff. It can be noted that AET_{Clim} is considered to be climatological AET which is region specific, while AET_{Hydro} is computed by the water balance, in which the runoff was simulated by the distributed hydrological model, PCRaster. To study the changes in AET, such calibration factors should not be a single factor for the entire basin as followed by [43, 5]. However, the calibration factors should be function of P, AET and runoff at catchment scale as given by the variability of precipitation and temperatures under the variability of climate signals. Therefore, calibration factors are modelled by relating with the variation of P, AET and runoff observed and simulated, using a data driven model which can be employed further to study the changes in AET under climate and hydrological aspects corresponding to climate signals.

2.2.1 Ensemble Regression Model(ERM) for Modelling the Calibration Factors

To investigate the changes in the hydro-climatological induced AET (AET_{Hydro} under climate variability, the dependence between Precipitation (P), AET_{Clim}, $R_(Observed, outlet)$ and X_{Cal} for the entire river basin for the present climate was estimated. The uncalibrated runoff ($R_(Calculated, outlet)$) at the basin outlet will be used in the estimation of calibration factor X_{Cal} (Eq. 2.20) by comparing with observed basin outlet runoff($R_(Observed, outlet)$). Here a data driven model based on Ensemble Regression Model (ERM) algorithm [27, 11] was trained using P, $AET_{Clim,i}$ and $R_(Observed, outlet)$ as independent variables with calibration factor X_{Cal} as the response variable. The ERM is based on the principle that a diverse set of models can make better decisions in comparison to an individual model [27] and gained much attention in the hydrological assessments in recent years [79]. Ensemble regression models help to decrease variance between the predicted and observed values and will produce a more reliable estimate than a single regression model.

The least squares gradient boosting ensemble regression used for predicting calibration factors was used in the present study. Consider an input training dataset of 'N' points {X,Y} = $(x_i, y_i)_{i=1}^{N}$. where x_i is the set of predictors and y_i is the observed predictand value at the i^{th} timestep. Initially, all the input points(predictors) in the data set are given equal weighting coefficients(equal importance - $\alpha_i = \frac{1}{n}$) and a base model($F_0(x)$) to predict values of the form y=F(x) is trained on this dataset. At every iteration m, errors/residuals(r_{im}) are calculated between the observed predictand value y_i and model predicted value($F_{m-1}(x_i)$) and a base-learner(hm) is fitted to these residuals(r_{im}) using the Loss function 'L' in the direction of steepest gradient i.e Weight α_i of point 'i' is increased corresponding to a higher value of errors/residual (r_{im}). The model is sequentially updated as shown in Equation

$$F_m(x) = F_{m-1}(x) + argmin_h \sum_{i=1}^{n} L(y_i, F_{m-1}(x) + h(x_i)) \tag{2.22}$$

where the least squares loss function $L(y, F(x)) = \frac{1}{2}(y - F(x))^2$ is used to update the errors/residuals(r_{im}). This can be described in more detail using the following steps.
Initialize $F_0(x) = \bar{y}$. $F_0(x)$ or $\bar{y}$ is the set of observed predictand values.
For m=1 to M where M is the number of iterations do:

- $(r_{im}) = -\left[\frac{d\left(\sum_{i=1}^{N}[y_i - F_{m-1}(x_i)]^2\right)}{d(F_{m-1}(x_i))}\right]$ for i=1..N
- Fit a base learner(h_m) to these residuals using the training set $(x_i, r_{im})_{i=1}^{N}$
- $(\gamma_m) = argmin_\gamma \sum_{i=1}^{N}[y_i - (F_{m-1}(x_i) + \gamma h(x_i))]^2$
- $F_m(x) = F_{m-1}(x) + \gamma_m h_m(x)$

Where h is base learner added to $F_{m-1}(x)$ to improvise the model, an ideal value of h after m iterations would imply $F_m(x)$=y. γ is the multiplier that is applied to the base learner to get the updated model

The trained ensemble regression model will be further used to predict the calibration factors for the future scenarios under projected changes of P, AET_{Clim} and $R_{Calculated,outlet}$. Here, the $R_{Calculated,outlet}$ will be the uncalibrated runoff from the hydrological model with projected P and AET_{Clim} for the future scenarios. To quantify the changes in the dependence between calibration factors and hydrological variables, coefficient of determination (R^2) was selected as performance measure. Finally, the hydrological induced regional AET under climate change over the basin was estimated with projections of various Regional Climate Model (RCM) outputs of precipitation and temperatures and estimated AET_{Clim} and $R_{Calculated,outlet}$. The changes of climatological AET_{Clim} and hydrometeorological AET, AET_{hydro} at river basin scale was analyzed under climate signals with regional climate change projections. The estimated climatological induced AET (AET_{Clim}) and hydrological induced AET (AET_{hydro}) can be further used to develop a meteorological and hydrological drought index respectively (Chapter 3).

2.3 Results and Discussions

2.3.1 Variation of IMD Precipitation and Temperature: Current Scenario

The daily gridded precipitation dataset available from 1901 to 2015 and daily gridded temperature dataset available from 1951 to 2014 over the Indian land mass were cropped for Krishna River Basin. About 348 grid points at at 0.25°* 0.25°resolution encompassing the entire basin were considered in the drought and hydrological analysis. A common data period of 1951 to 2014 was considered for the drought analysis, which is further divided into three time slices of 1951-1971, 1972-1992 and 1993-2014, over Krishna River basin to understand the climate variability over the basin whereas the period

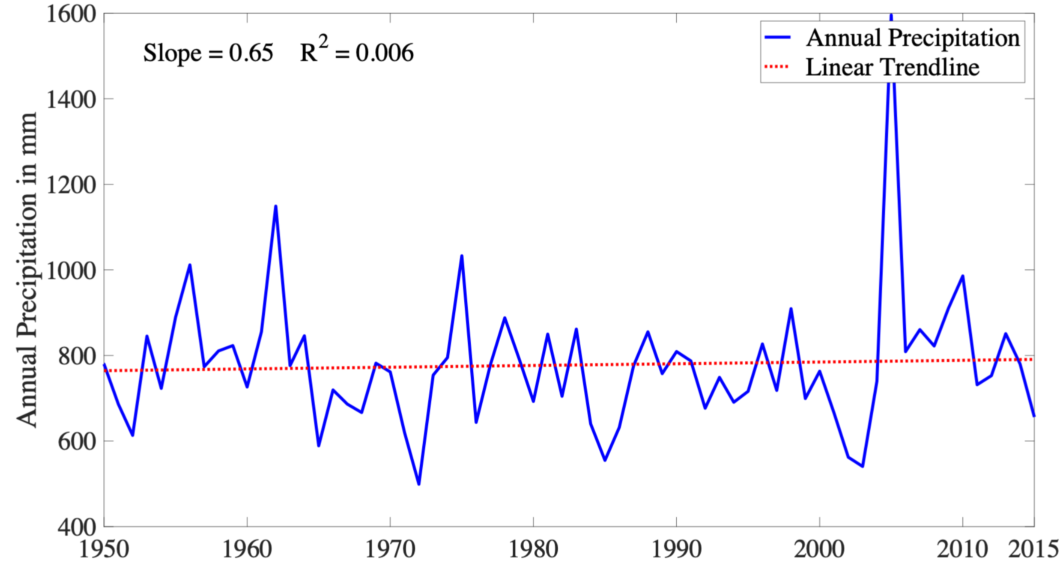


Figure 2.1: Average annual precipitation trend from 1951-2015

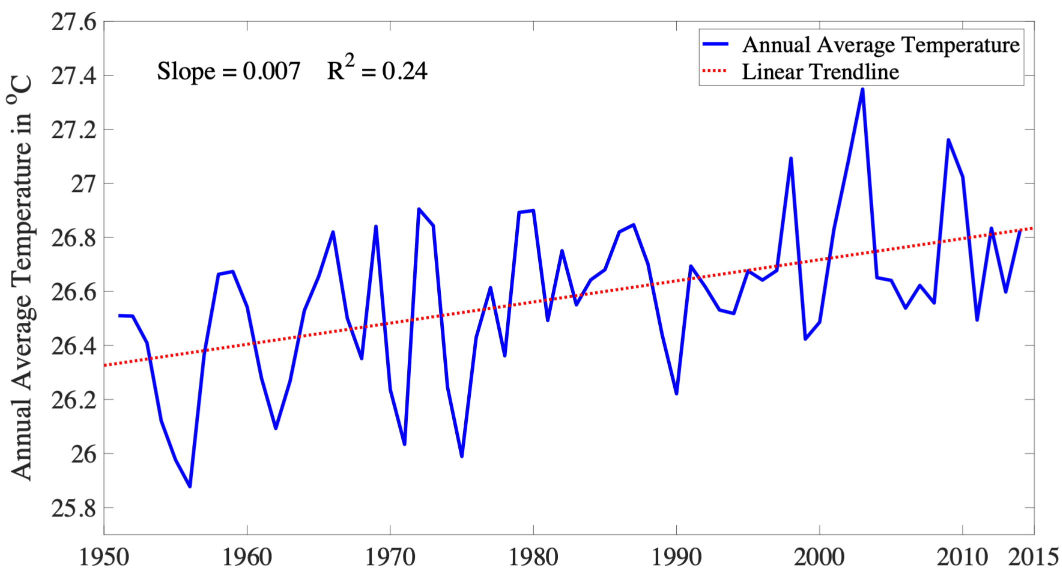


Figure 2.2: Annual Average Temperature trend from 1951-2014

between 1966 to 2014 is considered for Hydrological analysis based on the availability of discharge data(Runoff data). Further, to assess the period from where significant change has occurred in precipitation and temperatures, Pettitt's test [70] has been performed on the basin-averaged annual time series data from 1966 to 2014. To illustrate the changes in spatial distribution and temporal variability of regional water and energy variables over KRB, two-time intervals of 1965-2003 and 2004-2014 were selected as before and after the change-point of precipitation year of 2003 (Figures 2.6, 2.7).

The spatially averaged annual precipitation over Krishna river basin was estimated as 778.35 mm and varying between 499.1 to 1595.7 mm with Pettitt change point detection year as 2003 for the period of 1951 to 2014. The annual total precipitation has shown an increasing trend at a rate of 6.5 mm/decade with a significance level of 0.05 (Figure 2.1). The basin-averaged annual average temperatures have shown an increasing trend over Krishna river basin. Correspondingly, temperature also has shown an increasing trend of 0.1°C/decade (Figure 2.2) with Pettitt change point detection year as 1992.

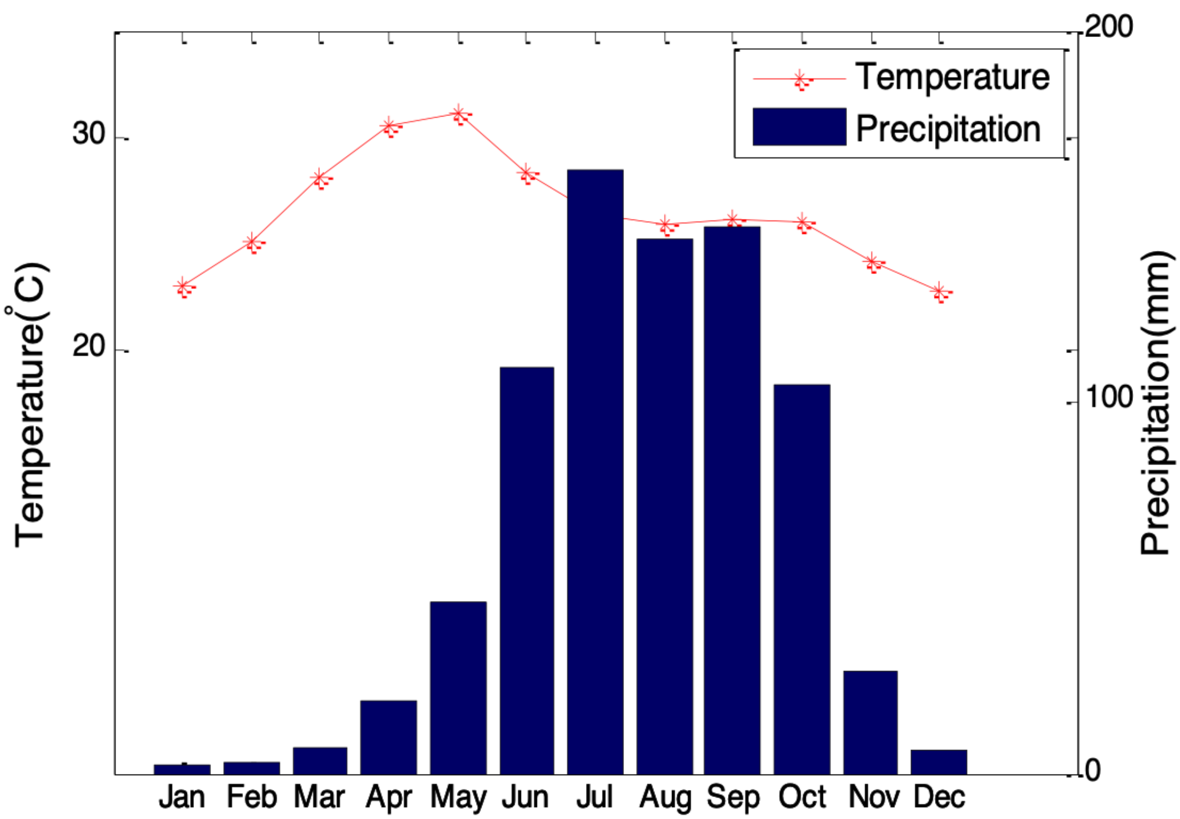


Figure 2.3: The spatial monthly average temperature and precipitation over Krishna River Basin for 1951-2014

The spatial averaged monthly variation of precipitation and Temperature for the period of 1951 to 2014 is shown in the figure 2.3, with rainfall contributing months as June to October, whereas, the dry months as March, April and May.

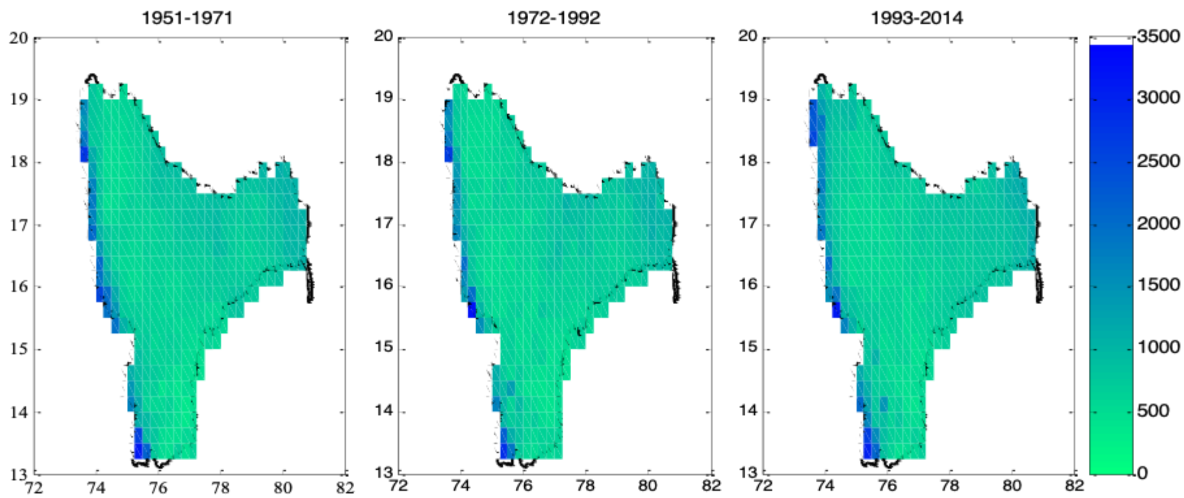


Figure 2.4: Spatial variation in annual total Precipitation in mm for 1951-1971, 1972-1992 and 1993-2014

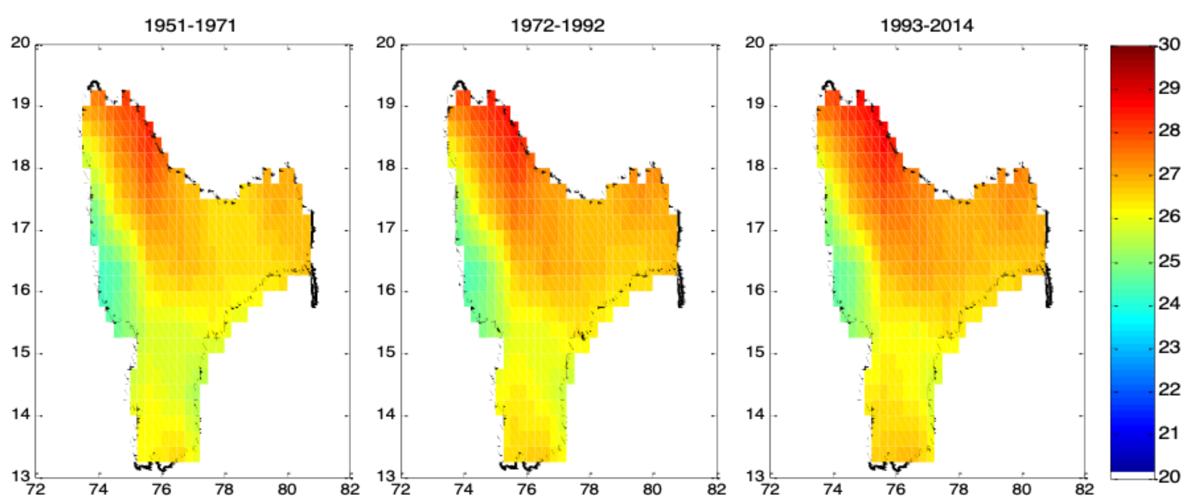


Figure 2.5: Spatial variation in annual average Temperature in °C for 1951-1971, 1972-1992 and 1993-2014

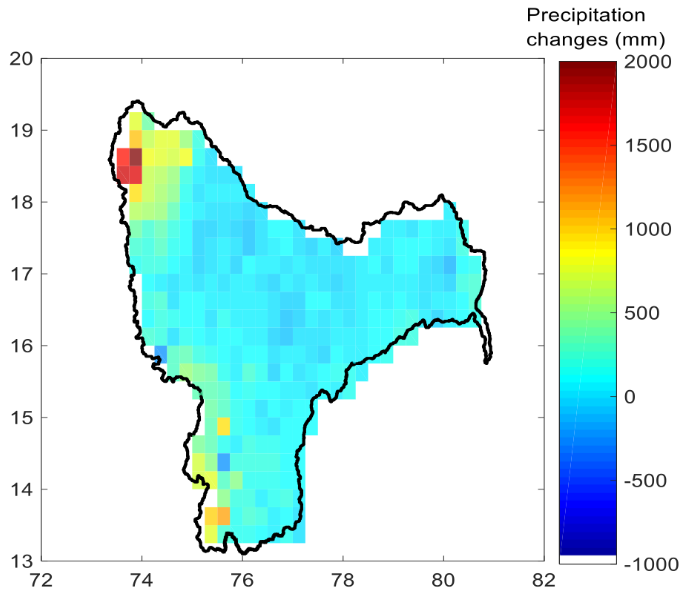


Figure 2.6: Spatial distribution of Average Precipitation changes between 1966-2003 and 2004-2014

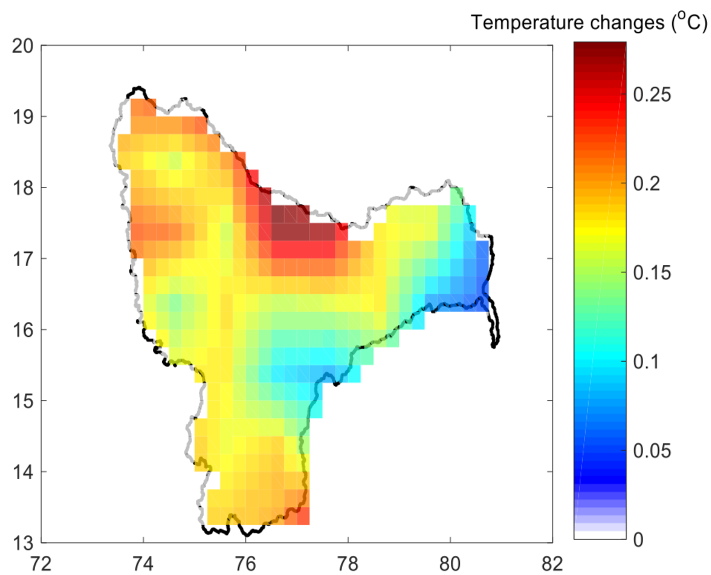


Figure 2.7: Spatial distribution of Average Temperature changes between 1966-2003 and 2004-2014

To examine the spatial pattern of precipitation and temperature changes, the annual total precipitation(Figure 2.4) and Temperatures (Figure 2.5) were estimated for three periods of 1951-1972, 1973-1992, and 1993-2014. The total annual precipitation amount was observed to be higher towards the Western Ghats boundary of the Krishna river basin. The annual total precipitation has not shown much spatial variation over the basin for the three time periods. Higher average temperatures were observed towards the upper most portion of the basin covering few districts of Maharashtra and Telangana.

Also, differences in average precipitation (Figure 2.6) and temperature (Figure 2.7) after and before the change point of 2003 have been plotted to understand the spatial variation. The upper and lower regions of KRB including few districts of Maharashtra and Karnataka especially the Western Ghats region has shown increase in precipitation from 1965-2003 to 2004-2014 (Figure 2.6), whereas the central and north-east parts of the basin has shown a decrease in precipitation, where arid climate persists(Figure 1.2). Overall, the precipitation has been increased by 161 mm from 1965-2003 to 2004-2014 over KRB. A positive change of annual average temperature from 1965-2003 to 2004-2014 was identified over the entire KRB with highest increase of about 0.3°C over few districts of Telangana.

The precipitation and temperature data from 1966 to 2014 were used to estimate the water-energy balance variables of PET and AET_{Clim}, AET_{hydro} and R over KRB and studied as current climate scenario.

2.3.2 Changes in surface water-energy balances over the basin: current scenario

Basin-averaged hydro-meteorological variables of precipitation, temperature, PET (estimated by Thornthwaite equation 2.1), AET (estimated by Budyko model 2.10) and R (estimated by PCRaster model 2.16) during periods of 1965-2003 and 2004-2014 were summarized in Table 2.1. The PCRaster model was applied for 348 cells of the KRB basin at 0.25°X 0.25°resolution at annual scale to estimate the uncalibrated basin outlet runoff at Vijayawada 2.16. The local drainage direction (ldd) was processed from DEM using 'lddcreate' PCRaster function. The entities were given in format of raster maps for spatio-temporal attributes, time-series for temporal non-spatial data, and look up tables (Figure 2.8). The entire river basin was discretized with a basic hydrologic unit of 1 km X 1 Km rectangular cell following the resolution of DEM and assigned rainfall and estimated AET (Eq. 2.10) for each cell.

The accuthresholdflux function [1] of PCRaster accumulates the precipitation surplus or flow in terms climatological (P-AET) residual available water or precipitation surplus for each cell and transported to the basin outlet from the upstream cells by following the elevation map (Figure 2.8(a)) and local drainage map (Figure 2.8(b)).

Then this uncalibrated runoff $R_{Calculated,outlet}$ at the basin outlet estimated using the period of 1966 to 2003 was used to estimate the calibration factors (Eq. 2.20) for AET_Clim by comparing with observed runoff at annual scale. The ERM was trained with annual P, AET, $R_{Calculated,outlet}$ and X_{Cal}, for a period from 1966 to 2003 (correlation coefficient as 0.95) and validated with 2004 to 2014 period (correlation coefficient as 0.59).

[1] http://pcraster.geo.uu.nl/pcraster/4.2.0/documentation/pcraster_manual/sphinx/op_accuthreshold.html

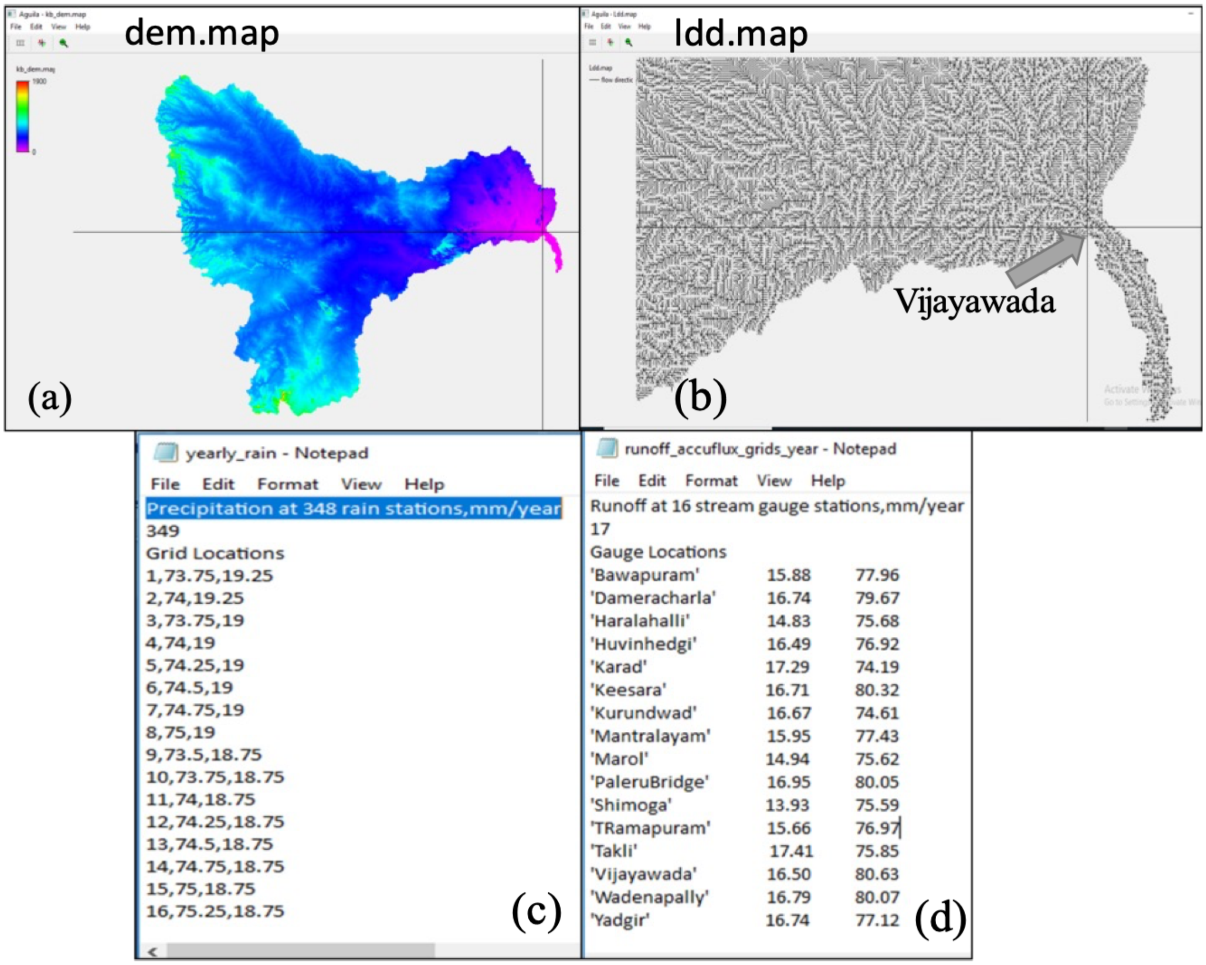


Figure 2.8: Example of entities: (a) digital elevation map (DEM) map, (b) local drainage direction (ldd) map (c) rainfall time series, (d) discharge locations

Hydrological Variable	**Time period**	
	1966-2003	**2004-2014**
Average Annual Precipitation(mm)	733.12	894.47
Annual Average Temperature (°C)	26.6	26.7
Total modelled PET - Thornthwaite(mm)	1803.74	1818.32
Total modelled AET - Budyko(mm)	644.21	724.35
Total hydrologically calibrated AET (mm)	669.81	788.54
Average observed runoff at Vijayawada (m3/s)	525.92	367.49
Average modelled runoff at Vijayawada (m3/s)	738.42	1036.71
Average calibrated runoff at Vijayawada (m3/s)	526.97	361.47

Table 2.1: Summary of spatial average annual water-energy variables for current (1966-2003, 2004-2014) and future period (2021-2040, 2041-2060) for KRB

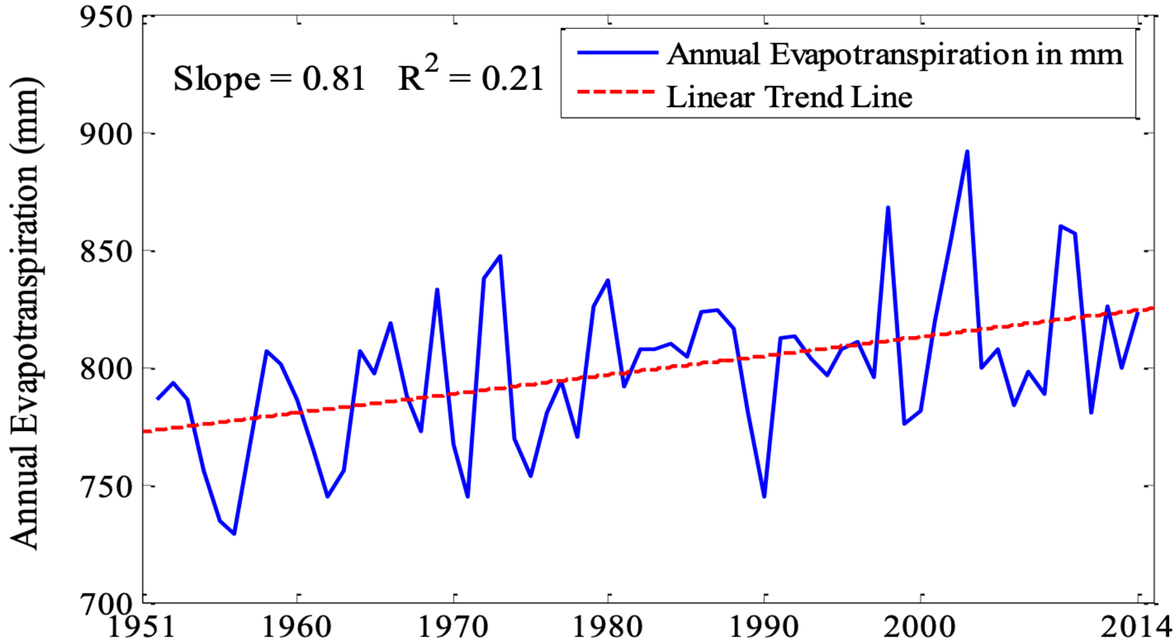


Figure 2.9: Temporal variation of Annual Potential Evapotranspiration(PET) from 1951-2014

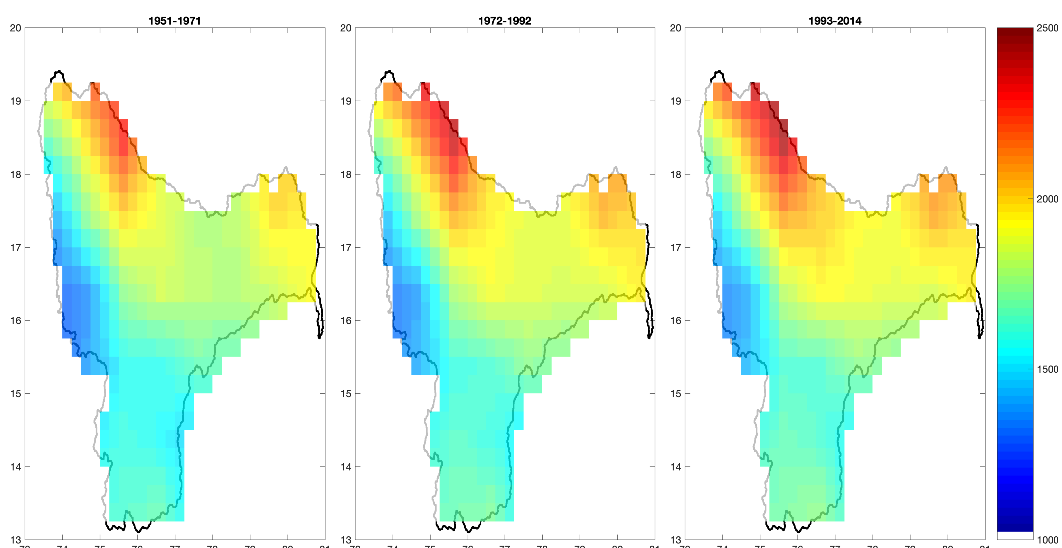


Figure 2.10: Spatial Variation of Potential Evapotranspiration for 1951-1971, 1972-1992 and 1993-2014

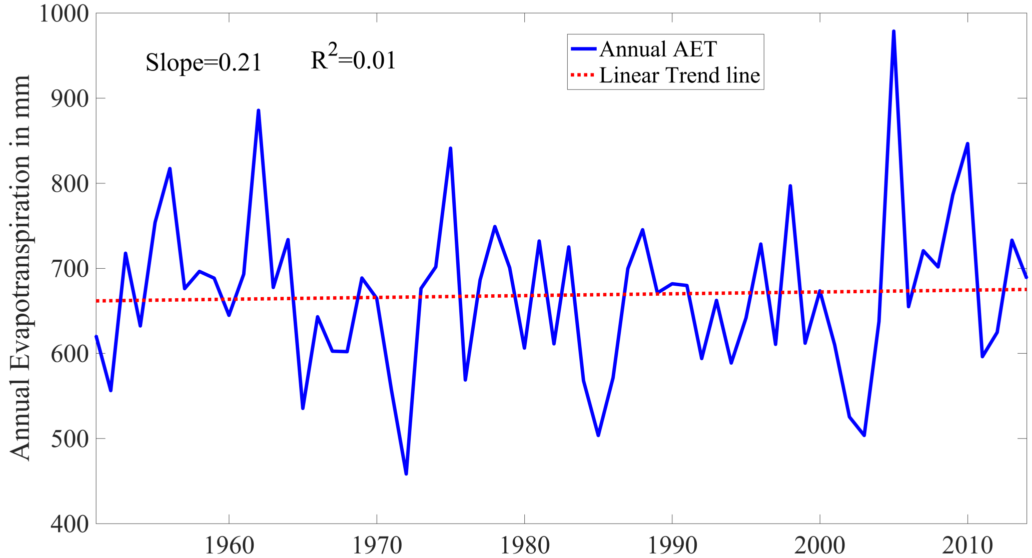


Figure 2.11: Temporal Variation of Actual Evapotranspiration(AET_{Clim}) from 1951-2014

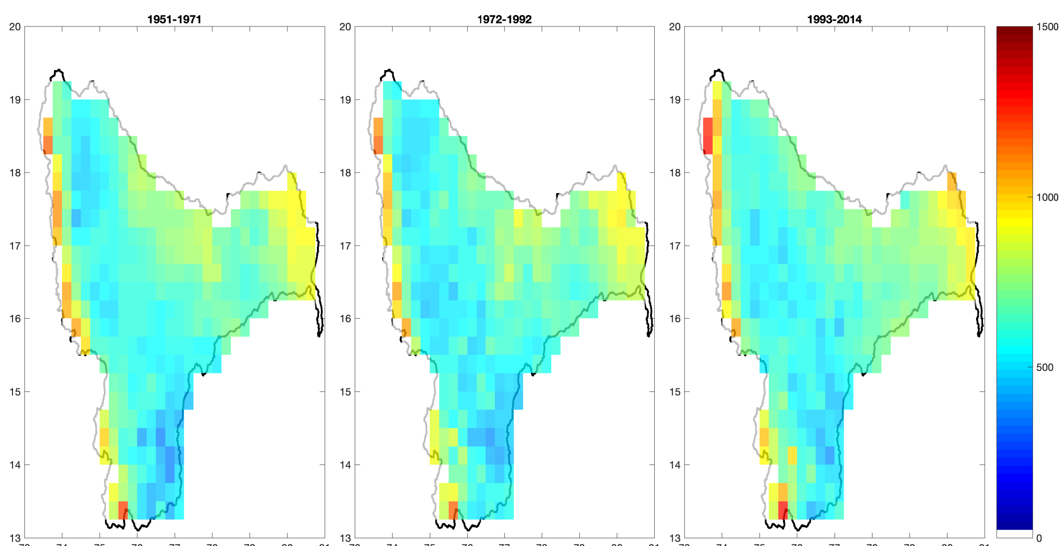


Figure 2.12: Spatial Variation of Actual Evapotranspiration(AET_{Clim}) for 1951-1971, 1972-1992 and 1993-2014

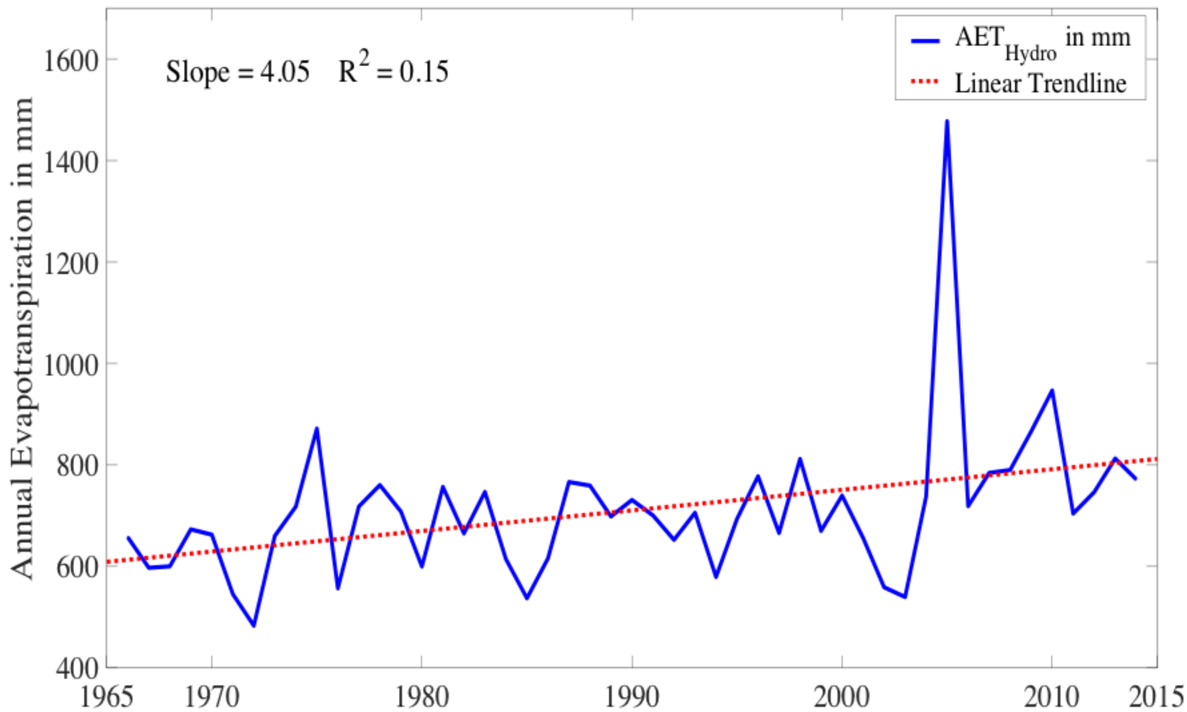


Figure 2.13: Average Actual Evapotranspiration(Hydro) trend from 1966-2014

The Krishna river basin was identified as semi-arid with aridity index ($\frac{P}{PET}$) as 0.44, with annual basin averaged precipitation (778 mm) and PET (1773 mm) estimated from 1951 to 2014. Increasing trends were observed in the basin averaged annual simulated PET by 18.1mm/decade at a significance level of 0.21 (Figure 2.9). Moreover, the spatial variation of average air temperature (Figure 2.5) was found to follow the spatial variation of PET (Figure 2.10) for three time slices.

The AET_{Clim} estimated based on water available (P) and energy available (PET) has shown an increasing trend for the entire basin at a rate of 2.1 mm/decade from 1951 to 2014 (Figure 2.11) is presented for three time periods of 1951-1972, 1973-1992, and 1993-2014 in figure 2.12 whereas the AET_{Hydro} estimated based on water available (P) and energy available (AET) has an increasing trend of 40.5 mm/decade (Figure 2.13) and is presented as the average annual AET before and after the change point(Figure 2.14).

Highest annual PET estimates and positive changes from 1966-2003 to 2004-2014 were observed over the north and north-east parts of KRB (Figure 2.15). The AET_{Clim} has shown positive changes varying from 11 mm to 692 mm with few cells scattered with negative changes from 1965-2003 to 2004-2014. Similarly, the spatial variability of AET_{Hydro} which is hydrologically calibrated AET with surface water and energy balances of P, PET and runoff (R) has also shown positive changes varying from 21.5 mm to 910 mm (Figure 2.15).

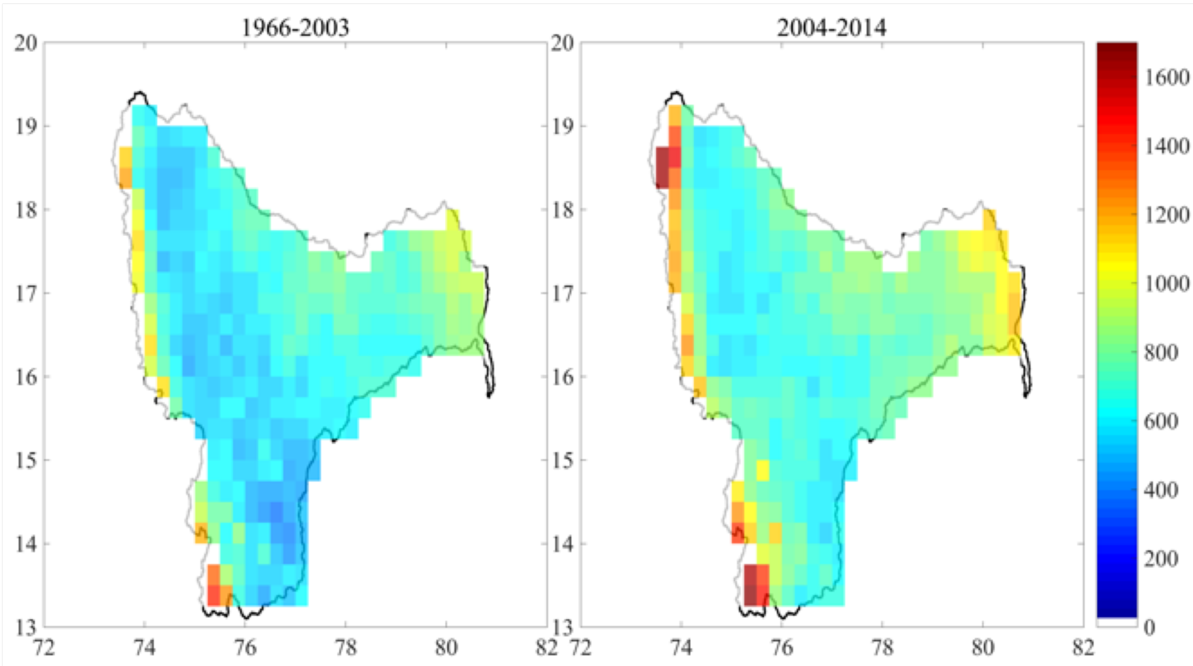


Figure 2.14: Spatial Variation of Actual Evapotranspiration(AET_{Hydro}) for 1966-2003 and 2004-2014

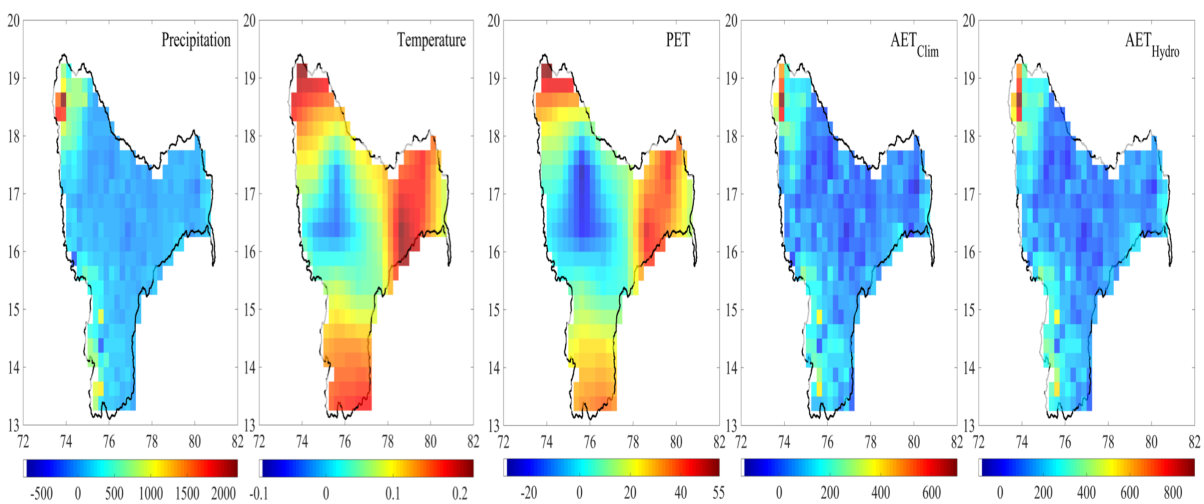


Figure 2.15: Spatial distribution of changes of average conditions for 1965-2003 and 2004-2014.

Both AET_{Clim} and AET_{Hydro} has followed similar spatial and temporal patterns with higher magnitudes for AET_{Hydro} as it accounted for the water-use from the river basin along with water (P) and energy (PET) variability. This can also be observed in terms of spatial variation of modeled precipitation surplus for two time periods of 1966-2003 and 2004-2014 in terms of (a) climatological residual available water (P-AET_{Clim}) and (b) hydrological residual available water (P-AET_{Hydro}) in figure 2.16 and (Figure 2.17 respectively. Overall, the KRB basin has suffered severe decrease in water availability from 1966-2003 to 2004-2014 both in terms of climatological and hydrological (Figures 2.16. 2.17) for the current climate scenario with increase in precipitation, temperature and ET flux over KRB. Particularly, the period 2004-2014 has suffered huge water shortages for the years 2002, 2003, 2009 and 2014 and also reported as drought years over KRB with significant increasing trends of droughts for the period of 1948 to 2012 [81, 61].

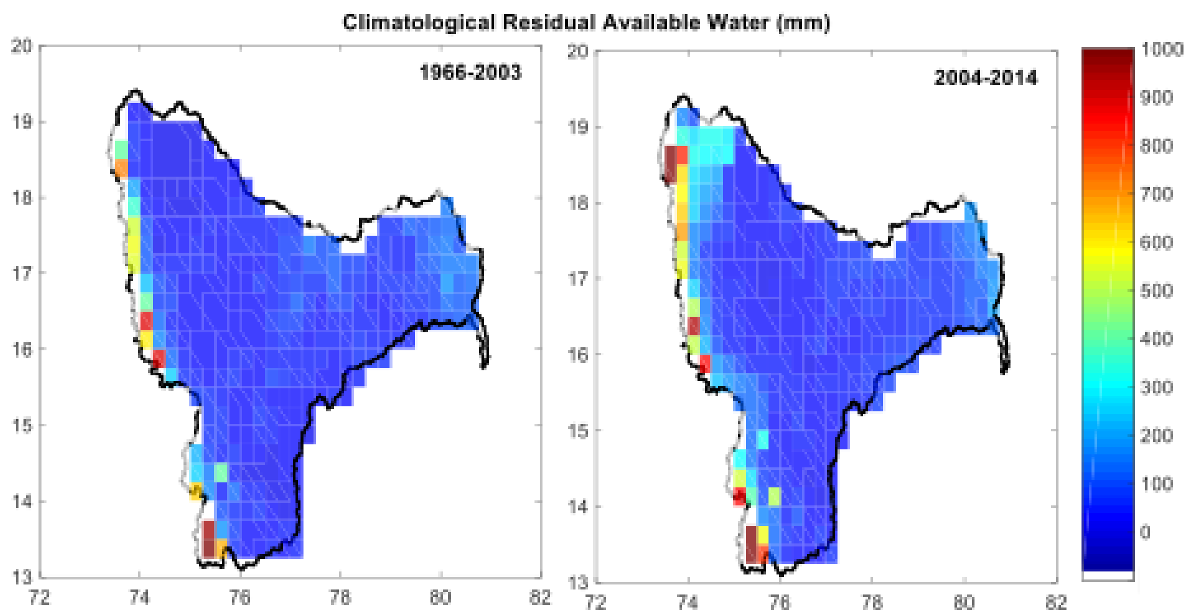


Figure 2.16: Spatial distribution of modeled precipitation surplus in terms of climatological residual available water ($P - AET_{clim}$)

The monthly climatic water demand based on PET (P-PET) and AET (P-AET) (Figure 2.18) reaffirmed the acute unevenness in seasonal precipitation distribution over Krishna River basin. From January to May the amount of available precipitation is less thereby representing water deficiency months and negative water balances. Starting from June, when the monsoon season begins the basin slowly getting to accumulate the storages with positive values of (P-PET) and therefore high amounts of (P-AET) till the end of monsoon season up to October. During post-monsoon and winter months (November, December, January, and February), pre-monsoon (March-May) season, acute water shortfall was observed, due to inadequate precipitation occurrence, therefore, climatic water demand in terms of PET was noticed to be more in these months. It should be noticed that the monthly climatic water balance based on AET is following the monthly variations of precipitation over the Krishna river basin. In general, it

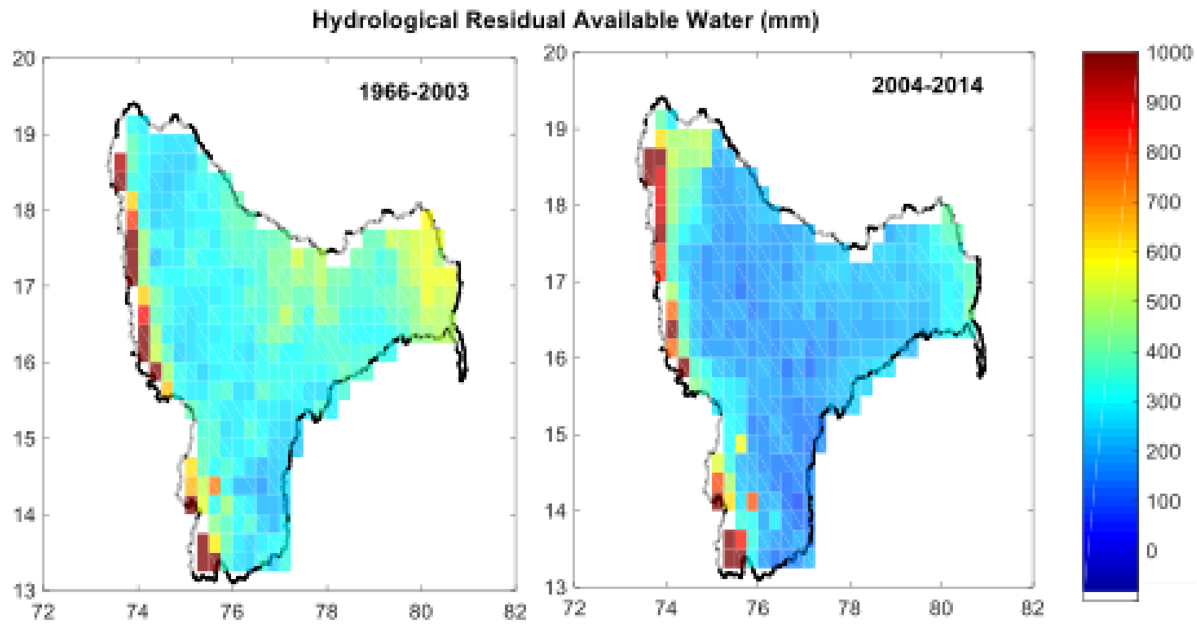


Figure 2.17: Spatial distribution of modeled precipitation surplus in terms of hydrological residual available water ($P - AET_{hydro}$)

was noticed that for the KRB, the spatio-temporal variation of precipitation followed AET more than PET. Further, climatic water balances based on PET has shown negative water balances and AET with positive water balances over the basin.

The AET_{Clim} has shown less water availabilities over the basin compared to AET_{Hydro} as it accounts for atmospheric water supply (P) and energy demand (PET) (Figure 2.16), (Figure 2.17). While, higher precipitation surpluses were observed based on hydrologically calibrated AET due to the consideration of closure of water balance of KRB accounting for water-use in terms of runoff. Further, to validate such results, the observed discharge at the basin outlet was compared with the uncalibrated and hydrologically calibrated discharges resulting water balances of (P-AET) and (P-X_{Cal} AET) respectively for the observed period from 1966 to 2014. The comparison of observed runoff with uncalibrated and calibrated runoff at annual scale resulted in RMSE (R-square) values as 444 (0.4) and 312 (0.6) respectively. The use of hydrologically calibrated AET has significantly improved the runoff prediction and can be considered as more reliable term in the long-term annual water-energy balance studies.

Also, there are various temperature-based approaches [92, 38] to estimate PET. Each empirical model provides different model PET estimates, leading to model uncertainty which may be further pronounced for future scenarios, if model predictions were made based on such single model outputs (Thompson,1977). Therefore, to study the uncertainty in the PET estimation based on various temperature-based models, the present study compared the PET estimated from Thornthwaite method with the Hargreaves model PET estimates [38] as shown in figure 2.19. The basin averaged annual PET estimates of Thornthwaite model (based on average temperature), which has followed the temporal variation of annual

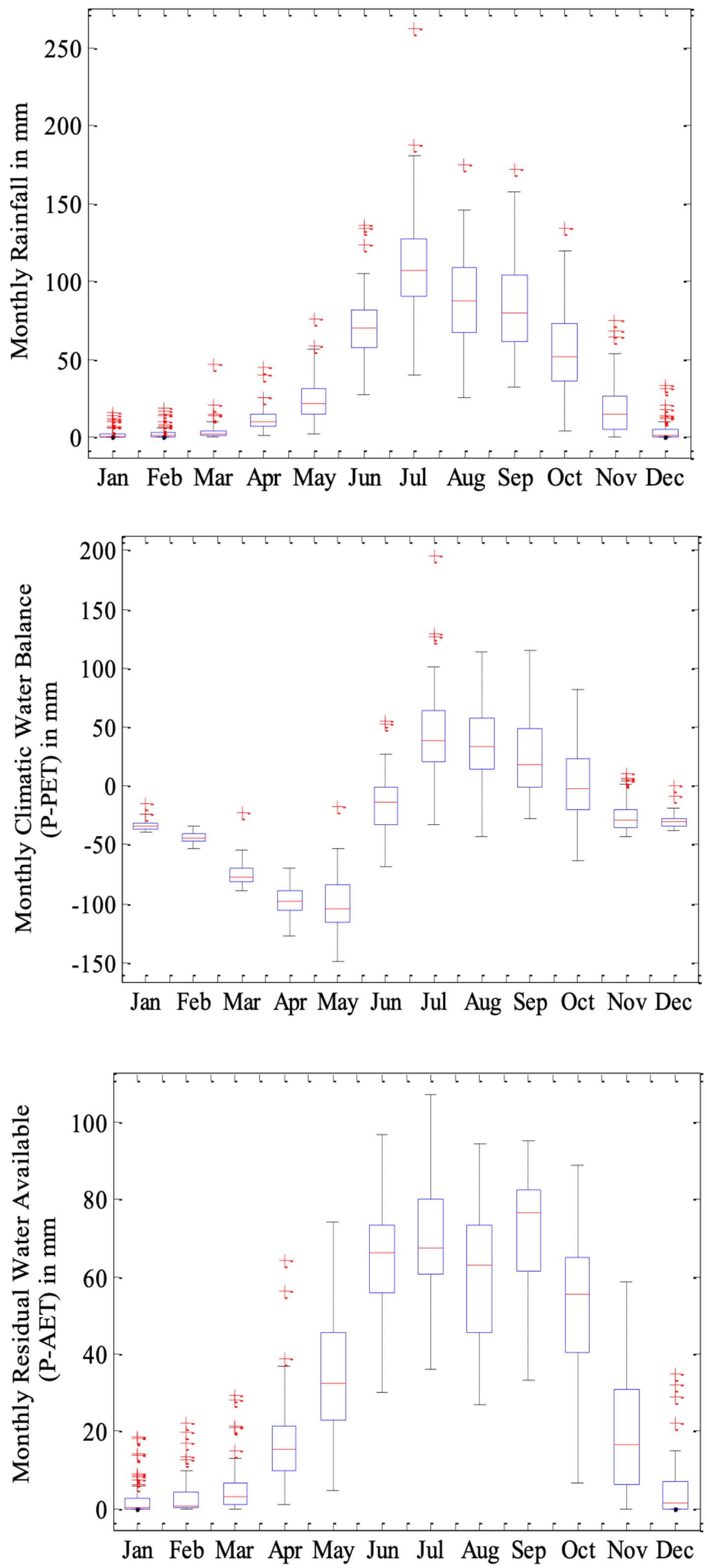


Figure 2.18: Monthly Variation of Precipitation, (P-PET) and (P-AET) over Krishna River Basin

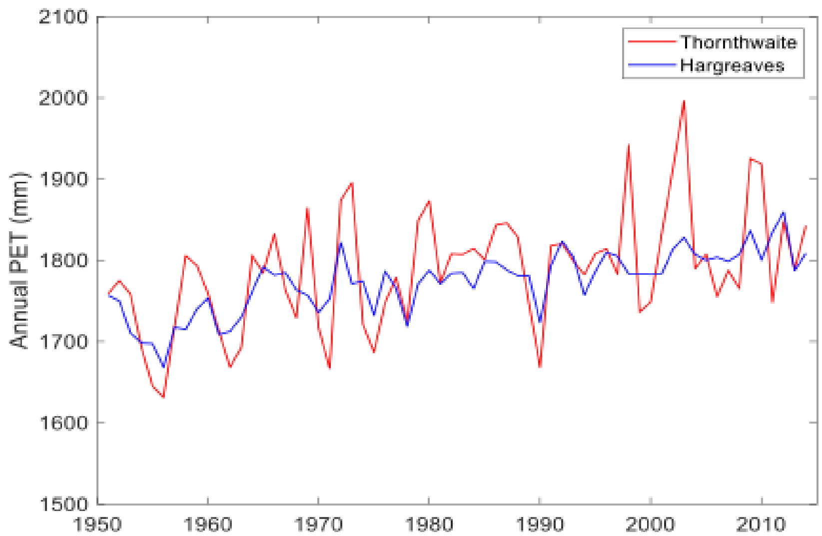


Figure 2.19: Basin averaged annual PET estimates with Thornthwaite and Hargreaves models.

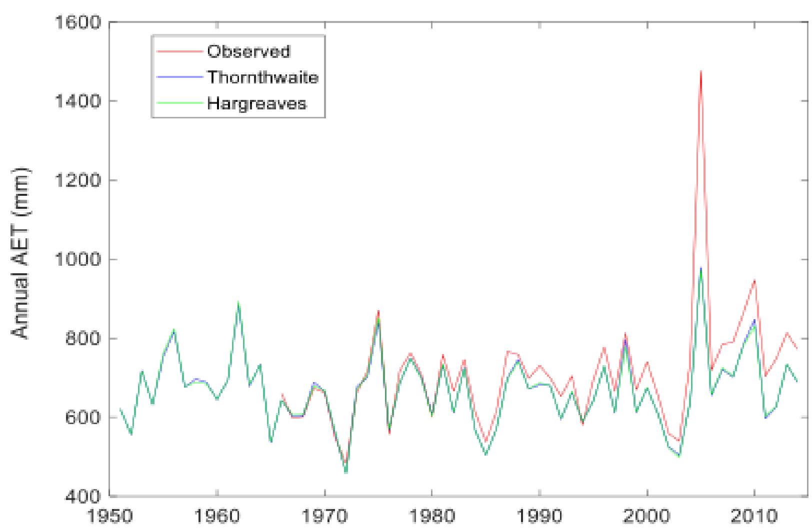


Figure 2.20: Basin averaged annual AET estimates with Thornthwaite and Hargreaves models. The observed annual AET is estimated from ET = P-R at catchment scale.

average temperature, are slightly more than Hargreaves PET estimates (based on maximum and minimum temperatures). Here, to compare the uncertainty level with both PET models, the basin averaged annual PET estimate from 1951 to 2014 based on Thornthwaite and Hargreaves models were compared as 1788 and 1773 mm respectively. Therefore, the uncertainty levels with various PET models can be considered as less significant in the assessment of water-energy variables at river basin scales. Furthermore, the uncertainty levels are also less significant at the level of AET estimates, based on both the PET models of Thornthwaite and Hargreaves estimates with Budyko AET formulation as shown in figure 2.20. Here, the study compared the uncertainty level at AET model estimates with both the PET models of Thornthwaite and Hargreaves models, the basin averaged annual Budyko formulation AET estimates from 1951 to 2014 were compared figure 2.20. The annual average AET from 1951 to 2014 based on Thornthwaite and Hargreaves models were noted as 668.4 and 668.3 mm respectively for KRB. Therefore, the uncertainty levels with various PET models can be considered as less significant in the assessment of AET estimates at river basin scale for KRB.

Further, the study used Budyko hypothesis to estimate evapotranspiration, (AET_{Clim}), which works at annual scale with a reasonable assumption as storage changes can be considered as constant [29]. It should be noted that, if the water balances were estimated at monthly scale, then the assumption of constant storage changes is no longer valid [102]. Further, as Budyko model adopted is parametric (w = 0.5), it is essential to test the applicability of the model for the river basin. Therefore, to validate the applicability of Budyko hypothesis at annual scale, the observed ET, which can be estimated with catchment scale water-balance equation, ET = P-R, can be compared with the ET estimated with parametric equation of Budyko hypothesis. For this, the present study compared the AET (Eq. 2.10) estimated with Budyko hypothesis at annual scale with the ET estimated with water balance equation of ET = P-R, where R is considered at the basin outlet, Vijayawada from 1965 to 2003 as shown in figure 2.20.

The annual basin averaged AET estimates from both Thornthwaite and Hargreaves PET estimates were compared with the observed ET as shown in figure 2.20. It can be noted that up to 2003, the observed, Thornthwaite and Hargreaves model AET estimates were comparable and after this considerable variation between observed and modeled AET estimates were noted. It should be noted that, the year 2003 has been identified as change point detection year with Pettit's test for precipitation figure 2.1. Further, temporal variability of the basin averaged annual AET is mainly depending on precipitation (Figure 2.1, 2.11 and 2.13). Because, the AET model formulation is based on the assumption that AET is limited by the water availability in terms of precipitation under dry conditions and energy availability under wet conditions [9, 107]. If the annual mean P is less than PET, then the zone is defined as water-limited zone; whereas, if the annual mean P is greater than PET, then the region is defined as energy-limited region [24]. Considering the annual mean P and PET over the basin is noted as 778.4 mm and 1773.3 mm respectively, with this the river basin can be identified as water limited. Therefore, the AET is limited by precipitation and thus the basin averaged AET is following the temporal variability of P rather than PET over KRB. Therefore, the modelling approach of estimation of AET with Budyko with estimates of PET proposed in the present study works particularly for arid catchments and not for humid catchments.

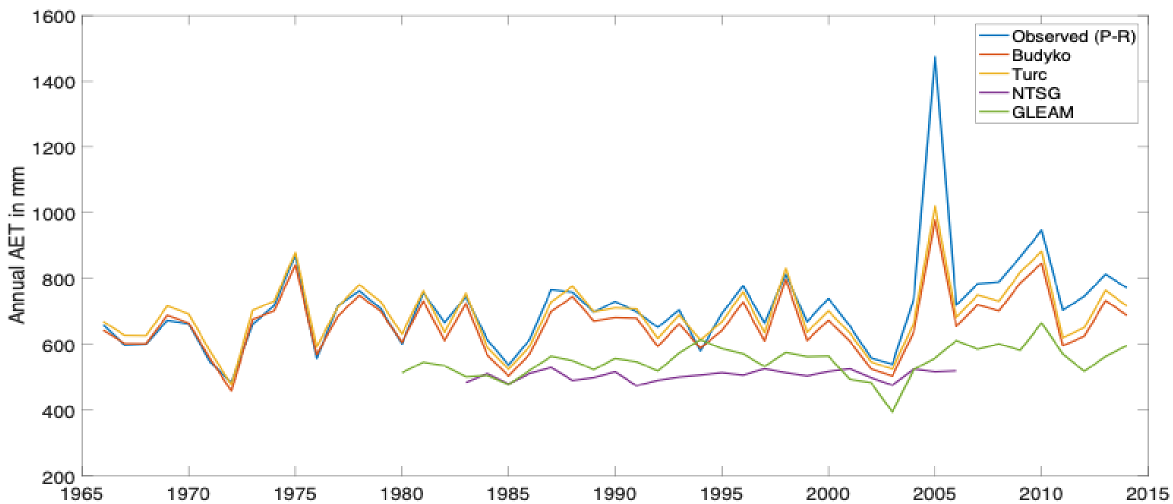


Figure 2.21: Temporal variation of annual AET over KRB from remote sensing-based data, Budyko hypothesis Turc model , hydrologically calibrated and Observed AET$P - R_{Obs}$

Moreover, to validate and to assess the strength of the AET estimates provided in the present study we compared the satellite-based ET data with the PET, AET_{Clim} and AET_{Hydro} estimates over KRB (Figure 2.21). The 2002 year was considered to be all India drought year [59] and the study aimed at to test the strength of the simulated ET at river basin scale for a water shortage year (Figure 2.22).

Further, to study the strength of predictability of modelled ET estimates of the present study, correlation coefficients between remote sensing-based data and ET estimates of river basin were estimated for the period of 1983-2006 (Figure 2.23). All over KRB, the PET estimated from Thornthwaite model has been over predicted (Figure 2.23) with negative correlation coefficients most of the basin except for lower portion of the basin (around 0.3), over few districts of Karnataka, where tropical climate persists.

The AET estimates based on Budyko hypothesis(AET_{Clim}) and hydrologically calibrated(AET_{Hydro}) were observed to be reasonably comparable with remote sensing-based ET data with higher correlation coefficients ranging from 0.3 to 0.7 for most of the KRB. Specifically, AET_{Hydro} estimates were observed to be more comparable with the remote sensing-based ET estimates with higher correlation coefficients for most of the basin compared to uncalibrated AET_{Clim} estimates. Furthermore, AET_{Hydro} estimates have shown better predictability for regions with higher temperatures and low to moderate precipitation i.e regions where arid climate prevails. More than 80% of the basin has shown positive correlations between remote sensing-based ET data and uncalibrated AET_{Clim}, whereas AET_{Hydro} has shown 85% of the basin with positive correlations. Also, about 50% (42%) and 25% (21%) of the basin has shown positive correlations of 0.3 and 0.4 between remote sensing-based ET data and AET_{Hydro} (AET_{Clim}) respectively. Further, it should be noted that the correlation measures between

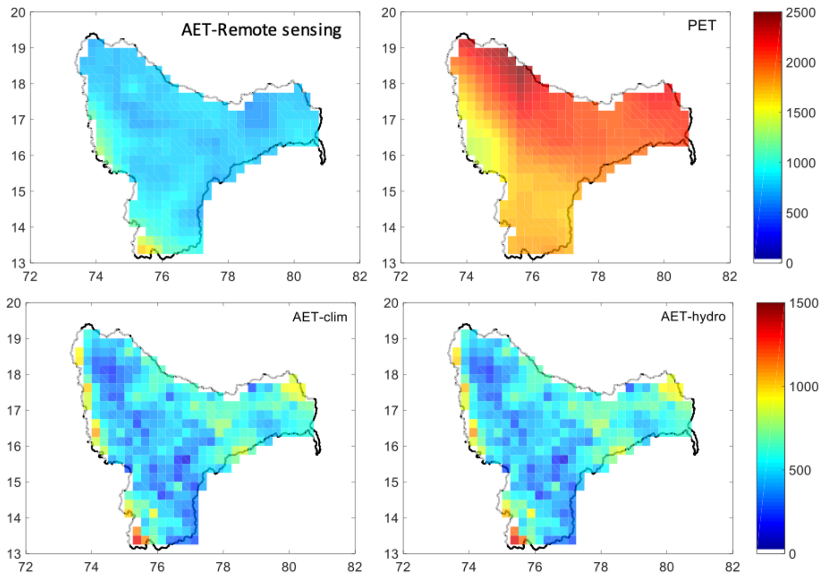


Figure 2.22: Spatial distribution of annual ET over KRB from remote sensing-based data, Thornthwaite model (PET), Budyko hypothesis AET_{clim}, hydrologically calibrated AET_{hydro} for the year 2002

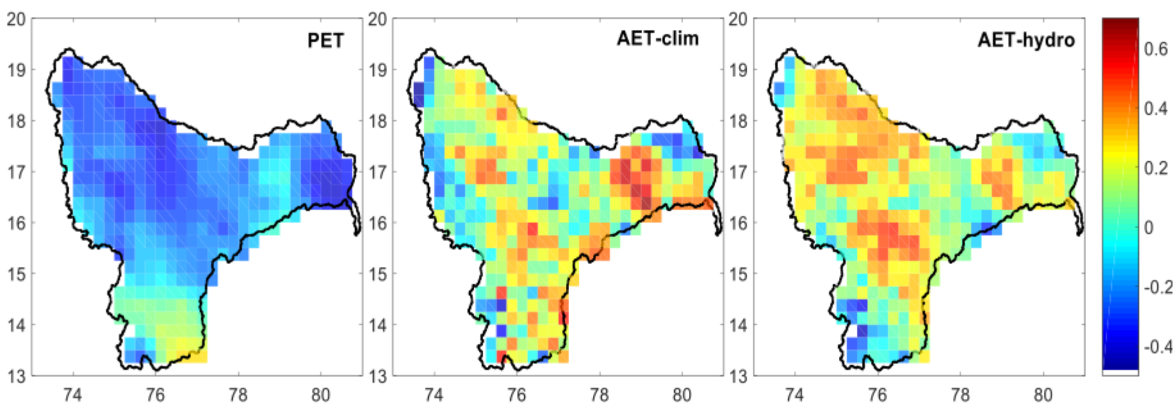


Figure 2.23: Correlation coefficients between Remote-sensing-based data and Thornthwaite model (PET), Budyko hypothesis AET_{clim}, hydrologically calibrated AET_{hydro} for the period of 1983-2006 over KRB

remote sensing-based ET data and modeled hydrological induced AET depends on the spatial variability of land-cover type and temporal scales [105].

Conclusion

An ensemble regression model to hydrologically calibrate ET (AET_{Hydro}) accounting for water-use (R) along with water supply (P) and energy (PET), has been developed. The study compared how this hydrologically calibrated ET estimate accounting for various major long-term water-energy balance variables is appropriate for predicting current water availabilities at regional (or catchment) scale by comparing with satellite-based land surface ET estimates. The proposed modelling framework of AET estimation can be implemented with any region specific empirical AET models to study the water balances at catchment scale. The climatological and hydrological modelled AET data will be used with Standardized Precipitation Evapotranspiration Index (SPEI) to develop a new drought index of Standardized Precipitation Actual Evapotranspiration Index (SPAEI) which can represent the climate and hydrological variability in the meteorological drought analysis as described in Chapter 3.

Chapter 3

A Comparative Analysis of Hydrological and Meteorological Drought Indices over Krishna river basin

Introduction

Potential Evapotranspiration (PET) based drought indices such as Standardized Precipitation Evapotranspiration index (SPEI) have gained much attention in the recent times given the ability to capture the anomalies of both precipitation and PET in the drought estimation. Although PET-based drought indices consider the climatic water demand, it is limited towards the inclusion of the effects of regional land surface changes and actual moisture availability in the drought estimation. Given that Actual Evapotranspiration (AET) represents the transfer of moisture from the surface to the atmosphere in response to both the energy demand and available moisture supply, the drought indices estimated based on AET will consider both climatic water demand and actual available moisture. Here, the study proposed a water balance-based drought indicator, Standardized Precipitation Actual Evapotranspiration Index (SPAEI), which uses hydrologically calibrated AET estimated with empirical formulations . The study compared the performance of the proposed hydrometeorological drought index with existing meteorological and hydrological drought indicators of Standardized Precipitation Evapotranspiration Index (SPEI), Standardized Precipitation Index (SPI) and Standardized Runoff Index (SRI) as well as with drought index calculated with remote sensing based AET estimates. The study investigated major droughts over the Krishna river basin, for which most of the basin is in arid climate using the hydrologically calibrated AET estimates as explained in Chapter 2. As SPAEI is more reasonable in reflecting the surface water-energy balance it enables better characterization of meteorological and hydrological droughts at regional scales. The different drought indices used along with the detailed methodology and formulations are described in Section 3.1 followed by the comparison of drought indices at variable time scales and their findings in Section 3.2.

3.1 Modelling of Meteorological and Hydrological Drought Indices

The different drought indices Standardized Precipitation Evapotranspiration Index(SPEI), Standardized Precipitation Actual Evapotranspiration Index(SPAEI), Standardized Precipitation Index(SPI) and Standardized Runoff Index(SRI) are explained in the following sections along with the drought categorization table.

3.1.1 Standardized Precipitation-Evapotranspiration Index(SPEI)

Following to the methodology proposed by [99], the SPEI is based on fitting a three-parameter log-logistic probability distribution for the accumulated climatic water balance(D), the difference between the accumulated precipitation(P) and Potential Evapotranspiration(PET) $D = P - PET$.The estimated D values represents the water demand or surplus(P-PET), while the evapotranspiration is the result of complex relationship between atmosphere and surface water available, vegetation and soil characteristics [8]).

The probability density function (pdf) (f(x)) and cumulative distribution function (CDF) (F(x)) of the three-parameter log-logistic distribution are given as follows:

$$f(x) = \frac{\beta}{\alpha}\left(\frac{x-\gamma}{\alpha}\right)^{\beta-1}\left[1+\left(\frac{x-\gamma}{\alpha}\right)^{\beta}\right]^{-2} \tag{3.1}$$

where α, β and γ are the scale,shape and origin parameters, respectively, for D and RAW values in the range of ($\gamma > D, RAW > \infty$). The parameters of the log-logistic distribution are obtained by following the L-moment procedure as follows:

$$\beta = \frac{2w_1 - w_0}{6w_1 - w_0 - 6w_2} \tag{3.2}$$

$$\alpha = \frac{(w_0 - 2w_1)\,\beta}{\Gamma\left(1+\frac{1}{\beta}\right)\Gamma\left(1-\frac{1}{\beta}\right)} \tag{3.3}$$

$$\gamma = w_0 - \alpha\Gamma\left((1+\frac{1}{\beta}\right)\left(1-\frac{1}{\beta}\right) \tag{3.4}$$

where Γ is the gamma function of β, and w_0, w_1 and w_2 are the probability weighted moments calculated based on Sheng and Hashino (2007), as in the Equation 3.5:

$$w_s = \frac{1}{n}\binom{n-1}{r}^{-1}\sum_{j=1}^{n-r}\binom{n-j}{r}x_j \tag{3.5}$$

where r=0,1,2 and n is the sample size x_j and is the ordered vector of observations in descending order. Next, the cumulative distribution function F(x) of log-logistic distribution can be calculated with the estimated parameters of Pearson-III distribution.

$$F(x) = \left[1+\left(\frac{x-\gamma}{\alpha}\right)^{-\beta}\right]^{-1} \tag{3.6}$$

With the values of F(x), the SPEI values were calculated in the Equation 3.7:

$$SPEI = W - \frac{C_0 + C_1W^1 + C_2W^2}{1 + d_1W^1 + d_2W^2 + d_3W^3} \tag{3.7}$$

where $W = -2ln(P)$ for $P \leq 0.5$ where P is the probability of exceeding a determined D value, P = 1-F(x). If $P > 0.5$, P is replaced by 1-P and the sign of the resultant SPEI is reversed.

The constants are:

$$C_0 = 2.515517, C_1 = 0.802853, C_2 = 0.010328$$
$$d_1 = 1.432788, d_2 = 0.189269, d_3 = 0.001308.$$

By substituting the C_0, C_1 and C_2 values in Eq. 3.7, the SPEI values at various time scales can be found. Generally, SPEI can be expressed at different time scales as SPEI (6), SPEI (12), etc., where the number in the bracket indicates the timescale in months for which the P-PET values are accumulated and the estimated SPEI at these timescales.Table 3.1 gives the range of SPEI values to identify the extreme weather as drought or wet conditions.

Moisture Category	**SPEI or SPAEI Value**
Extremely wet (EW)	2.00 and above
Very wet (VW)	1.50 to 1.99
Moderately wet (MW)	1.00 to 1.49
Near Normal (NN)	-0.99 to 0.99
Moderately dry (MD)	-1.00 to -1.49
Severely dry (SD)	-1.50 to -1.99
Extremely dry (ED)	-2.00 and less

Table 3.1: Criteria for Identification of Drought and Wet Conditions

3.1.2 Standardized Precipitation Actual Evapotranspiration Index(SPAEI)

SPAEI is estimated similar to SPEI however the three parameter log-logistic probability distribution to the accumulated residual water balance(RAW), the difference between the accumulated Precipitation(P) and Actual Evapotranspiration(AET)

$$RAW = P - AET$$

Depending on the AET chosen we name the index as $SPAEI_{Budyko}$ for AET_{Buydko}, $SPAEI_{Turc}$ for AET_{Turc} and $SPAEI_{Hydro}$ for AET_{Hydro}. Drought categorization for SPAEI is done similar to SPEI categorization as mentioned in the Table 3.1. To assess the credibility of joint effect of meteorological inputs, Precipitation and hydrological inputs, Runoff on drought, we also estimate drought indices Standardized Precipitation Index(SPI) solely dependent on Precipitation and Standardized Runoff Index(SRI) solely dependent on Runoff and compare it with the SPEI and SPAEI.

3.1.3 Standardized Precipitation Index

SPI is based on fitting a probability distribution to the accumulated precipitation(3.8) similar to SPEI.

3.1.4 Standardized Runoff Index

To characterize hydrological drought (Shukla and Wood) developed SRI by considering monthly stream flow data(Q_i) and fitting a probability distribution to the accumulated streamflow data(Q_i^k) similar to SPEI.

$$Q_i^k = \sum_{i-k+1}^{i} Q_i \tag{3.8}$$

The categorization criteria for the SPI and SRI is similar to SPEI as described in the Table 3.1. These drought indices are estimated at various time scales of 6-, 12-, 18, and 24-months and the drought severity, frequency and duration are calculated for which the results are described in Section 3.2.

3.2 Results and Discussions – Analysis of Meteorological and Hydrological Drought

The monthly D (P-PET) and RAW (P-AET) values described in 2.3.2 were used in the estimation of drought indices of SPEI and SPAEI respectively at various time scales of 6-, 12-, 18, and 24-months for the period of 1951 to 2014. The three-parameter log-logistic distribution was applied to model the time series of (P-PET) and (P-AET) i.e for $AET_c lim$ as well as $AET_H ydro$ for various time scales.

Furthermore, the fitted three parameter log-logistic distribution is validated with the Kolmogorov-Smirnov (K-S) goodness of fit test for both the climatic water balance time series of D and RAW. A rejection frequency was defined as the ratio of number of grid points which did not fit the time series of D and RAW for log-logistic distribution, to the total number of grid points in the river basin at a given significance level. The K-S rejection frequencies for the overall basin including all valid grid points were obtained as 6%, 8% and 7.7% for SPEI, $SPAEI_{Budyko}$ and $SPAEI_{Turc}$ respectively at a significance level of 0.01.

To assess SPEI and SPAEI, for meteorological drought detection, the years when the annual precipitation is less than 75 % of the annual average estimated over a period of 1951 to 2014 were considered based on the IMD definition of drought year [1]. Four drought years were identified based on the deviation of annual precipitation from the normal precipitation of the period 1951-2014. Based on figure 3.1(a), these drought years have been identified as 1972, 1985, 2002 and 2003. These drought years are among the major documented drought events over Indian monsoon region [20, 52]. Among these, the year 2002 was one of the most severe drought year in India, which has affected 56% of its geographical

[1] http://imd.gov.in/section/nhac/wxfaq.pdf

Year	Drought Type	SPEI	$SPAEI_{Budyko}$	$SPAEI_{Turc}$	$SPAEI_{Hydro}$	$SPAEI_{RS-ET}$
1972	Moderate	16.00	34.40	44.27	36.78	*
	Severe	33.3	33.3	19.47	29.6	*
	Extreme	28.53	6.67	2.40	0.86	*
1985	Moderate	28.80	39.73	45.87	53.10	27.01
	Severe	15.73	14.13	7.20	12.93	13.22
	Extreme	4.53	3.73	2.40	2.29	3.16
2002	Moderate	35.20	42.93	52.80	70.40	27.01
	Severe	35.47	13.87	11.73	15.51	10.92
	Extreme	11.47	2.40	1.87	2.01	2.30
2003	Moderate	13.60	31.20	36.53	53.73	25
	Severe	16.27	28.80	21.60	12.64	17.82
	Extreme	49.07	7.20	5.07	3.16	7.18

* - Data unavailable

Table 3.2: Percentage of Drought area in Krishna basin based on SPEI, $SPAEI_{Budyko}$, $SPAEI_{Turc}$, $SPAEI_{Hydro}$ and $SPAEI_{RS-ET}$

area, livelihoods of 300 million people [2]. Therefore, various drought indices of SPEI and SPAEI based on empirical models were studied for the drought affected years of 1972, 1985, 2002 and 2003 over the Krishna river basin in terms of areal extent, severity, frequency and duration.

Furthermore, as the satellite-based ET data used in the estimation of drought indices considered the surface energy balances, ET from vegetated areas, evaporation from water bodies, biome-specific NDVI-derived canopy conductance in the ET estimation [106], the study considered the drought indices estimated based on such data as a base for the comparison of empirical AET based drought indices [82]. The present study compared various drought characteristics estimated based on empirical AET models with Remote Sensing-ET (RS-ET) data for the drought affected years of the basin.

To study the spatial drought characterization, the areal extent of droughts represented as percentage of grids for moderate (-1 $\leq SPEI/SPAEI_{Budyko}/SPAEI_{Turc}/SPAEI_{Hydro}/SPAEI_{RS-ET} \leq$ -1.49), severe (-1.5 $\leq SPEI/SPAEI_{Budyko}/SPAEI_{Turc}/SPAEI_{Hydro}/SPAEI_{RS-ET} \leq$ -1.99) and extreme ($SPEI/SPAEI_{Budyko}/SPAEI_{Turc}/SPAEI_{Hydro}/SPAEI_{RS-ET} \leq$ -2) were studied out of total number of 348 grid points over the basin at 12-month time window (Fig 3.1, Table 3.2). The percentages of annual moderate and severe drought affected areas were observed to be increasing over the Krishna river basin for the period of 1951-2014. The areal extents of moderate, severe and extreme droughts were observed to be more for the drought years of 1972, 1985, 2002 and 2003 for all these drought indices. The remote sensing-based drought index, $SPAEI_{RS-ET}$, also has observed the years 1985, 2002 and 2003 as highly affected drought years in terms of higher percentage of areal extends for moderate, severe and extreme categories of droughts for the period of 1983 to 2006. The percentage of drought areal extents with SPEI, $SPAEI_{Budyko}$, $SPAEI_{Turc}$ for various categories of droughts were studied for the major drought years over the basin in the Table 3.2. The moderate

[2] https://public.wmo.int/en/bulletin/flood-and-drought-management-throughwater-resources-development-india

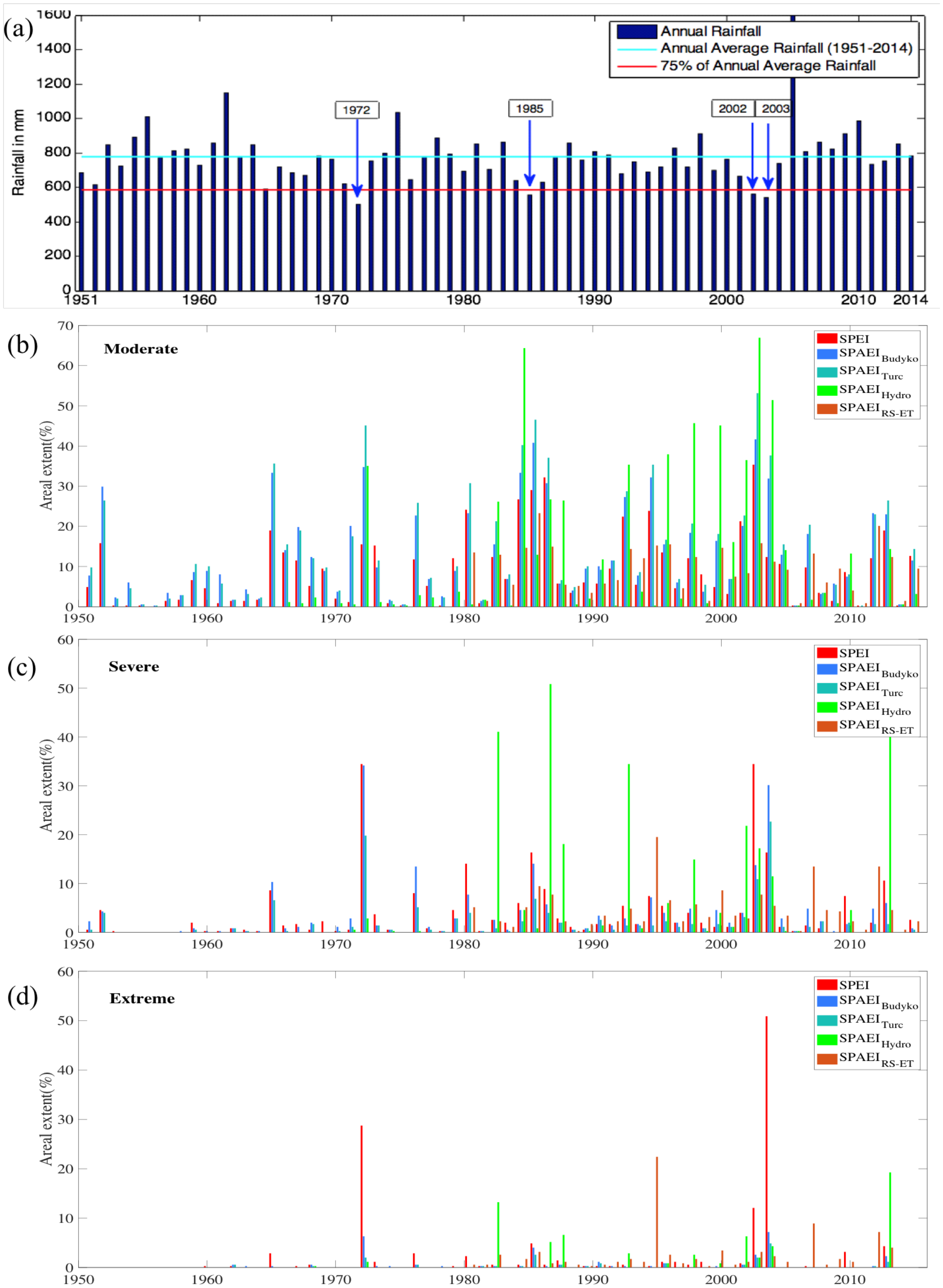


Figure 3.1: Annual precipitation of Krishna river basin compared to long term average annual precipitation (a), Areal extent of moderate (b), severe (c) and extreme (d) droughts represented as percentage of grids with SPEI and SPAEI <-1, <-1.5 and <-2 respectively at 12-month time window

drought areal extents with $SPAEI_{Budyko}$ and $SPAEI_{Turc}$ were observed to be more compared to SPEI for Krishna river basin (Figure 3.1(b), 3.2). However, the percentage of the severe and extreme drought areas were noted to be higher with SPEI compared to both $SPAEI_{Budyko}$ and $SPAEI_{Turc}$ (Figure 3.1(c), 3.2). As SPEI consider the residual water available for evaporation based on energy available, in terms of PET, higher severe and extreme drought areal extents were estimated. While, AET based drought indices accounts for the residual available water for evaporation based on both energy and water, moderate drought percentage areal extents were noted compared to SPEI. For example, for the recent consecutive drought years of 2002 and 2003, about 11.47%, 2.4%, 1.87%, 2.01%, 2.3% and 49.07%, 7.2%, 5.07%, 3.16%, 7.2% of area has been identified under severe drought with SPEI, $SPAEI_{Budyko}$, $SPAEI_{Turc}$, $SPAEI_{Hydro}$, $SPAEI_{RS-ET}$ respectively as shown in 3.2. Furthermore, $SPAEIBudyko$ and $SPAEIHydro$ indices were identified as a more reliable measure in the estimation of meteorological and hydrological drought areal extent by comparing with satellite-based land surface ET data ($SPAEI_{RS-ET}$) . While, the Turc model has under predicted the extreme drought areal extents compared to $SPAEI_{Budyko}$ and $SPAEI_{RS-ET}$.

Furthermore, to support such research findings, the spatial drought characterizations for the major drought years of 1972, 1985, 2002 and 2003 were presented in the figures 3.2, 3.3, 3.4, 3.5. The years 1972 and 2003 were noted as most severe historic droughts occurred over Krishna river basin as most of the basin was classified under extreme drought. For the 1972 drought year, the upper portion of the basin, particularly, Maharashtra (e.g. Ahmed nagar, Pune, Solapur, Satara, Sangi, Bijapur), North Karnataka (e.g. Gulbarga, Raichur), Telangana (e.g. Rangareddy, Mahabubnagar, Nalgonda, NagarKarnool) were classified under extreme drought regions with SPEI, whereas the AET based drought indices drive few of those regions into moderate (Figures 3.2 and Figure 1.3). Such noticeable deviation in the drought categorization from extreme/severe to moderate can also be seen for the drought years of 2002 and 2003 in the figures 3.4 and 3.5 respectively . Therefore, the driving of areal extents between SPEI and SPAEI (both Budyko and Turc) for severe and extreme droughts is more evident, than the moderate drought areal extents. For drought years of 1985, 2002 and 2003, the severe and extreme drought areal extents with $SPAEI_{Budyko}$ and $SPAEI_{Turc}$were found to be the more comparable with $SPAEI_{RS-ET}$ than with SPEI (Table 3.2). Thus, PET based drought indices categorize more percentage of area as severe or extreme, which were identified as moderate drought areas otherwise with AET based drought indices (Table 3.2).

The PET and AET based drought indices were compared in terms of drought intensity for 6-, 12-, 18-, and 24- months accumulation periods correspondingly (Figure 3.6). The SPEI,$SPAEI_{Budyko}$, $SPAEI_{Turc}$ and $SPAEI_{RS-ET}$ values accumulated at 6 months represent the monsoon seasonal precipitation variations over the basin. The SPEI is able to reconstruct most of the drought years at 6 months scale as moderate and severe, whereas, both $SPAEI_{Budyko}$ and $SPAEI_{Turc}$ has recognized them as mild drought years.

The drought severity for major droughts for various accumulated time periods were compared for the four drought indices 3.3. Both SPAEI formulations of empirical and remote sensing-based drought

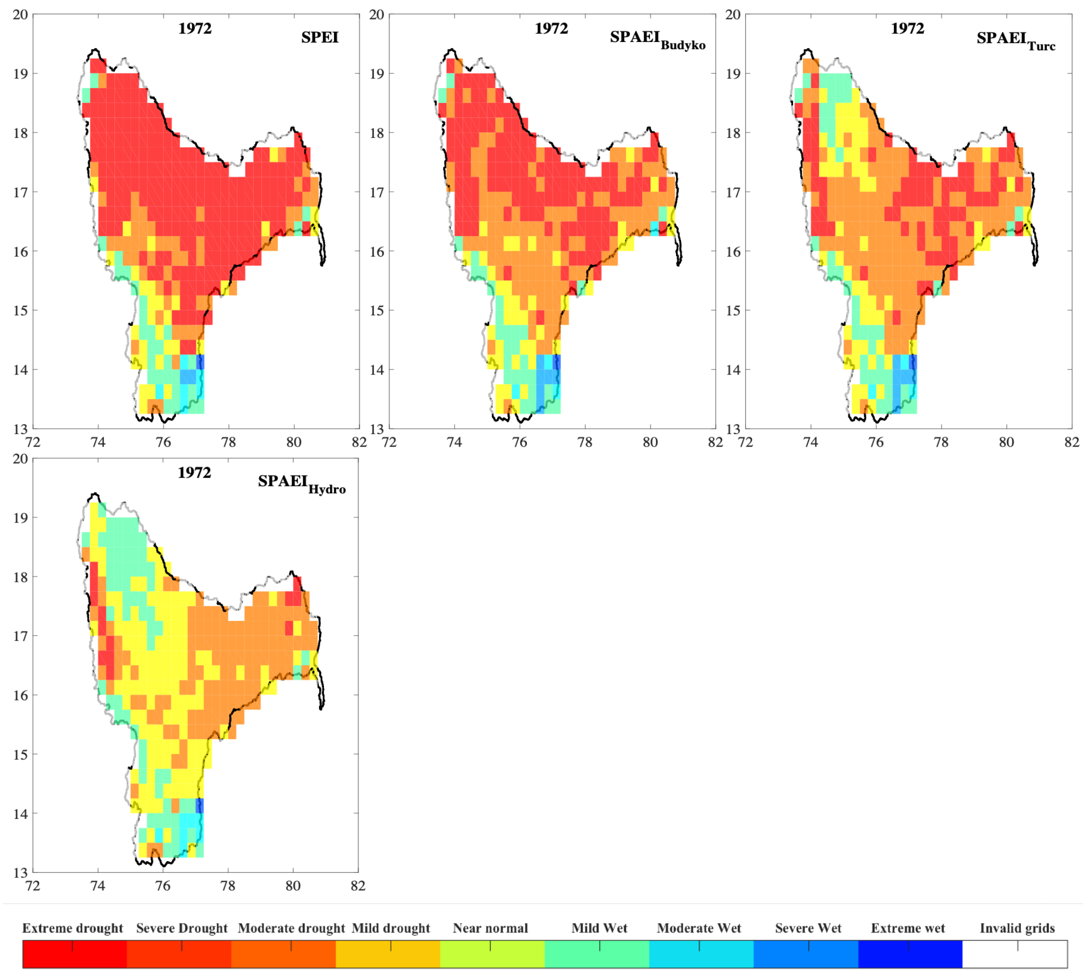


Figure 3.2: Spatial drought categorizations based on SPEI, $SPAEI_{Budyko}$, $SPAEI_{Turc}$, $SPAEI_{Hydro}$ and $SPAEI_{RS-ET}$ at 12-month scale over Krishna River basin for drought years of 1972

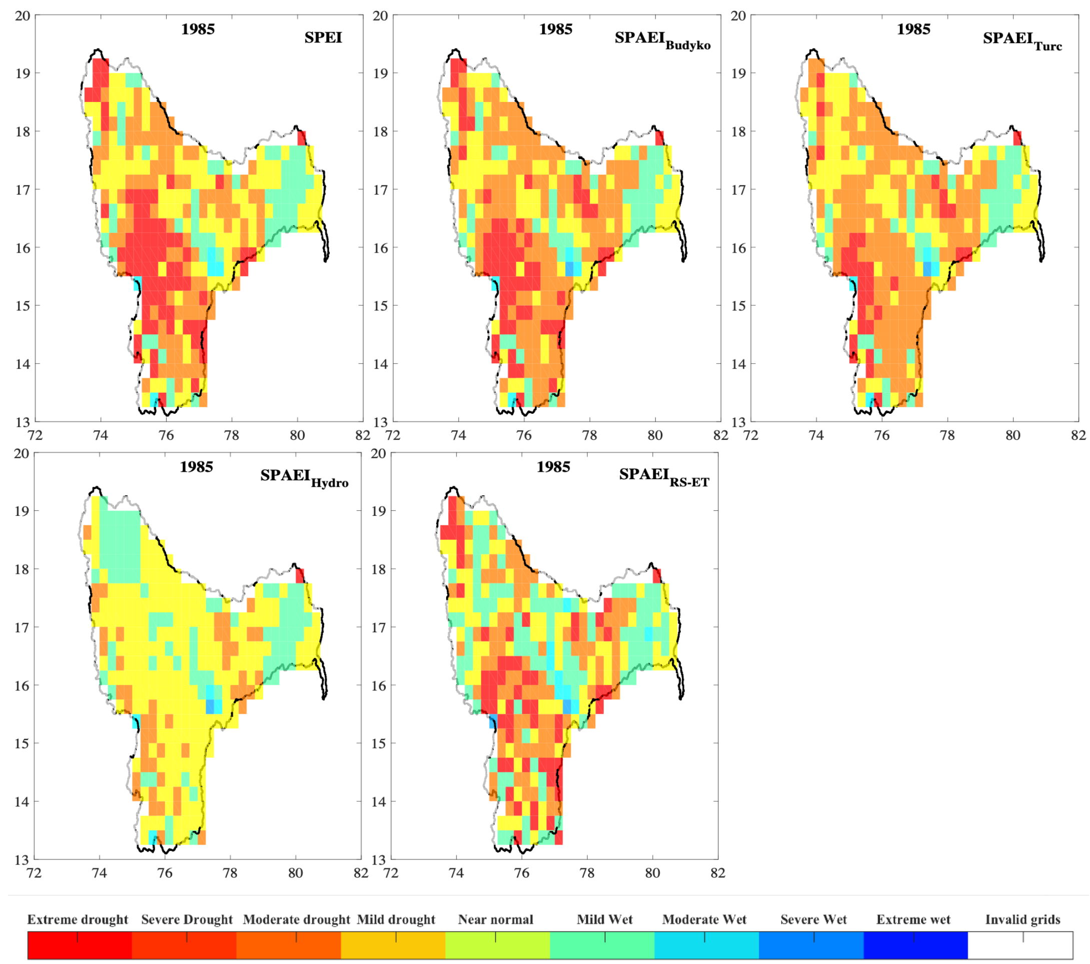


Figure 3.3: Spatial drought categorizations based on SPEI, $SPAEI_{Budyko}$, $SPAEI_{Turc}$, $SPAEI_{Hydro}$ and $SPAEI_{RS-ET}$ at 12-month scale over Krishna River basin for drought years of 1985

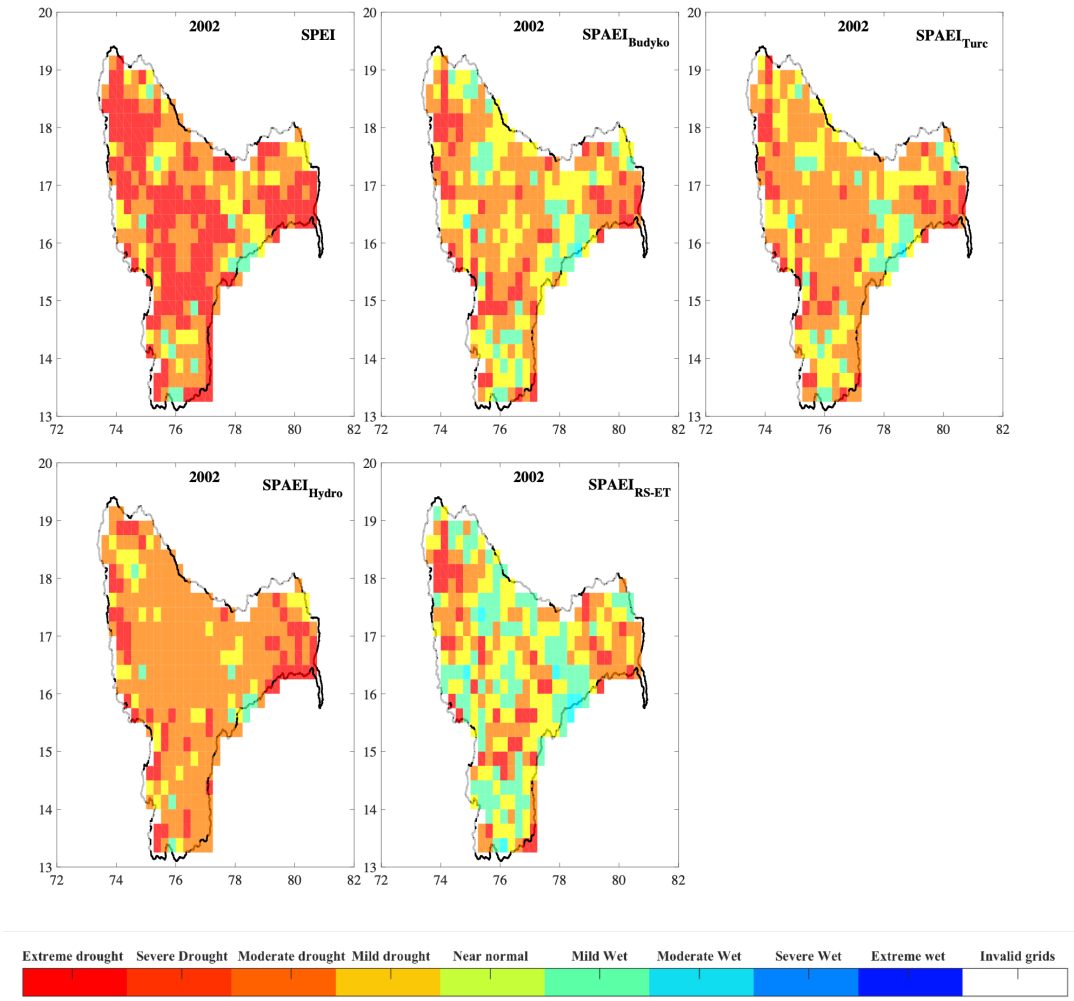


Figure 3.4: Spatial drought categorizations based on SPEI, $SPAEI_{Budyko}$, $SPAEI_{Turc}$, $SPAEI_{Hydro}$ and $SPAEI_{RS-ET}$ at 12-month scale over Krishna River basin for drought years of 2002

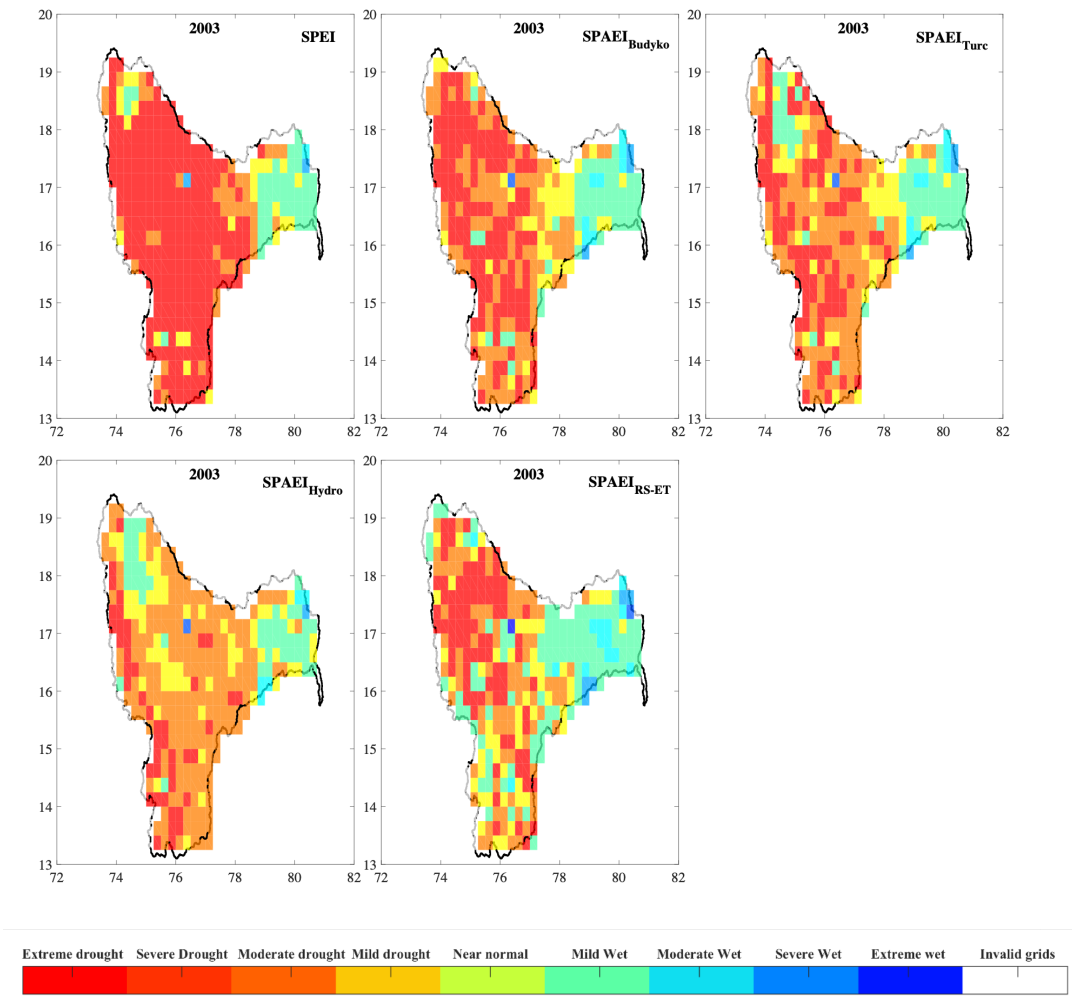


Figure 3.5: Spatial drought categorizations based on SPEI, $SPAEI_{Budyko}$, $SPAEI_{Turc}$, $SPAEI_{Hydro}$ and $SPAEI_{RS-ET}$ at 12-month scale over Krishna River basin for drought years of 2003

	6-Month scale			
Year	**SPEI**	$SPAEI_{Budyko}$	$SPAEI_{Turc}$	$SPAEI_{RS-ET}$
1972	-1.42	-0.69	-0.77	*
1985	-1.40	-0.98	-0.74	-0.30
2002	-1.57	-0.85	-1.01	-0.42
2003	-2.00	-0.97	-0.81	-0.23
	12-Month scale			
Year	**SPEI**	$SPAEI_{Budyko}$	$SPAEI_{Turc}$	$SPAEI_{RS-ET}$
1972	-1.51	-1.24	-1.05	*
1985	-0.98	-1.03	-0.98	-0.86
2002	-1.43	-1.06	-1.10	-0.83
2003	-1.80	-1.15	-1.06	-0.88
	18-Month scale			
Year	**SPEI**	$SPAEI_{Budyko}$	$SPAEI_{Turc}$	$SPAEI_{RS-ET}$
1972	-2.07	-1.33	-1.20	*
1985	-1.46	-1.21	-1.19	-0.39
2002	-2.21	-1.40	-1.38	-0.18
2003	-1.76	-1.15	-1.15	-0.89
	24-Month scale			
Year	**SPEI**	$SPAEI_{Budyko}$	$SPAEI_{Turc}$	$SPAEI_{RS-ET}$
1972	-1.77	-1.40	-1.35	*
1985	-1.08	-1.14	-1.11	-0.89
2002	-1.26	-0.99	-1.03	-0.77
2003	-1.99	-1.42	-1.40	-1.09

* - Data unavailable

Table 3.3: Drought intensity for major drought years of 1972, 1985, 2002 and 2003 drought years as estimated by SPEI, $SPAEI_{Budyko}$, $SPAEI_{Turc}$ and $SPAEI_{RS-ET}$ over Krishna river basin

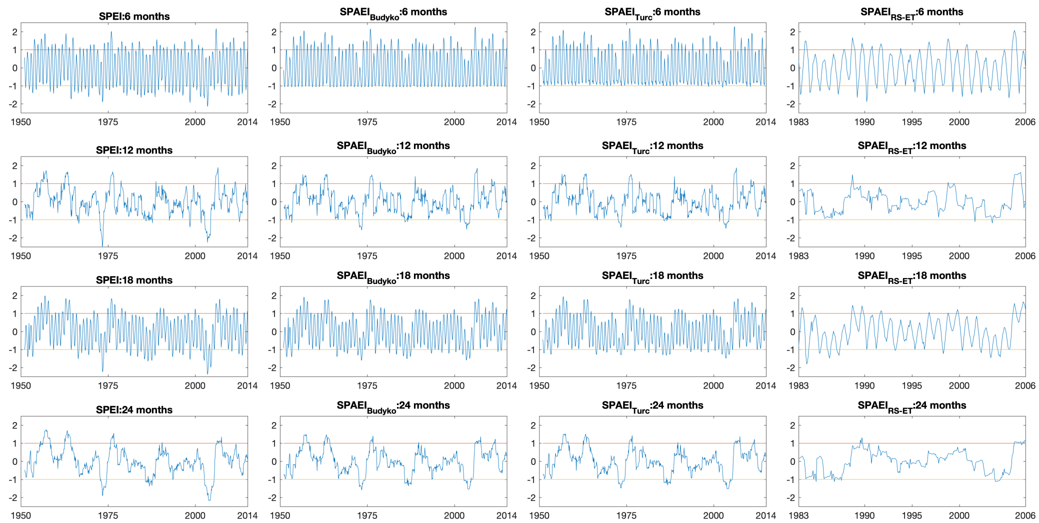


Figure 3.6: Time series of SPEI, $SPAEI_{Budyko}$, $SPAEI_{Turc}$ and $SPAEI_{RS-ET}$ for different accumulated periods 6, 12, 18, and 24 months for the period of 1951 to 2014 over Krishna river basin over Krishna River basin.

indices identified the major droughts as less intensified compared to SPEI. For example, the SPEI-12, $SPAEI_{Budyko}$-12, $SPAEI_{Turc}$-12 and $SPAEI_{RS-ET}$-12 values for the drought years of 2002 and 2003 were obtained as -1.43, -1.06, -1.10, -0.83 and -1.80, -1.15, -1.06, -0.88 respectively. The SPEI-18, $SPAEI_{Budyko}$-18 Budyko and $SPAEI_{Turc}$-18 values for the drought years of 1972 were obtained as -2.1, -1.33, -1.2 respectively. Over all, the severities of the drought indices at various time scales were found to be more with SPEI compared to $SPAEI_{Budyko}$, $SPAEI_{Turc}$ and $SPAEI_{RS-ET}$. As the calibration factor was estimated at an annual scale, $SPAEI_{Hydro}$ was estimated only at a 12-month scale for which the results are presented in the following paragraphs.

Therefore, the present study revealed that inclusion of AET in the drought assessment characterize the droughts as moderately intensified compared to severe droughts with PET based drought indices. Further, drought severities were noted to be less with $SPAEI_{Turc}$ compared to $SPAEI_{Budyko}$ due to the lower estimates of AET with Turc model 3.3, representing the dependence of severity or intensity of droughts on the quantification of AET estimates. Also, $SPAEI_{RS-ET}$ has identified less intensified droughts years with all accumulation periods (Table 3.3),revealing lower estimates of remote sensing-based ET values compared to empirical AET estimates. Overall, there exists a strong influence on drought severities with the use of various formulations of ET estimates in drought characterization.

Coming to the comparison of number of drought events over the basin capturing over a particular period, SPEI has estimated higher number of severe and extreme drought events compared to $SPAEI_{Budyko}$, $SPAEI_{Turc}$ and $SPAEI_{RS-ET}$. Figure 3.7 shows the comparison of severe and extreme drought frequencies (SPEI/ $SPAEI_{Budyko}$/ $SPAEI_{Turc}$/ $SPAEI_{RS-ET}$ ¡ -1.50) over Krishna river basin for 1983-2006. The PET based drought index (SPEI) has resulted in higher drought

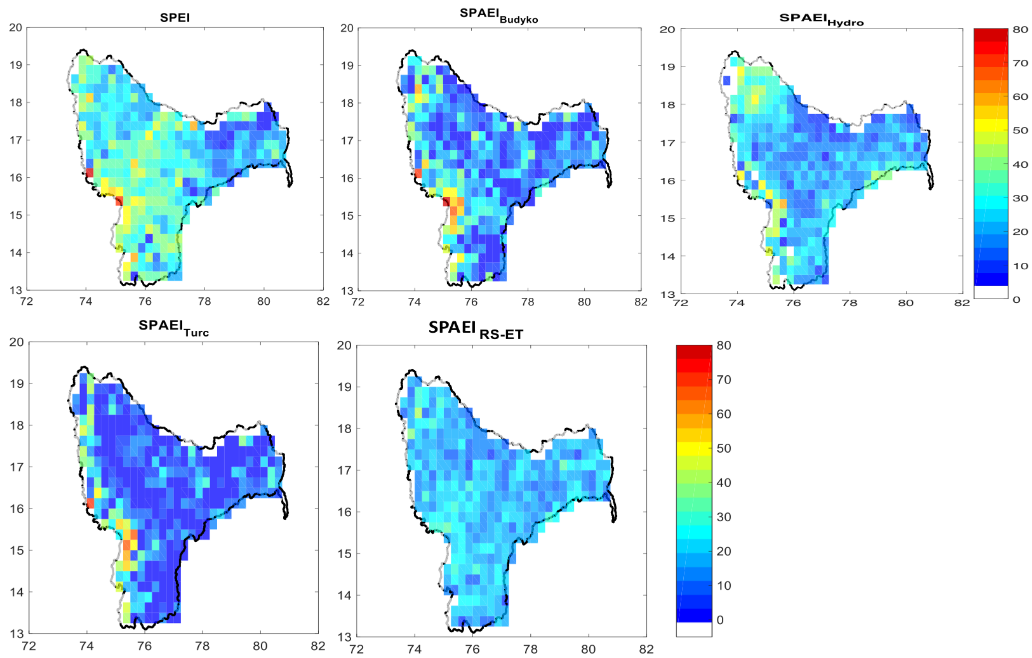


Figure 3.7: Number of severe and extreme drought events (SPEI/ $SPAEI_{Budyko}$/ $SPAEI_{TURC}$/$SPAEI_{RS-ET}$ $>$ -1.50) in months over Krishna river basin for 1983-2006

frequencies for the period of 1983-2006 compared to both empirical based AET ($SPAEI_{Budyko}$ and $SPAEI_{Turc}$) and remote sensing-based ($SPAEI_{RS-ET}$) drought indices over KRB. Further, lower severe and extreme drought frequencies were observed with Turc model compared to Budyko model. Whereas, remote sensing-based severe and extreme drought frequencies were more comparable with Budyko model drought frequencies for the river basin for the period of 1983-2006.

The present study considered a threshold of -1 for both SPEI and SPAEI formulations in drought duration estimation. Tthe drought duration was identified as the period of months which is continuous negative, starting from where the SPEI/ $SPAEI_{Budyko}$/ $SPAEI_{Turc}$/ $SPAEI_{RS-ET}$ values are more negative then -1 and ends when the SPEI/ $SPAEI_{Budyko}$/ $SPAEI_{Turc}$/ $SPAEI_{RS-ET}$ values turns out to be greater than -1. The duration and intensities of droughts were studied for the recent two consecutive drought years of 2002 and 2003 over the river basin for 6, 12, 18 and 24 months scales (Figure 3.8). The SPEI at 6-month drought duration was noted as from May to August and May to September for year 2002 and 2003 respectively with intensity as moderate. Whereas, the $SPAEI_{Budyko}$-6 and $SPAEI_{Turc}$-6 has identified 2002 and 2003 as normal conditions from March to September and February to September respectively. Whereas, the $SPAEI_{RS-ET}$ at 6-month time scale has identified the drought duration as April for 2002 and 2003 years. Similarly, for SPEI12 the duration of moderate drought months were noted from September 2002 to December 2003, whereas, $SPAEI_{Budyko}$ and $SPAEI_{Turc}$ has identified May and June 2003 as major drought affected months 3.8. The $SPAEI_{RS-ET}$ at 12-month scale has estimated the drought duration as June 2003. Similarly, such short and less intensified droughts were noted also with 18 and 24-month time scale of the drought

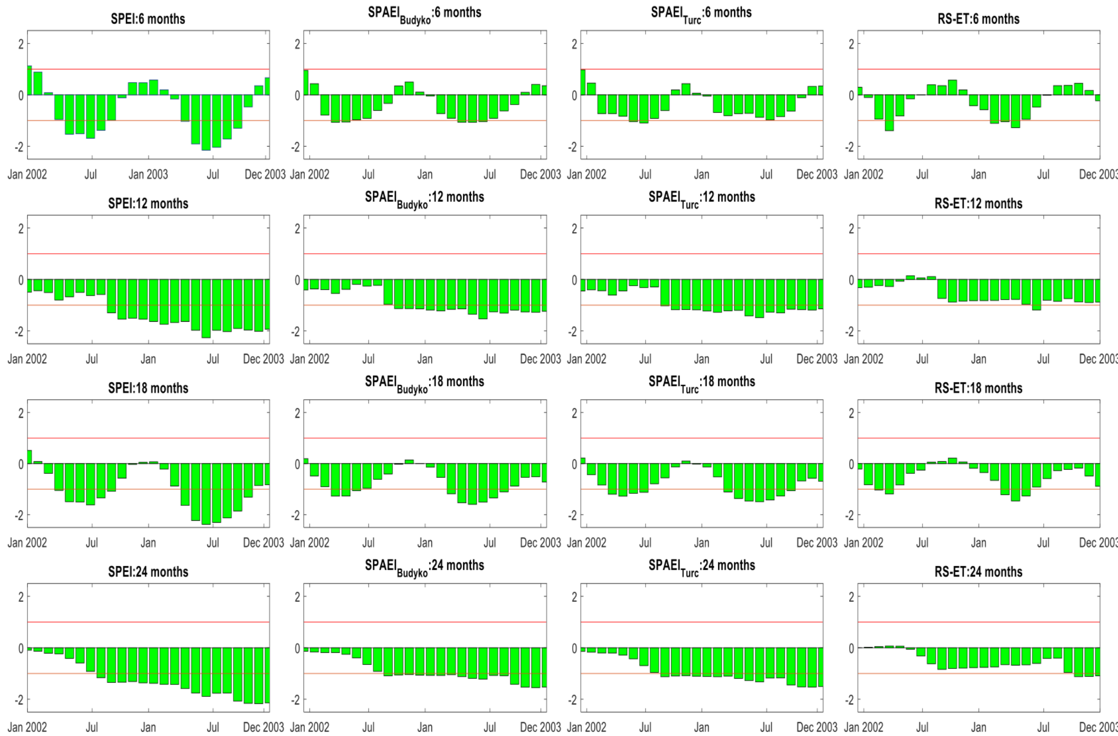


Figure 3.8: Duration and intensities of drought in months for the drought years of 2002 and 2003 over Krishna River basin for various time scales for SPEI and SPAEI

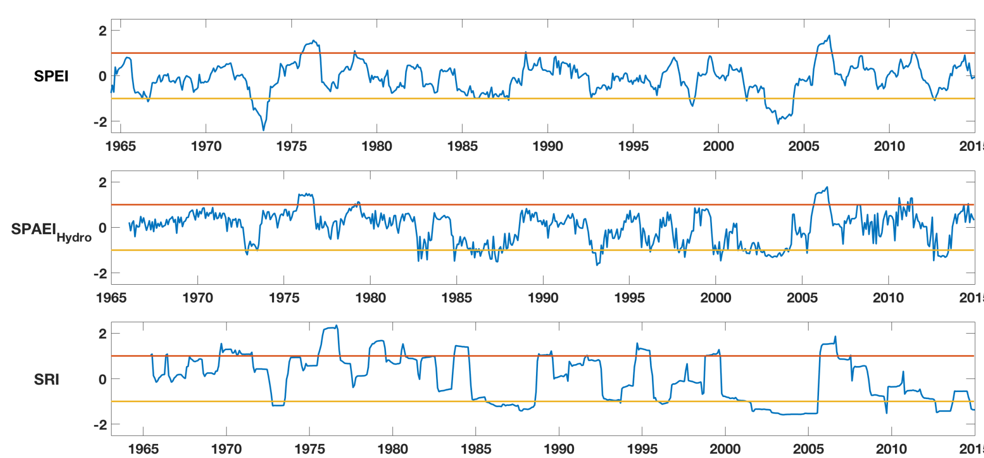


Figure 3.9: Time series of SPEI, $SPAEI_{Hydro}$ and SRI from 1966 to 2014 over Krishna river basin

indices. Further, the empirical based AET drought indices and remote sensing-based ET drought index has shown comparable drought durations at various accumulation periods. Overall, short drought durations were noticed with both AET based drought indices compared to PET based drought index.

The study further compared the hydrologically calibrated AET(AET_{Hydro}) based drought index ($SPAEI_{Hydro}$) with the SPEI and Standardised Runoff Index(SRI) (Figure 3.9) to show how the inclusion of runoff can play a role while estimating drought.

While SPEI overpredicts the drought in case of low rainfall and SRI overpredicts in case of low runoff even though there is high rainfall, $SPAEI_{Hydro}$ considers both rainfall and runoff and reliably estimates the drought.

Conclusion

The present study developed a drought index which can combine the structure of Standardized Precipitation and Evapotranspiration Index (SPEI) and actual evapotranspiration, the Standardized Actual Precipitation Evapotranspiration Index (SPAEI). The formulation of SPAEI based on P-AET, accounts for the water balance of the river basin representing the available water to fulfil the evaporation demands, where long-term storage losses are neglected. The use of AET in the drought estimation conceptually accounts for the soil water storage, water supply and energy available, therefore SPAEI can also characterize the hydrological drought conditions implicitly. The drought characterization based on two climatic water balances, one is with potential evapotranspiration and other with actual evapotranspiration were compared for Krishna river basin, India. The drought indices developed in the present study has revealed that inclusion of AET in the drought assessment will result in less intensified droughts compared with PET based drought indices. The PET based drought index, SPEI, overestimates the drought intensity as it is based on unlimited water supply and energy, whereas, the SPAEI is a reliable measure as it agrees better with the natural water budget of a river basin. The AET based drought indices were able to drive the areas into moderate, which or otherwise categorized under severe drought regions. These indices can be reliable measure to estimate the future meteorological as well as hydrological drought extent, severity and frequency. The next chapter provides a method to statistically downscale the meteorological inputs of precipitation and temperature using RCM and GCM projections and the drought impact assessment for the future scenarios.

Chapter 4

Climate Change Impact Assessment: Downscaling

Introduction

Climate change impact assessment studies are necessary in the context of changes in hydrological cycle under increase of anthropogenic global greenhouse gases in the atmosphere. Regional Circulation Models (RCMs) and General Circulation Models (GCMs) are two such climate models which simulates historical and future projections of various climate variables accounting for the greenhouse gases in the atmosphere which can be further used to assess the impacts of climate change. However, both the RCMs and GCMs are generally associated with systematic biases, resulting from inadequate physics and bias in GCM simulations. The present chapter demonstrates the use of RCM and GCM climate change projections in the drought impact assessment over Krishna river basin.

In this chapter, a quantile bias correction technique developed by [48] is presented in Section 4.1.1, followed by a statistical downscaling technique by [80] for downscaling the GCM outputs in Section 4.2.1. An analysis of the modelled meteorological and hydrological variables is carried out with the corrected RCM and the downscaled GCM outputs and the best models with accurate predictions are identified in Section 4.3. Then, the areal extent of drought along with the severity, frequency and duration is analyzed for these models in Section 4.4.

4.1 Methodology

4.1.1 Regional Circulation Model (RCM) Projections

The RCM projections provided by Coordinated Regional Downscaling Experiment(CORDEX) model outputs over South Asian domain can be used directly for the drought impact assessment after checking the performance of RCM outputs with the observed climate data. Given the systematic bias existing between the modeled and actual climate values, the RCM data sets can be used after correcting for bias [86]. The present study adopted quantile-based mapping method developed by [48], with the comparison of Cumulative Distribution Functions (CDFs) of observed and RCM simulated data of precipitation and temperatures for the historical and future scenarios.

- **Bias correction for Historical GCM,RCM data**

  Here, the CDFs of RCM and observed IMD data sets were compared to correct the bias present in RCM historical and future data sets [48], where Gamma distribution is used to calculate the CDFs of each time series as follows:

$$X_{m-p.adjust} = F_{o-c}^{-1}\left(F_{m-c}\left(x_{m-p}\right)\right) \tag{4.1}$$

  where $X_{m-p.adjst}$ is the bias-corrected climate variable for current period(RCM-historical)
  x_{m-p} is the biased RCM variable
  F_{m-c} is the CDF of RCM Historical data
  F_{o-c} is the CDF of Observed data
  F_{o-c}^{-1} is the inverse CDF of observed data which gives the observed variable at the corresponding equal CDF level.

- **Bias Correction for Future GCM/RCM Data**

  For a given percentile, we assume that the difference between the model and observed value during the training period also applies to the future period, which means the adjustment function remains the same. However, the difference or shift between the CDFs for the future and historic periods is also taken into account .

$$X_{m-p.adjst}^{'} = x_{m-p} + F_{o-c}^{-1}\left(F_{m-c}\left(x_{m-p}\right)\right) - F_{m-c}^{-1}\left(F_{m-p}\left(x_{m-p}\right)\right) \tag{4.2}$$

  where $X_{m-p.adjst}$ is the bias corrected climate variable for future period(RCM-future)
  x_{m-p} is the biased RCM future variable
  F_{m-p} is the CDF of RCM Future data
  F_{o-c} is the CDF of IMD(Observed) data and F_{o-c}^{-1} is the inverse CDF of IMD data which gives the IMD variable at the corresponding equal CDF level.
  F_{m-c} is the CDF of GCM/RCM Historical data and F_{m-c}^{-1} is the inverse CDF of RCM data which gives the RCM variable at the corresponding equal CDF level.

This ensures that the RCM data is standardized and can perform hydrological simulations with higher accuracy for models trained with IMD data. The statistical downscaling procedure to predict the precipitation and temperature using the GCM outputs is described in the following section.

4.2 Global Circulation Model (GCM) Climate Projections

GCMs are climate models designed to simulate time series of climate variables globally, accounting for the greenhouse gases in the atmosphere. Most of the climate change impact assessment studies mainly focus on the use of GCM outputs of various climatological variables and their integration with hydrological modelling [78, 90, 15]. The statistical downscaling technique used to predict precipitation and temperature from GCM projections for the assessment of climate change impact is explained in Section 4.2.1.

4.2.1 Statistical Downscaling

The statistical techniques used to bridge the spatial and temporal resolution gaps between what GCMs are currently able to provide and what impact assessment studies require are called as statistical downscaling methods. Downscaling is a technique by which the coarse resolution inputs of GCM can be re-estimated to a finer resolution of 0.25°(or) 0.5°(25-50km)(Spatial Downscaling) in monthly or lesser(sub monthly,daily) time scales(Temporal Downscaling). Statistical Downscaling used in this study establishes statistical relationships between historical and current largescale climate features (GCM predictor variables) and regional climate (Predictands - precipitation and temperature) to predict future local climate variables using GCM projections.

Following to the methodology proposed by [80], the present study established a statistical relationship between the predictor variables (the climatological variables - surface air temperature, wind speed, humidity etc) and the predictand variables (the hydrological variables to be predicted - rainfall, temperature). Based on the trained and tested statistical relationship established will be used with the GCM future climatological variables to predict the future projections of hydrological variables.

- **Data Preprocessing and Selection**

  - **Predictor data**

    NCEP/NCAR reanalysis data for surface air temperature, mean sea level pressure, specific humidity at 500 mb pressure level, zonal and meridional wind velocity at the surface level is extracted for the latitudes in the range of 12.5°to 20°N and the longitudes in the range of 72.5°to 82.5°E surrounding the entire Krishna basin. The data is extracted from ncfiles using the 'ncread' command in MATLAB. This data is present as a 3-dimensional matrix of longitude X latitude X timestep(monthly). As the observed discharge data is present from 1966 to 2014, the period from 1966-2003 is considered as the training period and the period from 2004-2014 as the validation period. The future period is divided into slices of 20 years for analysis, and the periods considered are 2021-2040, 2041-2060, and 2061-2080. The resolution and time period of the data used along with the number of latitudes and longitudes encompassing the basin is presented in the Table 4.1.

  - **Standardization and bias correction**

    Bias exists between GCM simulations and observed data which needs to be corrected. Hence, Bias Correction as mentioned in Section 4.1.1 is carried out using the NCEP predictors as the observed dataset and the GCM predictors in place of RCM data. This assumes

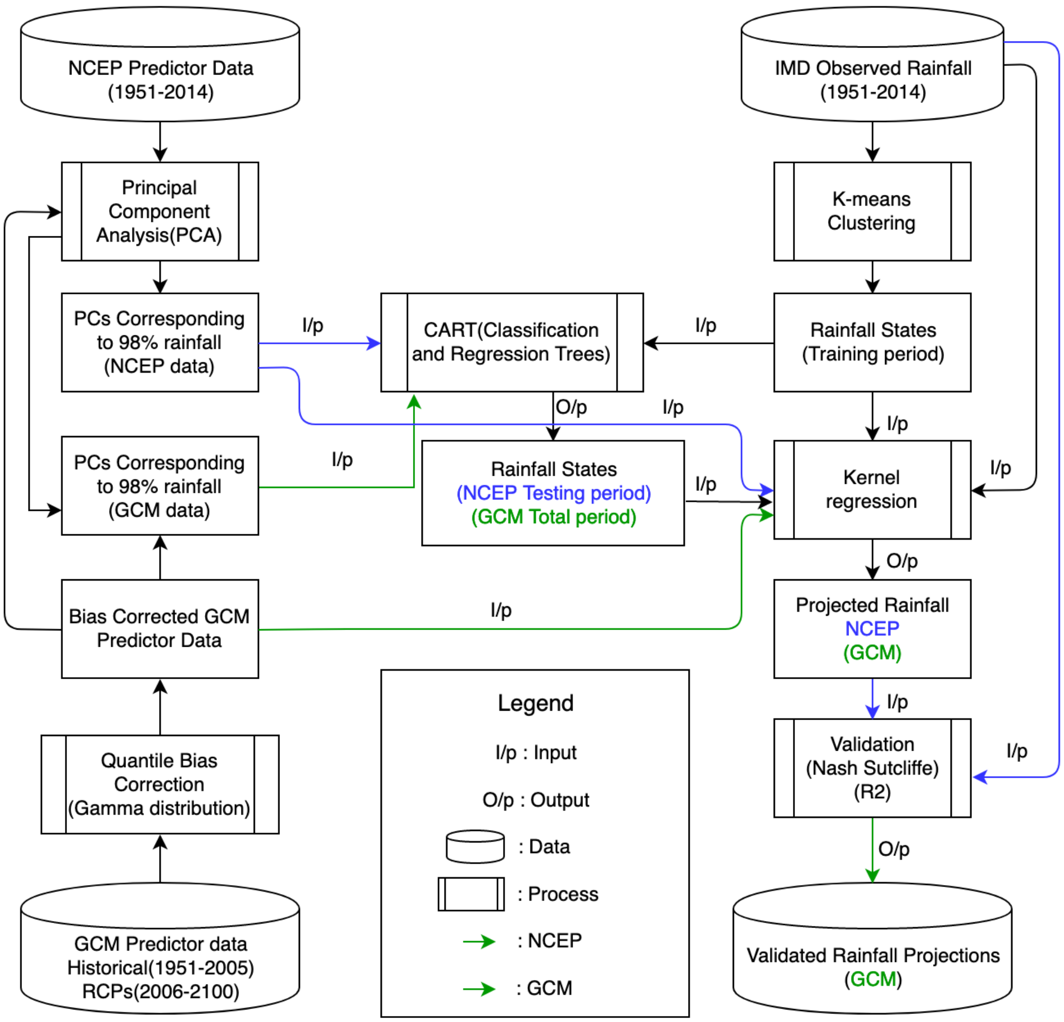


Figure 4.1: Statistical Downscaling of Rainfall

Dataset	Spatial resolution	Latitudes	Longitudes	Current period	Future period
IMD	0.25*0.25	25	31	1951-2014	
NCEP	2.5*2.5	4	5	1948-2017	
RCM	Variable(2.8*2.1)	5	6	1951-2005	2006-2100
GCM	Variable(2.8*2.8)	5	6	1951-2005	2006-2100

Table 4.1: Description of the resolution and period of availability of meteorological datasets required for downscaling

that the climate distribution doesn't change over time. Once the bias correction is done, the data is converted from a 3-dimensional matrix from the format longitude X latitude X time-step to a 2-dimensional matrix time-step X grid (each grid is taken as a column) for further analysis.

  - **Incorporating Seasonality**

    In addition to the bias corrected predictor data, as the precipitation and temperature each are distributed in a similar seasonal pattern across the months every year, the predictor data is modified to include the effect of seasonality by introducing one more predictor based on the month. This will ensure that the predictand variable (precipitation and temperature) across all years in a particular month follow a similarity.

Principal Components are extracted from the final predictor data and a statistical relationship is derived between observed high resolution (station level) variables and larger (NCEP/NCAR Reanalysis data) scale low resolution predictor variables using a transfer function.

- **Principal Component Analysis**

  The extracted reshaped data cannot be used directly for downscaling considering multicollinearity (high correlation among two or more predictor variables which are spatially closer) as well as multidimensionality (in case of high-resolution downscaling in large basins) problem which lead to an increase in computational time as well as resources. Principal Component Analysis(PCA) is a technique which will help to extract a reduced set of all features by a transformation of features without much loss of information. In other words, it helps us to reduce the problem of multidimensionality by considering the similar pattern of predictor data of a set of grids and replace these multiple columns of predictor data with a single column(time series) following the same pattern. This replacement does not majorly affect the classification or prediction models. The PCA algorithm is explained in detail in Appendix C.1. The requirement and procedure to establish this statistical relation from the extracted principal components is explained in more detail in the following section.

- **Rainfall State Estimation**

  One of the main assumptions in multisite downscaling is that all the stations in a region follow a homogeneous pattern of rainfall i.e the magnitude of rainfall received between two closely located stations or grids does not differ much in most of the cases. Excluding this assumption, if individual site downscaling is used, the effect of rainfall occurring in the surrounding sites is neglected which results in the model unable to capture this cross correlation. In order to rectify this, [56] introduced the concept of weather state which says that all the stations on a single time step follow the same categorization of rainfall. The three states considered are dry(nearly zero rainfall), medium and heavy rainfall. As a definite criterion of the magnitude of rainfall associated with each state is not categorized, and as it is entirely dependent on the basin characteristics and data, an unsupervised clustering algorithm needs to be adopted to identify the rainfall state for each month. . K-means clustering explained in Appendix C.2 is adopted in this study.

- **The Statistical Relationship**

  The bias-corrected predictors and the estimated rainfall states were considered as inputs for downscaling. A single multivariate regression model can't be applied to predict rainfall due to the high variance of rainfall data. Hence, observed rainfall has to be separated into categories or weather states mentioned previously following which individual regression models are to be applied separately to rainfall of each category.

  Classification and Regression Trees(CART)

  Classification and Regression Trees(CART) is a decision tree learning technique is used in prediction modelling. A decision tree is a machine learning model in which categorization of the data is done with each feature partitioning the data based on conditions or range of the features and works when the data has a finite set of values. In the case of continuous data, standard decision/classification trees will not be able to accurately capture the relationship between the predictor and predictand variables [51]. Hence Classification and Regression Trees(CART) explained in Appendix C.3 help in categorizing the rainfall into weather states by building a statistical relation between the continuous principal components extracted from predictor data and the rainfall states estimated using K-means clustering. The established relationship is assumed to be intact for the future predictors which are then taken as input for CART model and for which the future rainfall states are estimated. An advantage of Classification and Regression Trees over linear classification models is that they can capture non-parametric and non-linear relationships as well as yield simple models. Cross validation is carried in order to ensure there is no risk of overfitting the data.

- **Rainfall prediction using regression**

  Individual regression models are built on separating the predictor and observed data based on the weather state category into individual datasets. Kernel regression should be applied to capture the rainfall as it is a non-linear predictand variable. Kernel regression is a non-parametric technique which usually projects the non-linear data into a higher dimensional space where it can be linearly separated. This mapping is done using a function(Kernel). As the output expected is rainfall amount which is a real number, Support Vector Regression(SVR) explained in Appendix C.4 has been used instead of Support Vector Machine(SVM) Kernel regression. The key difference between linear regression and SVR is that in SVR, the model tries to fit the error within a certain threshold(ϵ), identifying a single separating hyperplane which maximizes the margin rather than solely minimizing the error which helps to find the best model.

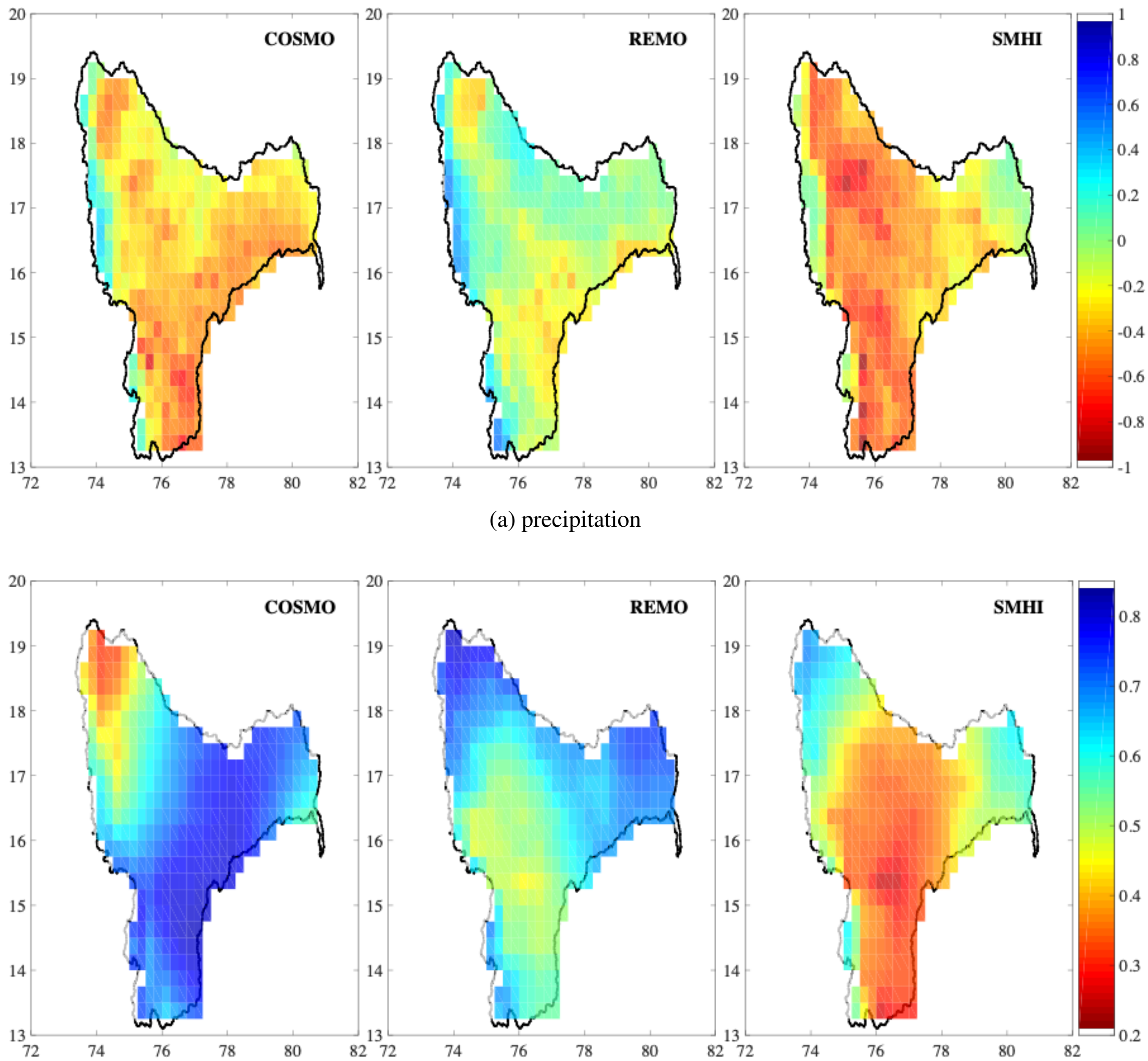
Linear SVR as used as the regression technique to predict the precipitation, temperature without overfitting. The observed IMD and modelled observations of meteorological, hydrological variables and drought with RCM and GCM outputs are presented in the following sections.

4.3 Results and Discussions

4.3.1 Climate Projections of Rainfall and Temperature with CORDEX RCM Data

For the assessment of precipitation and temperatures over KRB for the future scenarios, the CORDEX simulations were used. The precipitation and temperature data extracted from RCM outputs from 1965 to 2080, after bias correction were used to estimate the precipitation and temperatures over KRB. Basin-averaged precipitation and temperatures with the RCMs were studied for three future periods of 2021-2040, 2041-2060 and 2061-2080 along with historical period of 1965-2003 and 2004-2014 as given in Table 4.2. The study of the compatibility of RCM projections with observed data sets is a prominent step in the climate change projections assessment [86]. Therefore, the study compared the RCM climate projections with the observed data sets for historical period of 1965 to 2014 in reproducing the current climate variability. The bias corrected monthly precipitation and temperatures from each CORDEX RCM model were well compared with the observed IMD data for the period of 1965-2014 shown using the correlation between the two datasets (Figure 4.3). The RMSE (R-square) values estimated between observed precipitation and each RCM model outputs of COSMO, REMO, SMHI were estimated as 78.5 (0.2), 70 (0.23), 80 (0.15) respectively from 1966 to 2014. Whereas, the RMSE (R-square) values estimated between observed temperature and each RCM model outputs of COSMO, REMO, SMHI were estimated as 1.9 (0.6), 1.6 (0.7), 1.95 (0.52) respectively from 1966 to 2014. The NS values estimated between observed precipitation (temperature) and each RCM model outputs of COSMO, REMO, SMHI were estimated as 0.05 (0.79), 0.35 (0.67) and -0.28 (0.50) respectively from 1966 to 2014 as shown in figure 4.2.

Among the selected 3 RCMs the REMO model has shown best performance for simulating precipitation and temperatures over KRB. The precipitation has been predicted to increase under climate change signals with all 3 RCM models for the future periods of 2021 to 2080 over KRB. The increase in projections of annual average precipitation for the periods of 2021-2040 , 2041-2060 and 2061-2080 were compared with the observed data periods of 1966-2003 and 2004-2014 over KRB (Table 4.2) (Figure 4.4). The SMHI model has predicted highest increase of precipitation as varying from 101.6 to 136.7 mm to 217.1 mm for the future periods of 2021-2040 to 2041-2060 and 2061-2080 respectively compared to the current climate period of 1966-2003 over KRB. While, lowest precipitation projections, varying about 29.7 mm, 74.4 mm and 127.3 mm of increase for the period of 2021-2040 and 2041-2060 and 2061-2080 respectively compared to observed period of 1966-2003, were noted with COSMO model outputs. Moderate precipitation increasing projections were predicted with REMO for the RCP 4.5 for the future periods of 2021 to 2080. Overall, the projected increase of precipitation under climate



(a) precipitation

(b) temperatures

Figure 4.2: Comparison of Nash-Sutcliffe coefficients of observed and bias corrected RCM model outputs for period of 1966 to 2014

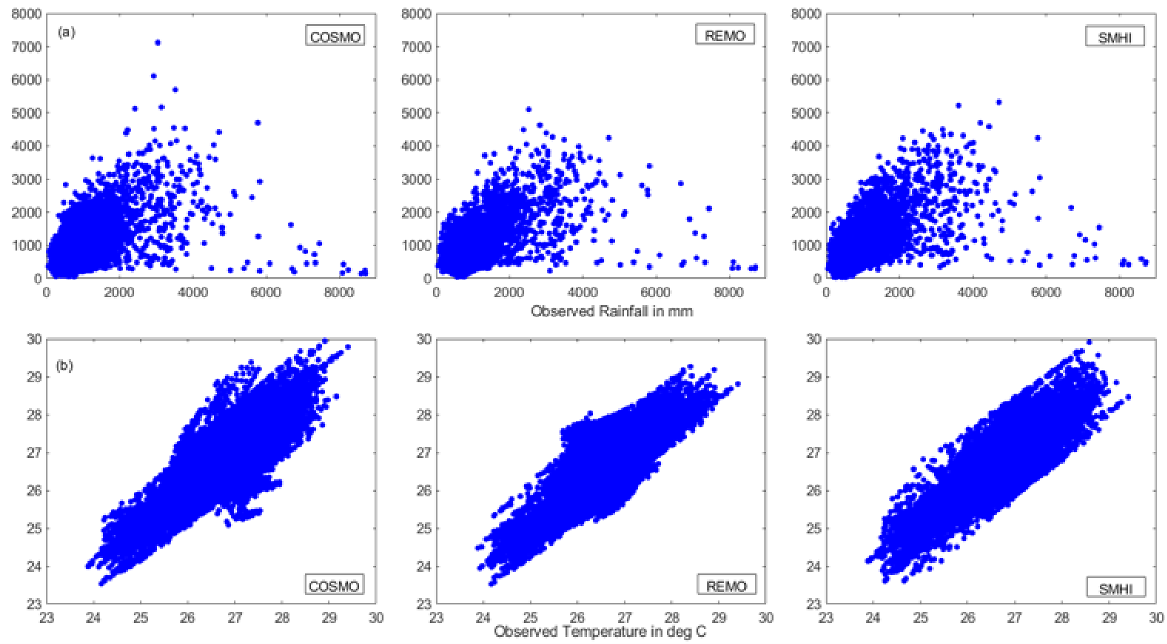


Figure 4.3: Scatter plots of observed and bias corrected RCM model outputs of precipitation and temperatures for period of 1966 to 2014

signals was predicted to be from 74.4 to 136.7 mm over KRB for the future period of 2041-2060 and 127.3 to 217.1 mm increase for the future period of 2061-2080 compared to the observed periods of 1966-2003.

About 1.11°C, 1.4°C to 1.85°C of increase in temperatures were predicted for the periods of 2021-2040, 2041-2060 and 2061-2080 respectively compared to the observed period of 1966-2014 over KRB (Figure 4.4) (Table 4.2).

The CORDEX RCM projections used for drought analysis were not able to give accurate results due to the underprediction of Precipitation extremes. Also, the RCM projections have shown negative NS coefficients with IMD precipitation data indicating that they are not able to capture the variation of rainfall. Hence the GCM projections were used to statistically downscale the meteorological inputs Precipitation and Temperature to obtain drought predictions from 1951 to 2080.

4.3.2 Climate Projections of Rainfall and Temperature with GCM Data

A rectangular region encompassing the entire basin is taken and it is assumed that rainfall follows a homogeneous pattern in this region. Depending on the NCEP resolution, for the Krishna river basin, the converted predictor matrix has around 100 columns(20 grids X 5 predictors). Considering that there will be a homogeneous pattern among these predictor variables as well as keeping the computational constraints in view, PCA is applied in order to reduce this problem of multidimensionality and multi-collinearity. For capturing 98% of variability of predictor data [80] about 15 principal components have

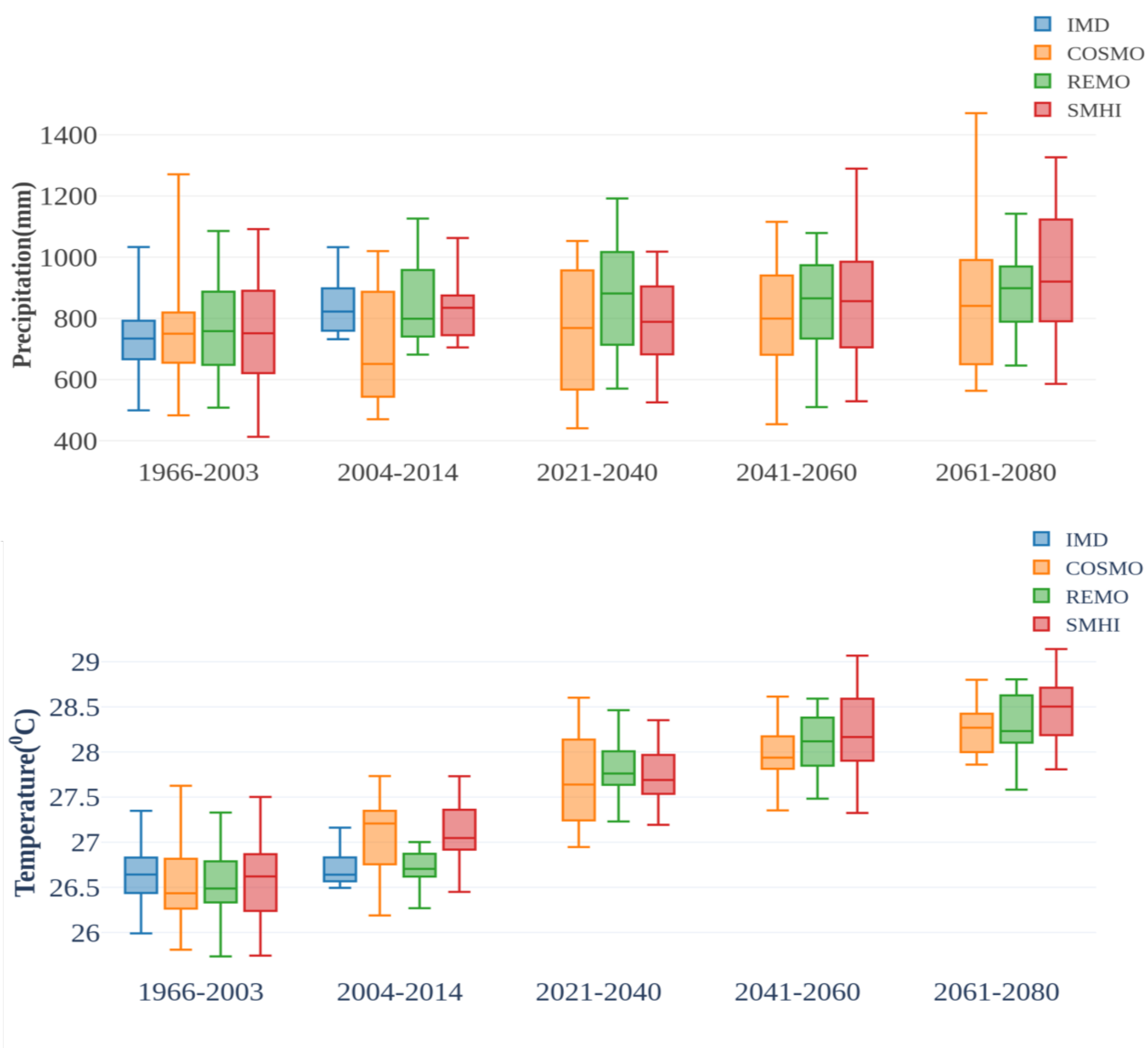


Figure 4.4: Basin averaged annual observed and projected precipitation and temperatures for the period of 1966-2003, 2004-2014, 2021-2040, 2041-2060 and 2061-2080 over KRB with various RCM model outputs

Hydrological Variable	RCM Name	Current period		Future period		
		1966-2003	2004-2014	2021-2040	2041-2060	2061-2080
Average Annual Precipitation(mm)	Observed	733.12	894.47	–	–	–
	COSMO	757.91	698.27	762.82	807.52	860.42
	REMO	755.36	841.84	864.29	850.7	887.18
	SMHI	773.33	834.73	834.73	869.82	950.2
Annual Temperature (°C)	Observed	26.6	26.7	–	–	–
	COSMO	26.53	27.05	27.71	28	28.25
	REMO	26.54	26.71	27.81	28.09	28.29
	SMHI	26.56	27.1	27.74	28.22	28.45

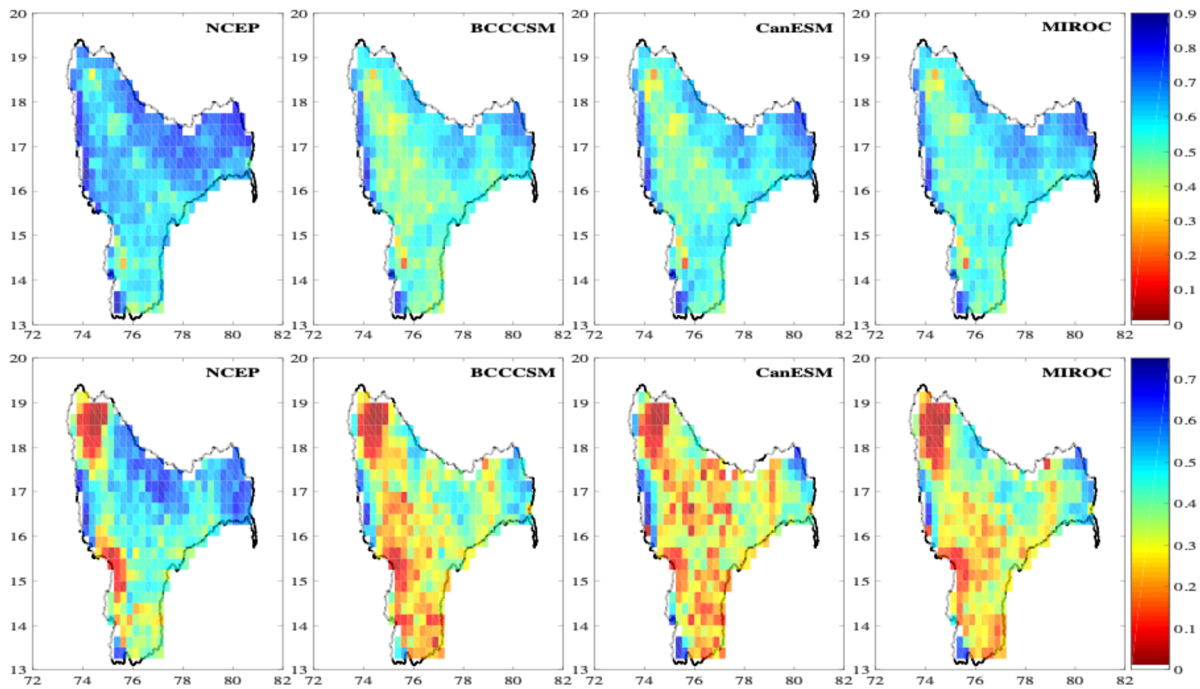
Table 4.2: Summary of spatial average annual water-energy variables for current (1966-2003, 2004-2014) and future period (2021-2040, 2041-2060) for KRB

been selected instead of 100 predictors in the statistical downscaling model. Then the rainfall states for the current period are obtained for the IMD-data using K-means clustering which are then given as inputs to CART model to find the rainfall states for GCM future datasets. Then, individual regression models are built for each grid for every state. Finally depending on the rainfall state of a month in the future, for every grid the corresponding regression model is identified and applied to get the amount of rainfall in that month. For temperature, a single regression model is built for each grid in contrast to multiple models for rainfall which is applied to get the future temperature predictions.

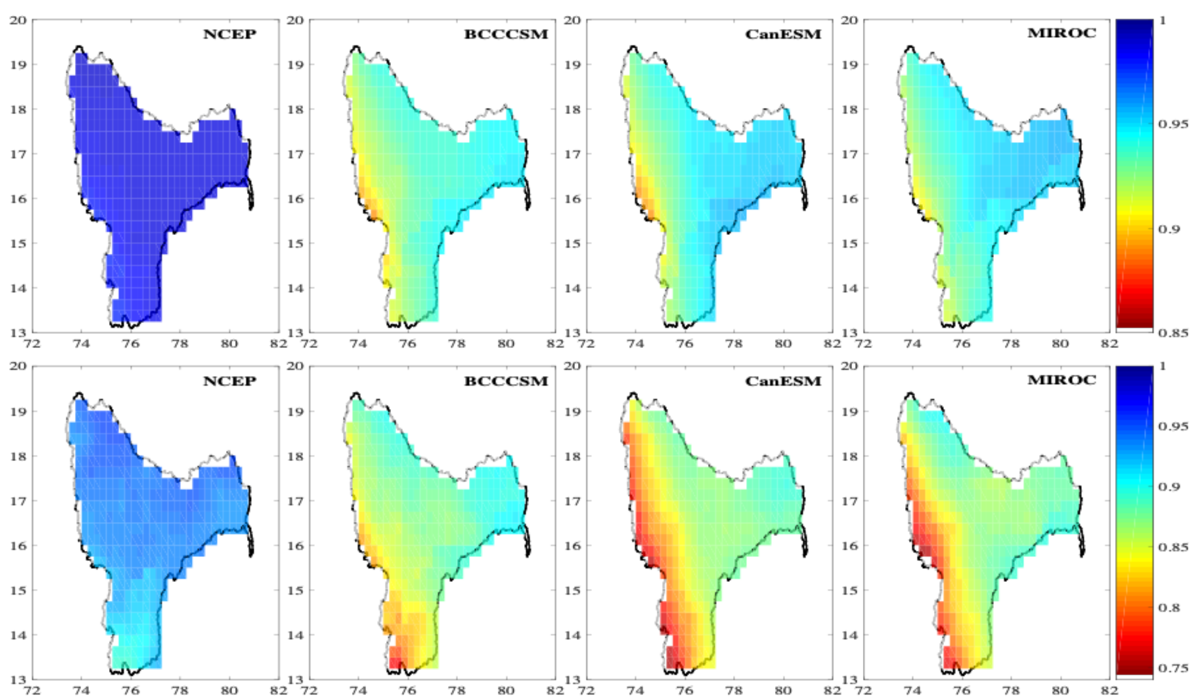
Basin-averaged precipitation, temperatures and drought indices were studied for three future time periods of 2021-2040, 2041-2060 and 2061-2080. The historical time period of 1951-1989 for training the downscaling model and the period 1990-2005 used for validation have been used as given in Table 4.3. The NS efficiency values were calculated for the IMD observed precipitation and temperatures with the downscaled GCM simulations. The average NS values for precipitation for the training (testing) data were estimated to be 0.65 (0.44) for NCEP, 0.54 (0.32) for BCCCSM, 0.56 (0.39) for CanESM, 0.57 (0.32) for MIROC for the period of 1951-1989(1990-2005). Whereas, the temperature values were estimated for training (testing) as 0.98 (0.93) for NCEP, 0.93 (0.87) for BCCCSM, 0.94 (0.85) for CanESM, 0.94 (0.86) for MIROC with observed data for the period of 1951-1989 (1990-2005). The mean values of estimated precipitation and temperatures for the considered time periods of training and testing were compared as shown in Table 4.3(Figure 4.6).

The increase in precipitation and temperatures for the future scenarios were predicted and compared with the observed time periods as given in Table 4.3. The BCCCSM, CanESM and MIROC GCMs have shown lower precipitation averages compared to the observed IMD data but an increase of 16.13 mm to 108.1 mm compared to the historical projections and an increase of 0.55 °C to 0.94 °C in temperature which is in comparison to various climate change reports [1].

[1]http://www.imd.gov.in/section/climate/StateLevelClimateChangeMonoFinal.pdf



(a) precipitation



(b) temperatures

Figure 4.5: Comparison of Nash-Sutcliffe coefficients of the observed and modelled GCM simulations for the Training and Validation period

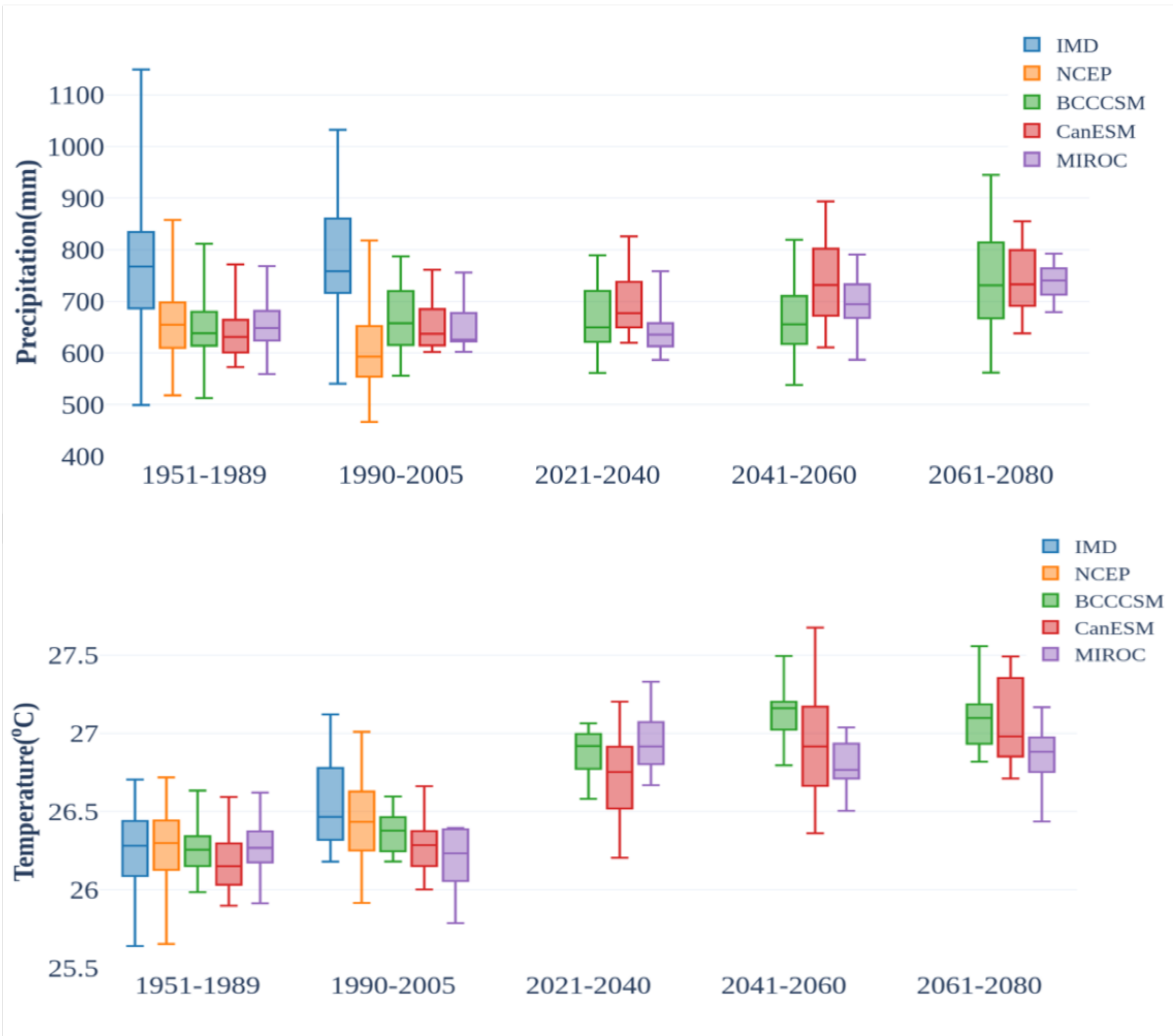


Figure 4.6: Basin averaged annual observed and projected (a) precipitation and (b) temperatures for the period of 1951-1989, 1990-2005, 2021-2040, 2041-2060 and 2061-2080 over KRB with various GCM model outputs

Hydrological Variable	GCM Name	Current period		Future period		
		1951-1989	1990-2005	2021-2040	2041-2060	2061-2080
Average Annual Precipitation(mm)	Observed	763.06	812.04	–	–	–
	NCEP	663.68	611.83	–	–	–
	BCCCSM	652.39	665.97	668.52	666.65	741.15
	CanESM	636.38	653.93	693.3	744.26	744.5
	MIROC	650.57	648.26	646.13	695.67	735.91
Annual Temperature (°C)	Observed	26.26	26.54	–	–	–
	NCEP	26.25	26.42	–	–	–
	BCCCSM	26.26	26.38	27.03	27.16	27.20
	CanESM	26.17	26.29	26.72	26.93	27.08
	MIROC	26.27	26.20	26.95	26.79	26.85

Table 4.3: Summary of spatial average annual water-energy variables for current (1966-2003, 2004-2014) and future period (2021-2040, 2041-2060) for KRB

4.3.3 Comparison of Estimated PET and hydrological induced AET with RCM and GCM projections

From the comparison of the performance of RCM and GCM (Figures 4.2, 4.5) simulations with observed data, the GCM projections which are based on the statistical relationships developed based on the observed climate have proven to be promising tools. The claim that GCM based statistical downscaling models work better than RCMs for hydrological drought prediction is further validated comparing the projected runoff, PET and hydrological induced AET estimated based on GCM projections with the estimates based on observed data (Table 4.4). While all the models have shown an increase in PET, RCM datasets have shown higher estimates of PET in the future with an increase of 300 mm to 500 mm in the future period whereas the GCM estimates have shown an increase of 100 mm to 200 mm. There is also an increase in the estimates of AET_{Hydro} where the increase is up to the range of 200 mm with RCM data, however there is an enormous increase of 250 mm to 500 mm in the AET_{Hydro} estimates of GCM data compared with the current period.

Modelled estimates of PET, AET_{Hydro} with GCM datasets are closer to the observed IMD estimations than the estimates of PET, AET_{Hydro} RCM datasets. Hence, drought analysis is carried out with the estimated meteorological and hydrological variables from the downscaled GCM Projections.

4.4 Comparison of drought estimations with GCM Projections

The standard structure of SPEI drought indices was compared with the proposed hydrometeorological drought index of $SPAEI_{Hydro}$. To find the best GCM for drought prediction of Krishna river basin, the drought areal extent for the historical drought years 1972, 1985, 2002 and 2003 is compared

Hydrological Variable	RCM/GCM Name	Current period		Future period		
		1966-2003	**2004-2014**	**2021-2040**	**2041-2060**	**2061-2080**
Annual Potential Evapotranspiration (mm)	Observed	1801.38	1816.00	–	–	–
	COSMO	1772.66	1911.00	2109.07	2197.44	2296.07
	REMO	1778.28	1823.63	2142.09	2226.67	2313.07
	SMHI	1782.09	1896.54	2093.22	2254.23	2329.27
		1951-1989	**1990-2005**	**2021-2040**	**2041-2060**	**2061-2080**
Annual Potential Evapotranspiration(mm)	Observed	1772.31	1831.13	–	–	–
	BCCCSM	1716.91	1758.94	1917.32	1965.66	1970.01
	CanESM	1712.73	1741.71	1871.67	1932.98	1971.68
	MIROC	1717.49	1697.60	1909.87	1859.31	1868.61
		1966-2003	**2004-2014**	**2021-2040**	**2041-2060**	**2061-2080**
Annual Calibrated Actual Evapotranspiration(mm)	Observed	383.54	533.00	–	–	–
	COSMO	356.02	430.69	383.95	428.82	481.38
	REMO	414.97	362.17	470.97	537.77	593.02
	SMHI	423.29	373.61	516.46	480.97	512.72
		1951-1989	**1990-2005**	**2021-2040**	**2041-2060**	**2061-2080**
Annual Calibrated Actual Evapotranspiration(mm)	Observed	340.22	495.96	–	–	–
	BCCCSM	571.62	611.81	611.43	621.58	674.76
	CanESM	564.31	594.07	654.35	710.18	684.63
	MIROC	575.09	592.16	580.63	613.75	648.49
		1966-2003	**2004-2014**	**2021-2040**	**2041-2060**	**2061-2080**
Annual Runoff at Vijayawada (m^3/s)	Observed	509.73	307.59	–	–	–
	COSMO	401.89	267.59	378.87	378.71	379.04
	REMO	340.39	479.68	393.32	312.94	294.17
	SMHI	350.04	461.12	265.99	388.86	437.49
		1951-1989	**1990-2005**	**2021-2040**	**2041-2060**	**2061-2080**
Annual Runoff at Vijayawada(m^3/s)	Observed	586.56	362.63	–	–	–
	BCCCSM	670.63	468.63	475.32	375.26	552.77
	CanESM	599.94	498.43	377.8	291.98	530.01
	MIROC	660.06	467.09	545.33	681.97	727.88

* Observed - IMD

Table 4.4: Summary of spatial average annual water-energy variables for current (1966-2003, 2004-2014) and future period (2021-2040, 2041-2060) for KRB

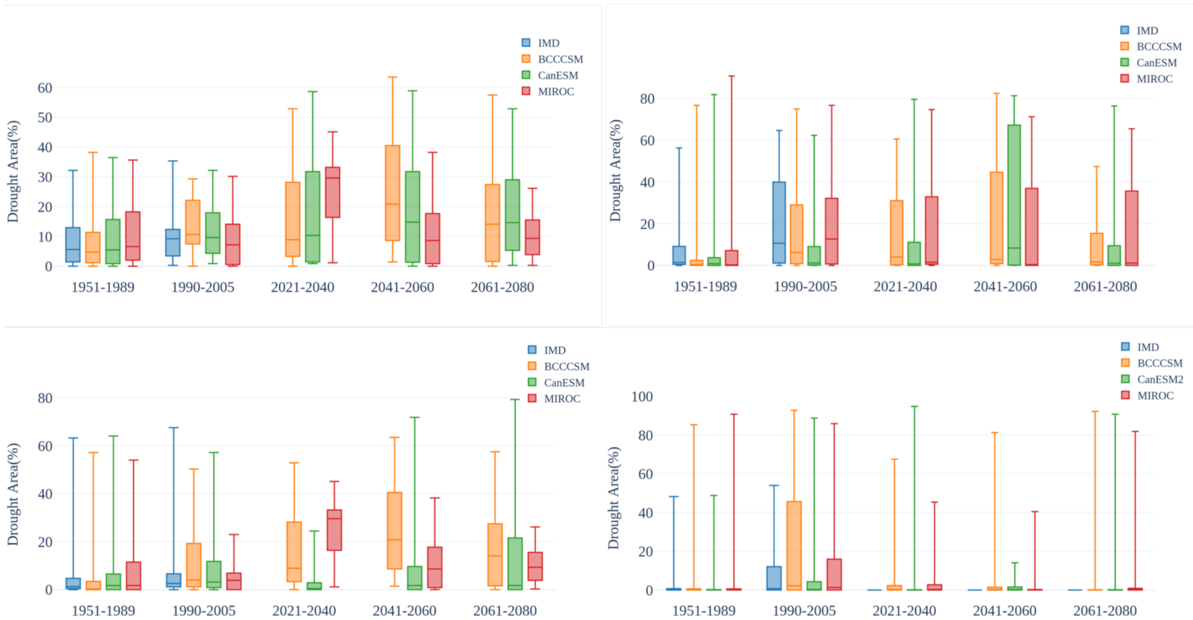


Figure 4.7: Comparison of moderate(upper), severe and extreme(lower) drought area estimated with SPEI(left) and $SPAEI_{Hydro}$(right) indices of GCM projections for the current(left) and future(right) periods.

with three GCM simulations in terms of drought areal extent. While BCCCSM was able to capture the drought for almost all of the drought years, followed by MIROC which was also to capture most of the drought events, the CanESM GCM was not able to capture the major drought events.

The estimated moderate and severe drought area has been increasing for the future periods compared to the current period however the extreme area is not captured properly attributed to the fact that the employed downscaling model do not model extreme data effectively. The drought area has increased from 14.8% to 23%-44% for the 2021-2040 period, 14%-39% for the 2041-2060 period and 15.4%-30.75% for the 2061-2080 period, there is a net increase of 25%-31% from the current to future periods.

Here, Drought is also analyzed in terms of average intensity estimated over the whole basin, frequency of droughts for every time period considered as well as the average duration of drought in the considered time period. Intensity(in terms of negative value) is highly negative for BCCCSM and MIROC as well as more frequently less than '-1' compared to the observed data. The time period after 2000 has a greater number of negative intensity values than the period before that indicating an increase in the drought frequency in the future period. However, the drought intensity is not as high for the severe and extreme drought years compared to the IMD data as the extremes are not captured accurately in downscaling models.

The drought frequency (the number of droughts occurring in a time period), the drought duration and the average extent of drought were analyzed for the three future time slices of 2021-2040, 2041-2060 and 2061-2080. The frequency of drought is more with SPEI compared to the drought frequency with

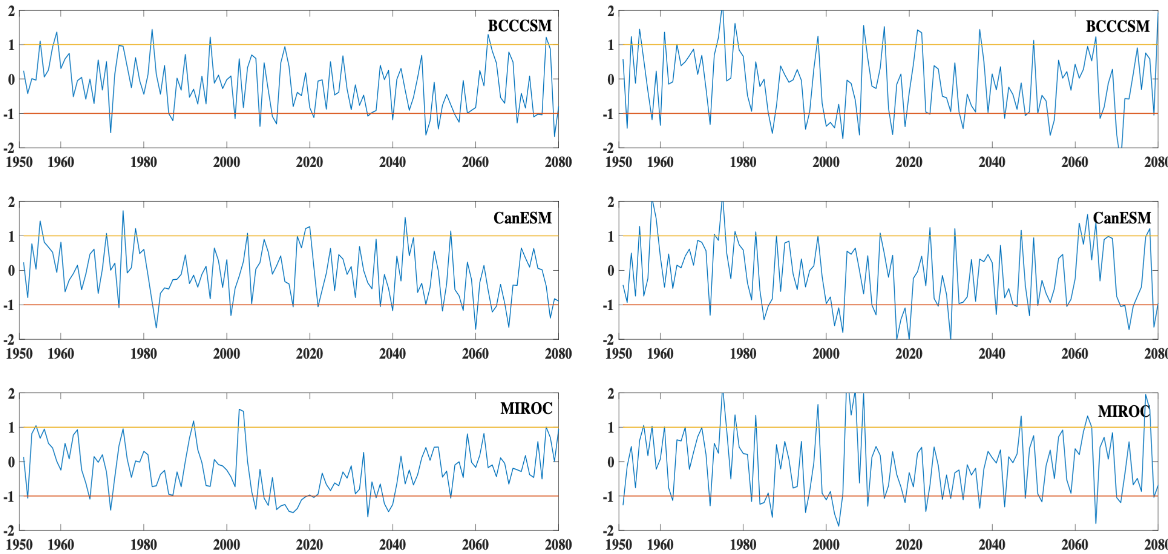


Figure 4.8: Drought Intensity of basin averaged SPEI(left) and $SPAEI_{Hydro}$(right) for the GCM for the time period 1951-2080

$SPAEI_{Hydro}$ for most of the basin and higher for the northern regions of the basin. While the average drought frequency for the current period 1951-2014 is 9.5 years with SPEI and around 9 years with $SPAEI_{Hydro}$. The analysis showed that in the future the number of drought years are increasing to 5 to 6 years for every 20 years with SPEI and 4.5 years of drought for every 20 years with $SPAEI_{Hydro}$. Comparing each GCM (Figure 4.9), the BCCCSM has shown higher drought frequencies in the 2041-2060 period and a basin average of 5, 8 and 6 years of drought with SPEI and 4, 5 and 4 years of drought with $SPAEI_{Hydro}$ respectively for the three future time slices 2021-2040, 2041-2060 and 2061-2080. MIROC has shown more droughts in the 2021-2040 period with SPEI in the central region of the basin and for both 2021-2040 as well as the period 2061-2080 with $SPAEI_{Hydro}$ in the north western region. The average drought frequency of MIROC GCM in the basin with SPEI is 9, 3 and 3 years and with $SPAEI_{Hydro}$ it is about 4.5, 4 and 5 years for the three time slices. However, CanESM has shown higher drought frequencies in the southern parts of the basin with $SPAEI_{Hydro}$ for the period 2041-2060. The average drought frequency of CanESM GCM in the basin with SPEI is about 4, 5.7 and 6.5 years and with $SPAEI_{Hydro}$ is about 4, 6.7 and 5 years for the three time slices. Overall, highest number of droughts were observed in the time period 2061-2080 with both SPEI and $SPAEI_{Hydro}$ drought indices, which directs for the proper water resources management of the river basin.

Finally, the drought duration for each GCM (Figure 4.10) was analyzed. While the average duration with SPEI is 6 months and with $SPAEI_{Hydro}$ is about 3 months for the current period, which is predicted to increasing for about 7 months with both SPEI and $SPAEI_{Hydro}$ with the estimations from GCM datasets. Indicating there will not be any considerable difference in the drought duration characteristics with SPEI and $SPAEI_{Hydro}$ for the current periods. The BCCCSM GCM has shown high basin averaged drought duration increasing from 7-9 months with SPEI and 2-3 months with $SPAEI_{Hydro}$

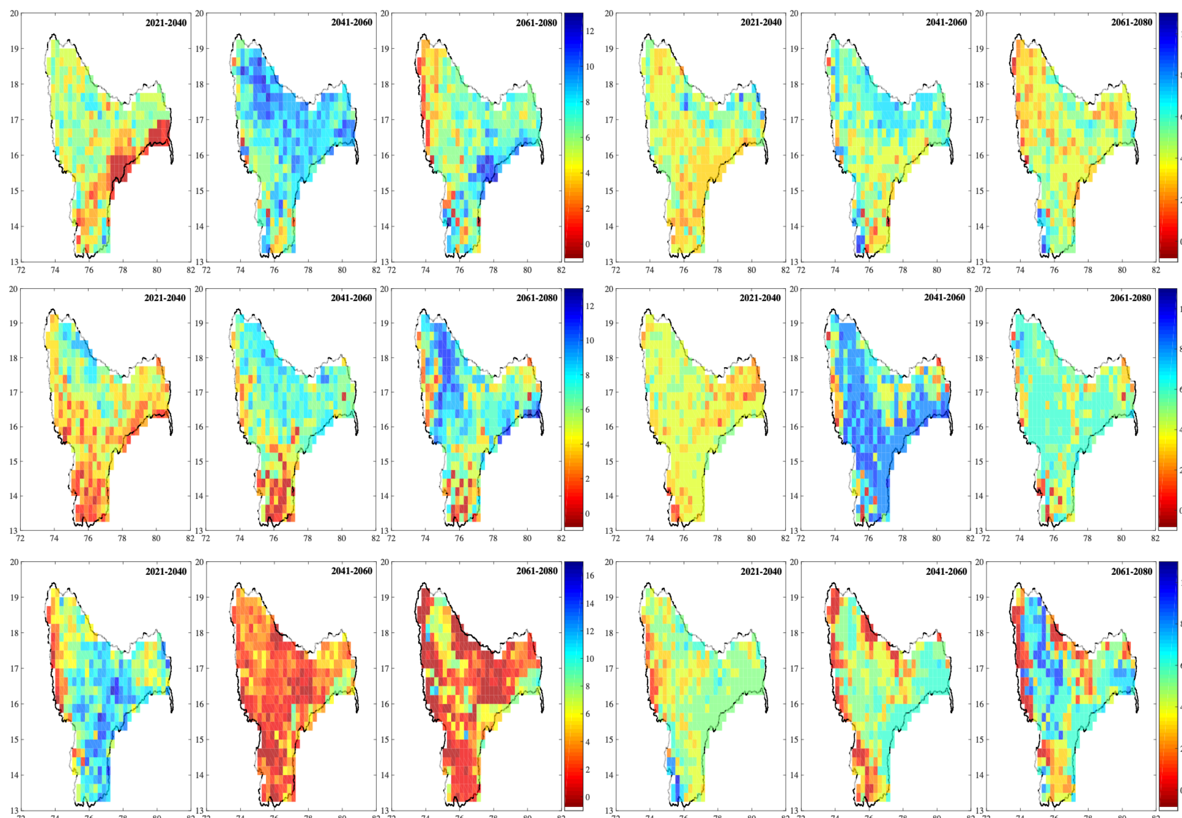


Figure 4.9: Frequency of Drought Occurrence (in years) for the GCMs with SPEI(left) and $SPAEI_{Hydro}$(right) for the time period 2021-2080

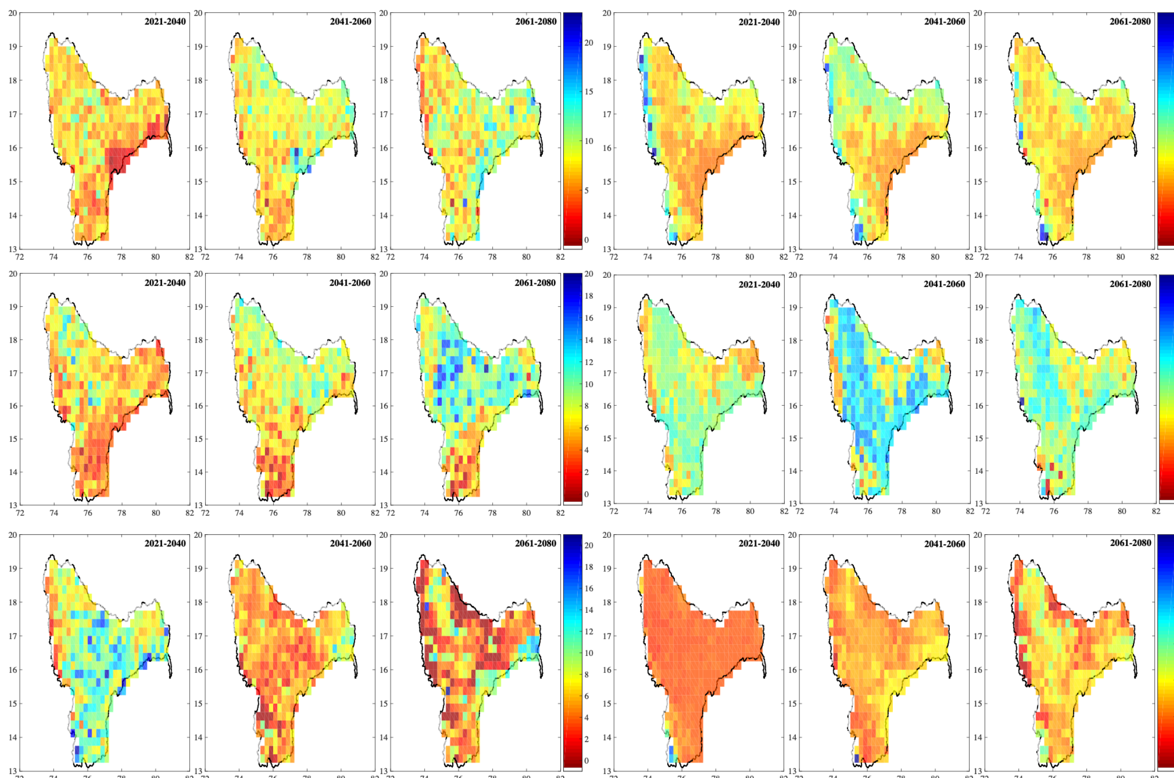


Figure 4.10: Drought duration (in months) in the basin with SPEI(left) and $SPAEI_{Hydro}$(right) for the GCM for the time period 2021-2080.

respectively over the 2021-2080 period. MIROC has shown more drought duration in the 2021-2040 period (10 months) with SPEI and 5 months for the other periods 2041-2060 and 2061-2080 but an increase in duration from 1.5-3.5 months with $SPAEI_{Hydro}$. CanESM has shown higher drought in the southern parts of the basin for the time period 2041-2060 with $SPAEI_{Hydro}$. The drought duration has been increasing from 6-9 months for the CanESM over the three time periods with SPEI and increasing from 4-5 months with $SPAEI_{Hydro}$ until 2060 and then decreasing back to 4 months for the period 2061-2080. High drought durations were observed over the basin for the time period of 2061-2080 with both the drought indices predominantly in the northern parts of the basin following to the projected increase of temperatures (Figure 2.7). Some of the regions have also shown drought durations of more than 20 months indicating up to 2 years of continuous drought which can directly affect the crop productivity and food security.

Conclusions

Droughts have been increasing alarmingly in the future. There is an increase in the average drought area, drought frequency and drought duration as well as the frequent increase in intensity over Krishna River basin. While the RCMs are not able to capture the increasing drought conditions exactly due to the inaccurate precipitation and temperature variability with IMD data, the statistical downscaling projections based on BCCCSM and MIROC GCMs are able to capture these increases accurately in terms of spatial variability in correlation with precipitation and temperature changes. Out of the 3 GCMs, it is recommended to use the BCCCSM GCMs as it has shown the continuous increase in drought over the time period as expected. Overall, highest number of droughts were observed in the time period 2061-2080 with both SPEI and $SPAEI_{Hydro}$ drought indices, which directs for the proper water resources management of the river basin. High drought durations were observed over the basin for the time period of 2061-2080 with both the drought indices predominantly in the northern parts of the basin following to the projected increase of temperatures.

Chapter 5

Conclusions

With the increasing intensity, frequency and duration of droughts worldwide under climate change, evaluation of variability associated with retrospective drought events will be valuable for the water and food security at river basin scales. Various meteorological drought indices have evolved to characterize the drought at larger scales. However, at river basin scales meteorological drought assessments may not be enough to understand the hydrological droughts which define the water availability of the crops. The present thesis aimed to develop a hydrometeorological drought prediction index by considering precipitation, evapotranspiration and runoff which is derived from hydrological inputs estimated from hydrological modelling. For this, the study aimed to estimate the Actual Evapotranspiration (AET) which is calibrated with precipitation, potential evapotranspiration and runoff to represent hydrologically calibrated AET, which is further used in the drought index Standardized Precipitation Evapotranspiration (SPEI) index to develop a new hydrometeorological drought index, Standardized Precipitation Actual Evapotranspiration Index (SPAEI). Further, the study used Regional Climate Models (RCMs) and General Circulation Model (GCM) outputs to study the climate change impact on hydrometeorological drought index. The proposed methodology was demonstrated on Krishna River Basin, India. The following paragraphs give the summary and conclusions of the study presented in the thesis.

- The study proposed hydrologically calibrated AET considering precipitation, Potential Evapotranspiration (PET) and runoff simulated by hydrological model. The hydrologically calibrated AET was observed to be comparable with remote sensing-based ET estimates and water balance approach of AET estimation.

- The hydrologically calibrated AET was then forced in the SPEI formulation to develop hydrometeorological drought index Standardized Actual Evapotranspiration Index (SPAEI). The study assessed the drought characteristics over Krishna river basin using SPEI and SPAEI and identified the years 1972, 1985, 2002 and 2003 as major drought years over Krishna river basin, which is also considered as major drought events all over India.

- The study compared various drought characteristics of intensity, severity, frequency and duration using both SPEI and SPAEI over Krishna river basin. The results of the study revealed that SPEI overestimates the drought intensity as it is based on unlimited water supply, whereas, the SPAEI is a reliable measure as it agrees better with the natural water budget of a river basin.

- The research findings of the present thesis reveal that use of AET may not have a major effect on the drought duration assessment. However, the drought severity, areal extent and frequency will have more pronounced effect by the inclusion of AET compared to PET in the drought characterization.

- The study used RCM and GCM climate change projections of precipitation and temperature to study the climate change impacts on drought characteristics using SPEI and SPAEI over Krishna River basin. The mean conditions of both the bias corrected RCM outputs, and statistically downscaled GCM outputs have been found to be well comparable with the observed IMD precipitation and temperature data.

- The study observed that compared to RCM climate projections, GCM based statistical downscaling models have performed well in capturing the historical observations, indicating reliable predictions based on GCM based observations at river basin scales. Although, statistical downscaling models has proved to be reliable projections for climate change impact assessment studies, the RCM data sets provides ease of quick implementation with less computational effort. Therefore, the study suggests the use of RCM projections for a preliminary understanding and GCM based projections for a detailed climate change impact assessment study.

- The projected drought characteristics based on GCM analysis has observed that while the intensity and duration of the drought has been found to be slightly increasing, the frequency and the drought areal extent have been increasing alarmingly over the Krishna river basin.

Related Publications

Journals:

1. S. Rehana, G. Sireesha Naidu, N.T. Monish, U.Sowjanya
   Modeling Hydro-climatic Changes of Evapotranspiration over an Arid River Basin of India
   Journal of Water and Climate Change, IWA Publishing, https://doi.org/10.2166/wcc.2020.173

2. Rehana, S., G. Sireesha Naidu, N.T. Monish
   Estimation of Annual Regional Drought Index Considering the Joint Effects of Climate and Water Budget for Krishna River Basin, India
   Environmental Monitoring and Assessment, MS. No: EMAS-D-19-03139 (Pub: Springer) https://link.springer.com/article/10.1007/s10661-020-08379-y

3. G. Sireesha Naidu, M. Pratik and S. Rehana
   Modelling Hydrological Responses under Climate Change using Machine Learning Algorithms – Semi-Arid River Basin of Peninsular India
   Evolutionary Algorithms in Water Resources, H2Open Journal, IWA Publishing (Under Review)

Book Chapters:

1. Shaik Rehana, Galla Sireesha Naidu, Nellibilli Tinku Monish
   Spatio-temporal Variations of Precipitation and Temperatures under CORDEX Climate Change Projections: A Case study of Krishna River Basin,India.
   Contemporary Environmental Issues and Challenges in Era of Climate Change, Edited by Singh, Pooja, Singh, Rajeev Pratap, Srivastava, Vaibhav, ISBN13: 9789813295940(Pub: Springer) https://www.springer.com/gp/book/9789813295940

Conferences:

1. Rehana, S., Sireesha Naidu, G., and N.T. Monish
   Climate change signals on regional water energy variables over Krishna river basin, India
   16th Annual Meeting, Asia Oceania Geosciences Society (AOGS), 28 Jul – 2 Aug, 2019, Singapore

2. Rehana, S., and Monish, N. T, and Sireesha, G.
   A Comparative Analysis of Regional Drought Characterization Over Krishna River Basin in India Using Potential and Actual Evapotranspiration
   15th Annual Meeting, Asia Oceania Geosciences Society(AOGS), 03 - 08 Jun, 2018, Honolulu, Hawaii, USA, .

3. Rehana, S., Sireesha Naidu, G., Apaar, A., Rajan, K.S.
   Spatio-temporal evaluation of evapotranspiration over Krishna river basin, India
   Hydro 2019 International Conference, Hydraulics, Water Resources & Coastal Engineering, 18 - 20 Dec , 2019

4. Rehana, S., Sireesha Naidu, G., Apaar, A., Rajan, K.S.
   Drought detection and assessment over Krishna River Basin with Standardized Precipitation Evapotranspiration Index
   Hydro 2019, International Conference, Hydraulics, Water Resources & Coastal Engineering, 18 - 20 Dec, 2019

5. Galla Sireesha Naidu, Shaik Rehana
   Impact of climate change on reservoir inflows using hydrological modelling framework and open source GIS softwares
   FOSS4G-Asia 2017- Oral Presentation

Poster Presentation

1. Shaik Rehana, Sireesha, G. Monish, N.T.
   Study of Observed Precipitation and Temperature Extreme Indices over India
   30th Conference on Climate Variability and Change, 24th Conference on Probability and Statistics, and 16th Conference on Artificial Intelligence, 28–29 July 2017, Baltimore, Maryland, U.S.A.

2. S. Rehana, G. Sireesha Naidu and N.T. Monish
   Assessment of Hydro-Climatic Changes over Krishna River Basin, India
   Workshop on Climate Extremes, Societal Resiliency and Krishna River Basin, 7-8 January 2020, Department of Civil Engineering, IIT Hyderabad, India

Bibliography

[1] S. Aadhar and V. Mishra. High-resolution near real-time drought monitoring in south asia. *Scientific data*, 4:170145, 2017.

[2] F. Abramopoulos, C. Rosenzweig, and B. Choudhury. Improved ground hydrology calculations for global climate models (gcms): Soil water movement and evapotranspiration. *Journal of Climate*, 1(9):921–941, 1988.

[3] K. Allen, J. Ogden, B. Buckley, E. Cook, and P. Baker. The potential to reconstruct broadscale climate indices associated with southeast australian droughts from athrotaxis species, tasmania. *Climate Dynamics*, 37(9-10):1799–1821, 2011.

[4] A. Anandhi, V. Srinivas, R. S. Nanjundiah, and D. Nagesh Kumar. Downscaling precipitation to river basin in india for ipcc sres scenarios using support vector machine. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 28(3):401–420, 2008.

[5] S. M. Asokan, J. Jarsjö, and G. Destouni. Vapor flux by evapotranspiration: Effects of changes in climate, land use, and water use. *Journal of Geophysical Research: Atmospheres*, 115(D24), 2010.

[6] H. N. Bhalme and D. A. Mooley. Large-scale droughts/floods and monsoon circulation. *Monthly Weather Review*, 108(8):1197–1211, 1980.

[7] O. Boucher, G. Myhre, and A. Myhre. Direct human influence of irrigation on atmospheric water vapour and climate. *Climate Dynamics*, 22(6-7):597–603, 2004.

[8] W. Brutsaert. Evaporation into the atmosphere: Theory. *History, and Applications. D. Reidel*, 1982.

[9] M. Budyko. The heat balance of the earth's surface, translated by nina a. *Stepanova (from Teplovoi balans zemnoi poverkhnosti, Gidrometeorologicheskoi izdatel'stvo, Leningrad, 1956, 255 pages), US Department of Commerce, Washington, DC*, 1958.

[10] M. Budyko. Climate and life. international geophysical series, vol. 18, 1974.

[11] P. Bühlmann. Bagging, boosting and ensemble methods. In *Handbook of Computational Statistics*, pages 985–1022. Springer, 2012.

[12] C. Buontempo, C. Mathison, R. Jones, K. Williams, C. Wang, and C. McSweeney. An ensemble climate projection for africa. *Climate dynamics*, 44(7-8):2097–2118, 2015.

[13] W. Buytaert and B. De Bièvre. Water for cities: The impact of climate change and demographic growth in the tropical andes. *Water Resources Research*, 48(8), 2012.

[14] S. Chander and H. Shanker. Unit hydrograph based forecast model. *Hydrological sciences journal*, 29(3):279–291, 1984.

[15] J. Chen, F. P. Brissette, A. Poulin, and R. Leconte. Overall uncertainty study of the hydrological impacts of climate change for a canadian watershed. *Water Resources Research*, 47(12), 2011.

[16] A. Chowdhury, M. Dandekar, and P. Raut. Variability in drought incidence over india–a statistical approach. *Mausam*, 40(2):207–214, 1989.

[17] B. I. Cook, J. E. Smerdon, R. Seager, and S. Coats. Global warming and 21 st century drying. *Climate Dynamics*, 43(9-10):2607–2627, 2014.

[18] A. Dai. Drought under global warming: a review. *Wiley Interdisciplinary Reviews: Climate Change*, 2(1):45–65, 2011.

[19] J. Das and N. Umamahesh. Spatio-temporal variation of water availability in a river basin under cordex simulated future projections. *Water Resources Management*, 32(4):1399–1419, 2018.

[20] U. De, R. Dube, and G. P. Rao. Extreme weather events over india in the last 100 years. *J. Ind. Geophys. Union*, 9(3):173–187, 2005.

[21] G. Destouni, F. Jaramillo, and C. Prieto. Hydroclimatic shifts driven by human water use for food and energy production. *Nature Climate Change*, 3(3):213, 2013.

[22] G. K. Devia, B. Ganasri, and G. Dwarakish. A review on hydrological models. *Aquatic Procedia*, 4:1001–1007, 2015.

[23] A. Dobler and B. Ahrens. Precipitation by a regional climate model and bias correction in europe and south asia. *Meteorologische Zeitschrift*, 17(4):499–509, 2008.

[24] R. Donohue, M. Roderick, and T. R. McVicar. On the importance of including vegetation dynamics in budyko? s hydrological model. *Hydrology and Earth System Sciences Discussions*, 3(4):1517–1551, 2006.

[25] F. Feser, B. Rockel, H. von Storch, J. Winterfeldt, and M. Zahn. Regional climate models add value to global model data: a review and selected examples. *Bulletin of the American Meteorological Society*, 92(9):1181–1192, 2011.

[26] H. J. Fowler, S. Blenkinsop, and C. Tebaldi. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International journal of climatology*, 27(12):1547–1578, 2007.

[27] J. H. Friedman. Greedy function approximation: a gradient boosting machine. *Annals of statistics*, pages 1189–1232, 2001.

[28] B. Fu. On the calculation of the evaporation from land surface. *Sci. Atmos. Sin*, 5(1):23–31, 1981.

[29] P. Gentine, P. D'Odorico, B. R. Lintner, G. Sivandran, and G. Salvucci. Interdependence of climate, soil, and vegetation as constrained by the budyko curve. *Geophysical Research Letters*, 39(19), 2012.

[30] B. George, H. Malano, B. Davidson, P. Hellegers, L. Bharati, and S. Massuel. An integrated hydro-economic modelling framework to evaluate water allocation strategies i: Model development. *Agricultural water management*, 98(5):733–746, 2011.

[31] B. George, H. Malano, B. Davidson, P. Hellegers, L. Bharati, and S. Massuel. An integrated hydro-economic modelling framework to evaluate water allocation strategies ii: Scenario assessment. *Agricultural Water Management*, 98(5):747–758, 2011.

[32] S. Ghosh and P. Mujumdar. Nonparametric methods for modeling gcm and scenario uncertainty in drought assessment. *Water Resources Research*, 43(7), 2007.

[33] S. Ghosh and P. P. Mujumdar. Statistical downscaling of gcm simulations to streamflow using relevance vector machine. *Advances in water resources*, 31(1):132–146, 2008.

[34] M. A. Giorgetta, J. Jungclaus, C. H. Reick, S. Legutke, J. Bader, M. Böttinger, V. Brovkin, T. Crueger, M. Esch, K. Fieg, et al. Climate and carbon cycle changes from 1850 to 2100 in mpi-esm simulations for the coupled model intercomparison project phase 5. *Journal of Advances in Modeling Earth Systems*, 5(3):572–597, 2013.

[35] F. Giorgi, E. Coppola, F. Solmon, L. Mariotti, M. Sylla, X. Bi, N. Elguindi, G. Diro, V. Nair, G. Giuliani, et al. Regcm4: model description and preliminary tests over multiple cordex domains. *Climate Research*, 52:7–29, 2012.

[36] B. Goswami, M. Madhusoodanan, C. Neema, and D. Sengupta. A physical mechanism for north atlantic sst influence on the indian summer monsoon. *Geophysical Research Letters*, 33(2), 2006.

[37] P. Guhathakurta. Droughts in districts of india during the recent all india normal monsoon years and its probability of occurrence. *Mausam*, 54(2):542–544, 2003.

[38] G. H. Hargreaves and Z. A. Samani. Reference crop evapotranspiration from temperature. *Applied engineering in agriculture*, 1(2):96–99, 1985.

[39] M. J. Hayes, M. D. Svoboda, D. A. Wiihite, and O. V. Vanyarkho. Monitoring the 1996 drought using the standardized precipitation index. *Bulletin of the American meteorological society*, 80(3):429–438, 1999.

[40] B. C. Hewitson and R. G. Crane. Large-scale atmospheric controls on local precipitation in tropical mexico. *Geophysical research letters*, 19(18):1835–1838, 1992.

[41] S. Huang, Q. Huang, G. Leng, and S. Liu. A nonparametric multivariate standardized drought index for characterizing socioeconomic drought: A case study in the heihe river basin. *Journal of Hydrology*, 542:875–883, 2016.

[42] J. Jarsjö, Y. Shibuo, and G. Destouni. Using the pcraster-polflow approach to gis-based modelling of coupled groundwater-surface water hydrology in the forsmark area. Technical report, Swedish Nuclear Fuel and Waste Management Co., 2004.

[43] J. Jarsjö, Y. Shibuo, and G. Destouni. Spatial distribution of unmonitored inland water discharges to the sea. *Journal of Hydrology*, 348(1-2):59–72, 2008.

[44] M. Jung, M. Reichstein, P. Ciais, S. I. Seneviratne, J. Sheffield, M. L. Goulden, G. Bonan, A. Cescatti, J. Chen, R. De Jeu, et al. Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature*, 467(7318):951, 2010.

[45] D. Kim and J. Rhee. A drought index based on actual evapotranspiration from the bouchet hypothesis. *Geophysical Research Letters*, 43(19), 2016.

[46] K. N. Kumar, M. Rajeevan, D. Pai, A. Srivastava, and B. Preethi. On the observed variability of monsoon droughts over india. *Weather and Climate Extremes*, 1:42–50, 2013.

[47] R. Kumar and H. R. Gautam. Climate change and its impact on agricultural productivity in india. *Journal of Climatology & Weather Forecasting*, 2014.

[48] H. Li, J. Sheffield, and E. F. Wood. Bias correction of monthly precipitation and temperature fields from intergovernmental panel on climate change ar4 models using equidistant quantile matching. *Journal of Geophysical Research: Atmospheres*, 115(D10), 2010.

[49] M. Liu, X. Xu, C. Xu, A. Y. Sun, K. Wang, B. R. Scanlon, and L. Zhang. A new drought index that considers the joint effects of climate and land surface change. *Water Resources Research*, 53(4):3262–3278, 2017.

[50] W. Liu, L. Wang, J. Zhou, Y. Li, F. Sun, G. Fu, X. Li, and Y.-F. Sang. A worldwide evaluation of basin-scale evapotranspiration estimates against the water balance method. *Journal of Hydrology*, 538:82–95, 2016.

[51] W.-Y. Loh. Classification and regression trees. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1(1):14–23, 2011.

[52] G. Mallya, V. Mishra, D. Niyogi, S. Tripathi, and R. S. Govindaraju. Trends and variability of droughts over the indian monsoon region. *Weather and Climate Extremes*, 12:43–68, 2016.

[53] B. Martens, D. Gonzalez Miralles, H. Lievens, R. Van Der Schalie, R. A. De Jeu, D. Fernández-Prieto, H. E. Beck, W. Dorigo, and N. Verhoest. Gleam v3: Satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development*, 10(5):1903–1925, 2017.

[54] K. McGuffie and A. Henderson-Sellers. *A climate modelling primer*. John Wiley & Sons, 2005.

[55] T. B. McKee, N. J. Doesken, J. Kleist, et al. The relationship of drought frequency and duration to time scales. In *Proceedings of the 8th Conference on Applied Climatology*, volume 17, pages 179–183. American Meteorological Society Boston, MA, 1993.

[56] R. Mehrotra and A. Sharma. A nonparametric nonhomogeneous hidden markov model for downscaling of multisite daily rainfall occurrences. *Journal of Geophysical Research: Atmospheres*, 110(D16), 2005.

[57] P. Milly. Climate, soil water storage, and the average annual water balance. *Water Resources Research*, 30(7):2143–2156, 1994.

[58] D. Miralles, T. Holmes, R. De Jeu, J. Gash, A. Meesters, A. Dolman, et al. Global land-surface evaporation estimated from satellite-based observations. *Hydrology and Earth System Sciences Discussions*, 7, 2011.

[59] V. Mishra, S. Aadhar, A. Asoka, S. Pai, and R. Kumar. On the frequency of the 2015 monsoon season drought in the indo-gangetic plain. *Geophysical Research Letters*, 43(23):12–102, 2016.

[60] V. Mishra, S. Mukherjee, R. Kumar, and D. A. Stone. Heat wave exposure in india in current, 1.5 c, and 2.0 c worlds. *Environmental Research Letters*, 12(12):124012, 2017.

[61] V. Mishra, R. Shah, and B. Thrasher. Soil moisture droughts under the retrospective and projected climate in india. *Journal of Hydrometeorology*, 15(6):2267–2292, 2014.

[62] V. Mishra, B. V. Smoliak, D. P. Lettenmaier, and J. M. Wallace. A prominent pattern of year-to-year variability in indian summer monsoon rainfall. *Proceedings of the National Academy of Sciences*, 109(19):7213–7217, 2012.

[63] Q. Mu, M. Zhao, J. S. Kimball, N. G. McDowell, and S. W. Running. A remotely sensed global terrestrial drought severity index. *Bulletin of the American Meteorological Society*, 94(1):83–98, 2013.

[64] V. Murray and K. L. Ebi. Ipcc special report on managing the risks of extreme events and disasters to advance climate change adaptation (srex), 2012.

[65] R. Nath, X. Cui, D. Nath, H. F. Graf, W. Chen, L. Wang, H. Gong, and Q. Li. Cmip5 multimodel projections of extreme weather events in the humid subtropical gangetic plain region of india. *Earth's Future*, 5(2):224–239, 2017.

[66] D. Pai, L. Sridhar, P. Guhathakurta, and H. Hatwar. District-wide drought climatology of the southwest monsoon season over india based on standardized precipitation index (spi). *Natural hazards*, 59(3):1797–1813, 2011.

[67] W. Palmer. Meteorological drought, research paper no. 45: Us weather bureau. *Office of Climatology, US Weather Bureau, Washington DC*, 1965.

[68] M. Parry, M. L. Parry, O. Canziani, J. Palutikof, P. Van der Linden, and C. Hanson. *Climate change 2007-impacts, adaptation and vulnerability: Working group II contribution to the fourth assessment report of the IPCC*, volume 4. Cambridge University Press, 2007.

[69] B. Parthasarathy, N. Sontakke, A. Monot, and D. Kothawale. Droughts/floods in the summer monsoon season over different meteorological subdivisions of india for the period 1871–1984. *Journal of Climatology*, 7(1):57–70, 1987.

[70] A. Pettitt. A non-parametric approach to the change-point problem. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 28(2):126–135, 1979.

[71] M. Pidwirny. The hydrologic cycle. *fundamentals of physical geography,*, 2006.

[72] P. Racherla, D. Shindell, and G. Faluvegi. The added value to global model projections of climate change by dynamical downscaling: A case study over the continental us using the giss-modele2 and wrf models. *Journal of Geophysical Research: Atmospheres*, 117(D20), 2012.

[73] D. Raje and P. Mujumdar. Reservoir performance under uncertainty in hydrologic impacts of climate change. *Advances in Water Resources*, 33(3):312–326, 2010.

[74] M. Rajeevan and J. Bhate. A high resolution daily gridded rainfall dataset (1971–2005) for mesoscale meteorological studies. *Current Science*, pages 558–562, 2009.

[75] K. S. Ray, M. Shewale, et al. Probability of occurrence of drought in various sub-divisions of india. *Mausam*, 52(3):541–546, 2001.

[76] S. Rehana and P. Mujumdar. Climate change induced risk in water quality control problems. *Journal of Hydrology*, 444:63–77, 2012.

[77] S. Rehana and P. Mujumdar. Regional impacts of climate change on irrigation water demands. *Hydrological Processes*, 27(20):2918–2933, 2013.

[78] S. Rehana and P. Mujumdar. Basin scale water resources systems modeling under cascading uncertainties. *Water resources management*, 28(10):3127–3142, 2014.

[79] F. Sajedi-Hosseini, A. Malekian, B. Choubin, O. Rahmati, S. Cipullo, F. Coulon, and B. Pradhan. A novel machine learning-based approach for the risk assessment of nitrate groundwater contamination. *Science of the total environment*, 644:954–962, 2018.

[80] K. Salvi and S. Ghosh. High-resolution multisite daily rainfall projections in india with statistical downscaling for climate change impacts assessment. *Journal of Geophysical Research: Atmospheres*, 118(9):3557–3578, 2013.

[81] H. L. Shah and V. Mishra. Hydrologic changes in indian subcontinental river basins (1901–2012). *Journal of Hydrometeorology*, 17(10):2667–2687, 2016.

[82] R. D. Shah and V. Mishra. Development of an experimental near-real-time drought monitor for india. *Journal of Hydrometeorology*, 16(1):327–345, 2015.

[83] J. Sheffield and E. F. Wood. Projected changes in drought occurrence under future global warming from multi-model, multi-scenario, ipcc ar4 simulations. *Climate dynamics*, 31(1):79–105, 2008.

[84] M. L. Shelton. *Hydroclimatology: perspectives and applications*. Cambridge University Press, 2009.

[85] Y. Shibuo, J. Jarsjö, and G. Destouni. Hydrological responses to climate change and irrigation in the aral sea drainage basin. *Geophysical Research Letters*, 34(21), 2007.

[86] S. Singh, S. Ghosh, A. Sahana, H. Vittal, and S. Karmakar. Do dynamic regional models add value to the global model projections of indian monsoon? *Climate dynamics*, 48(3-4):1375–1397, 2017.

[87] M. Sivapalan, M. A. Yaeger, C. J. Harman, X. Xu, and P. A. Troch. Functional model of water balance variability at the catchment scale: 1. evidence of hydrologic similarity and space-time symmetry. *Water Resources Research*, 47(2), 2011.

[88] L. Sun, H. Li, S. E. Zebiak, D. F. Moncunill, F. D. Filho, and A. D. Moura. An operational dynamical downscaling prediction system for nordeste brazil and the 2002–04 real-time forecast evaluation. *Journal of Climate*, 19(10):1990–2007, 2006.

[89] M. Svoboda, D. LeComte, M. Hayes, R. Heim, K. Gleason, J. Angel, B. Rippey, R. Tinker, M. Palecki, D. Stooksbury, et al. The drought monitor. *Bulletin of the American Meteorological Society*, 83(8):1181–1190, 2002.

[90] C. Teutschbein, F. Wetterhall, and J. Seibert. Evaluation of different downscaling techniques for hydrological climate-change impact studies at the catchment scale. *Climate Dynamics*, 37(9-10):2087–2105, 2011.

[91] S. E. Thompson, C. J. Harman, P. A. Troch, P. D. Brooks, and M. Sivapalan. Spatial scale dependence of ecohydrologically mediated water balance partitioning: A synthesis framework for catchment ecohydrology. *Water Resources Research*, 47(10), 2011.

[92] C. W. Thornthwaite. An approach toward a rational classification of climate. *Geographical review*, 38(1):55–94, 1948.

[93] S. Tripathi, V. Srinivas, and R. S. Nanjundiah. Downscaling of precipitation for climate change scenarios: a support vector machine approach. *Journal of hydrology*, 330(3-4):621–640, 2006.

[94] L. Turc. The water balance of soils. relation between precipitation, evaporation and flow. *Ann. Agron*, 5:491–569, 1954.

[95] G. Van der Schrier, P. Jones, and K. Briffa. The sensitivity of the pdsi to the thornthwaite and penman-monteith parameterizations for potential evapotranspiration. *Journal of Geophysical Research: Atmospheres*, 116(D3), 2011.

[96] H. Van Lanen, N. Wanders, L. Tallaksen, and A. Van Loon. Hydrological drought across the world: impact of climate and physical catchment structure. *Hydrology and Earth System Sciences*, 17:1715–1732, 2013.

[97] A. F. Van Loon. Hydrological drought explained. *Wiley Interdisciplinary Reviews: Water*, 2(4):359–392, 2015.

[98] J.-P. Venot. Drawing water for thirsty lands stories of the closing krishna river basin in south india. *Oral presentation, Water availability, use and management*, 2008.

[99] S. M. Vicente-Serrano, S. Beguería, and J. I. López-Moreno. A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of climate*, 23(7):1696–1718, 2010.

[100] S. M. Vicente-Serrano, S. Beguería, J. I. López-Moreno, M. Angulo, and A. El Kenawy. A new global 0.5 gridded dataset (1901–2006) of a multiscalar drought index: comparison with current drought index datasets based on the palmer drought severity index. *Journal of Hydrometeorology*, 11(4):1033–1043, 2010.

[101] S. M. Vicente-Serrano, G. Van der Schrier, S. Beguería, C. Azorin-Molina, and J.-I. Lopez-Moreno. Contribution of precipitation and reference evapotranspiration to drought indices under different climates. *Journal of Hydrology*, 526:42–54, 2015.

[102] D. Wang and Y. Tang. A one-parameter budyko model for water balance captures emergent behavior in darwinian hydrologic models. *Geophysical Research Letters*, 41(13):4569–4577, 2014.

[103] R. L. Wilby, S. Charles, E. Zorita, B. Timbal, P. Whetton, and L. Mearns. Guidelines for use of climate scenarios developed from statistical downscaling methods. *Supporting material of the Intergovernmental Panel on Climate Change, available from the DDC of IPCC TGCIA*, 27, 2004.

[104] D. A. Wilhite and M. H. Glantz. Understanding: the drought phenomenon: the role of definitions. *Water international*, 10(3):111–120, 1985.

[105] X. Yang, B. Yong, L. Ren, Y. Zhang, and D. Long. Multi-scale validation of gleam evapotranspiration products over china via chinaflux et measurements. *International Journal of Remote Sensing*, 38(20):5688–5709, 2017.

[106] K. Zhang, J. S. Kimball, R. R. Nemani, and S. W. Running. A continuous satellite-derived global record of land surface evapotranspiration from 1983 to 2006. *Water Resources Research*, 46(9), 2010.

[107] L. Zhang, K. Hickel, W. Dawes, F. H. Chiew, A. Western, and P. Briggs. A rational function approach for estimating mean annual evapotranspiration. *Water Resources Research*, 40(2), 2004.

[108] L. Zhang, N. Potter, K. Hickel, Y. Zhang, and Q. Shao. Water balance modeling over variable time scales based on the budyko framework–model development and testing. *Journal of Hydrology*, 360(1-4):117–131, 2008.

[109] S. Zhou, B. Yu, Y. Huang, and G. Wang. The complementary relationship and generation of the budyko functions. *Geophysical Research Letters*, 42(6):1781–1790, 2015.

[110] Y. Zhu, W. Wang, V. P. Singh, and Y. Liu. Combined use of meteorological drought indices at multi-time scales for improving hydrological drought detection. *Science of the Total Environment*, 571:1058–1068, 2016.

Appendix A

Data Preprocessing for running a distributed Hydrological Model

For running any hydrological model, we need to know some of the basin properties such as extent, topography and boundary. A watershed/basin is an area of land that drains all the streams and rainfall to a common outlet such as the outflow of a reservoir,sea etc. Here we explain the process of creating a mask file and fraction file using QGIS which is to be used for preparing other inputs for distributed hydrological models. A mask file is a binary grid file and contains information which shows whether the grid cell of a particular resolution and extent is present in the basin(indicated with 1) or not (indicated with 0) whereas the fraction file represents the area of the grid falling inside the basin. These files in turn can be used to perform various hydrologic estimations.

- Load the basin shapefile into QGIS using the Layer ⟶ Add vector Layer option.

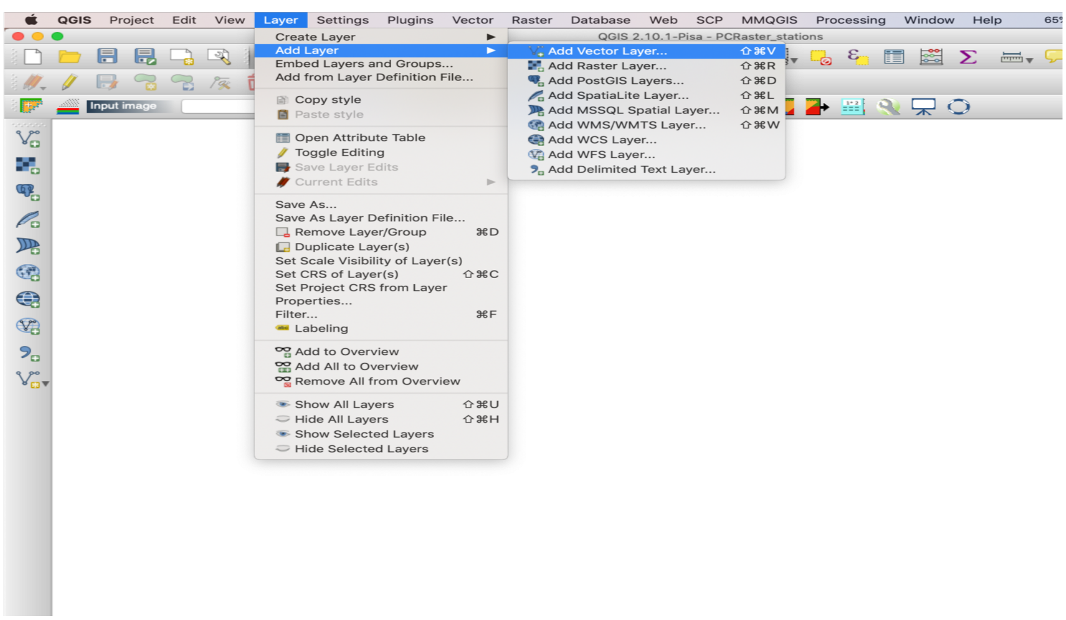


Figure A.1: Adding a vector layer in QGIS

- Now create a vector grid overlapping the basin shapefile using the Vector ⟶ Research Tools ⟶ Vector Grid. Fill the extent and resolution details as shown in the window in Figure A.3. This will be dependent on the selected basin. We are doing the analysis at 0.25°resolution. In the bottom right section set projection to EPSG:4326. This is the standard WGS84 projection.

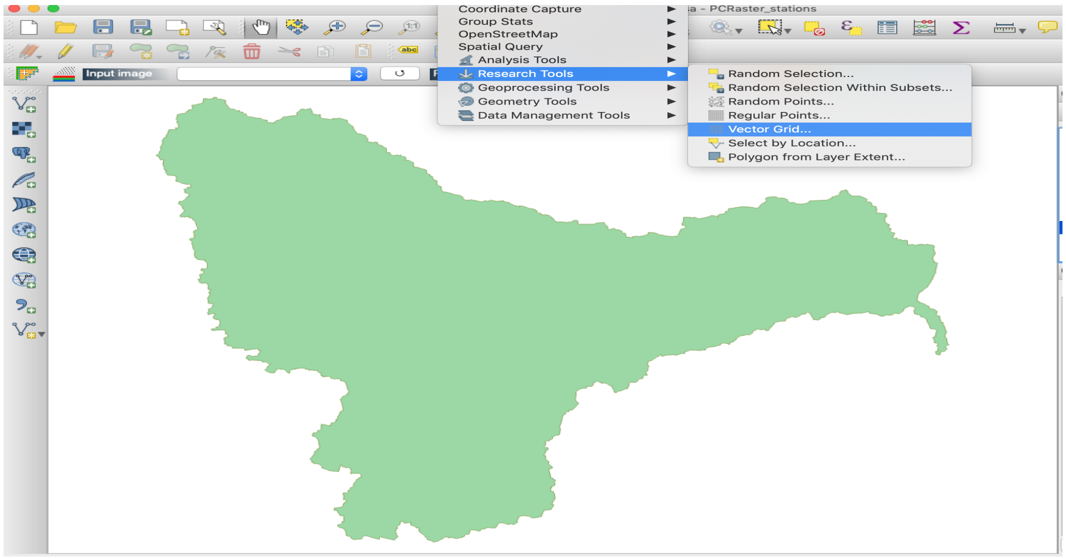


Figure A.2: Creating a vector grid

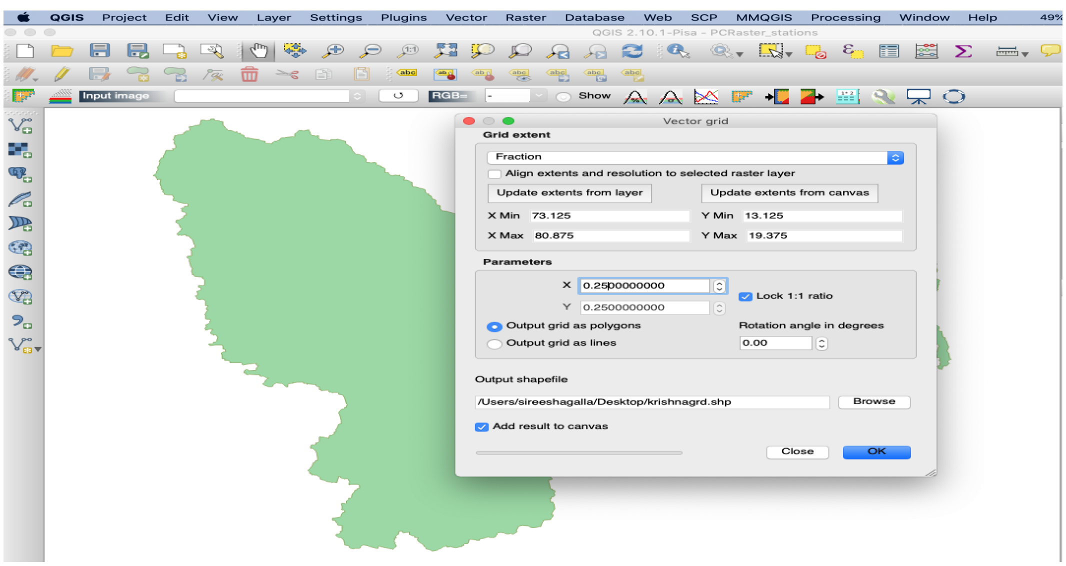


Figure A.3: Setting options for vector grid

- The vector grid is created as shown in Figure A.4.
- Now find the intersecting portion of the grid and shape file as shown in Figure A.5. This is used to calculate the fraction of the grid present in the basin.

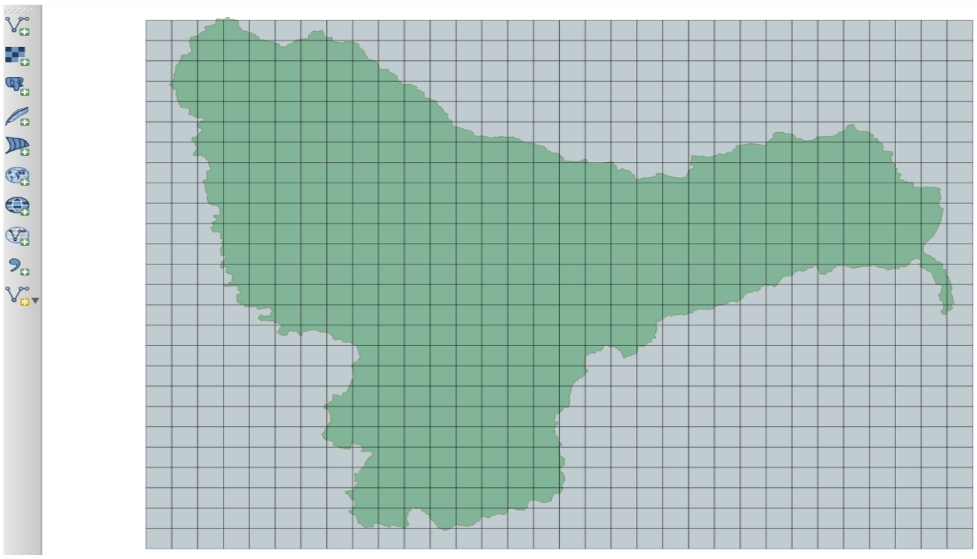


Figure A.4: Vector grid at 0.25°for Krishna basin

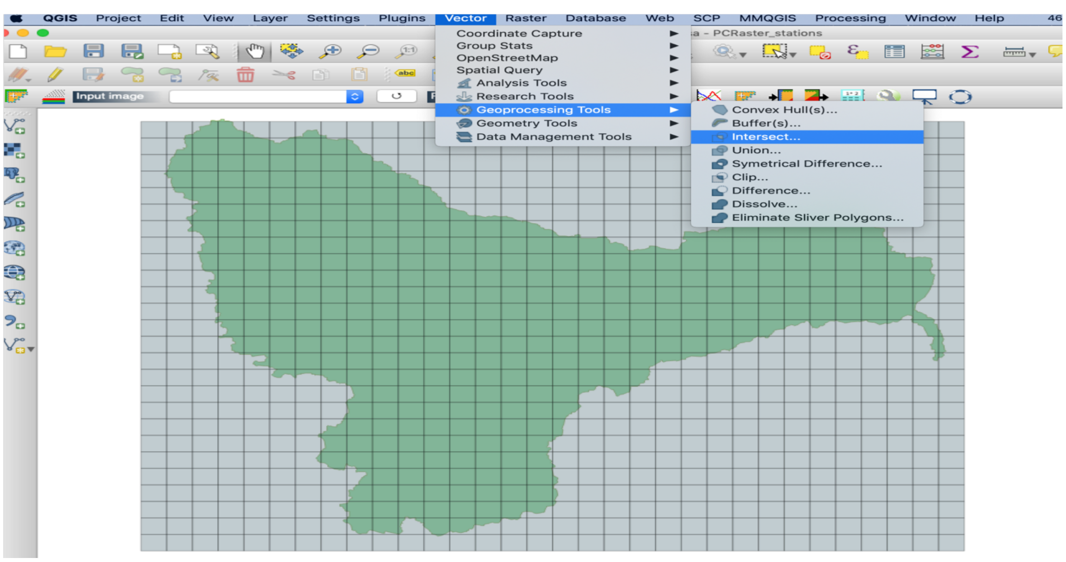


Figure A.5: Intersecting two vector layers(shapefiles)

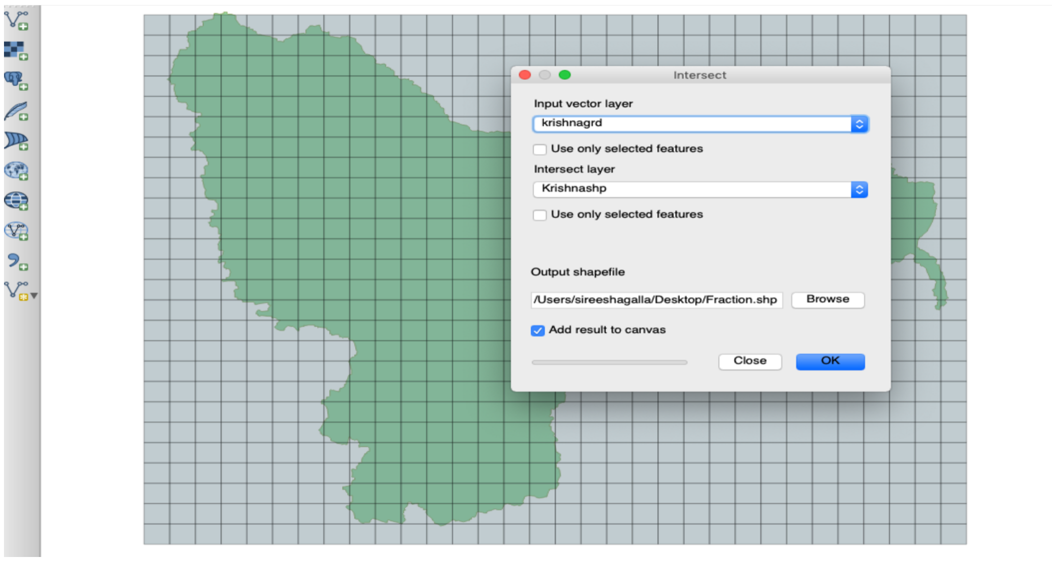


Figure A.6: Intersecting two vector layers(shapefiles)

- Open the attribute table for this intersecting layer and add a new field which contains the fraction of grid cell present in basin. Enter the field values as shown in Figure A.9. $area function in geometry section gives the area of the grid cell in sq.deg(depends on the projection used). As we want the fraction of the cell present in a 0.25°* 0.25°grid, we need to divide this cell area by a single grid area.

- The final fraction of grid cell appears as a column in the attribute table as shown in Figure A.10. We need to rasterize the basin vector shapefile using this column.

- The steps to convert a vector file into a raster based on an attribute is shown in the Figures A.11 and A.12.

- This can be converted into an ascii fraction file to view the fraction values at each grid and create a mask file using the steps shown in the Figure A.14 and the Figure A.15.

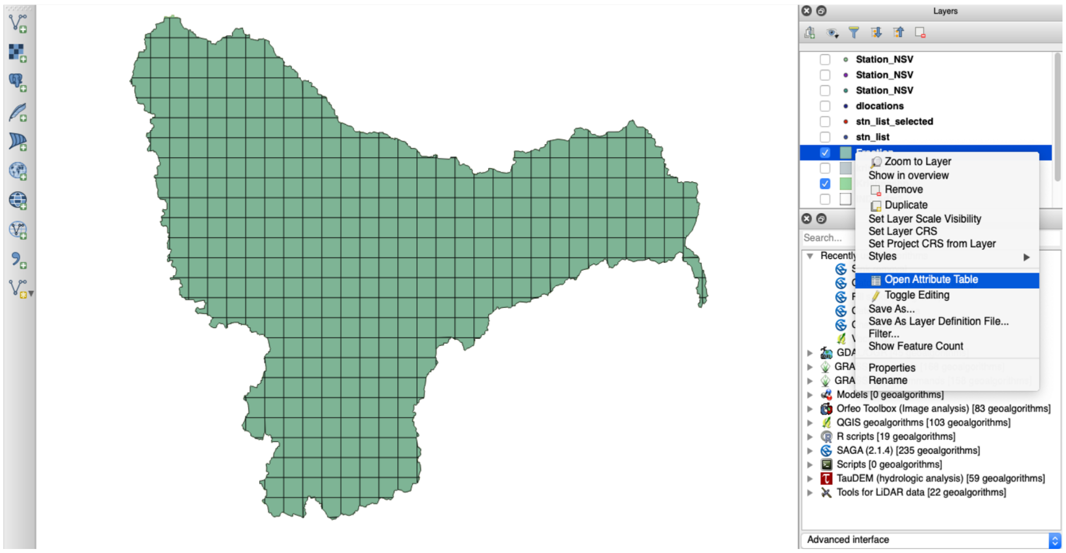


Figure A.7: View attribute table

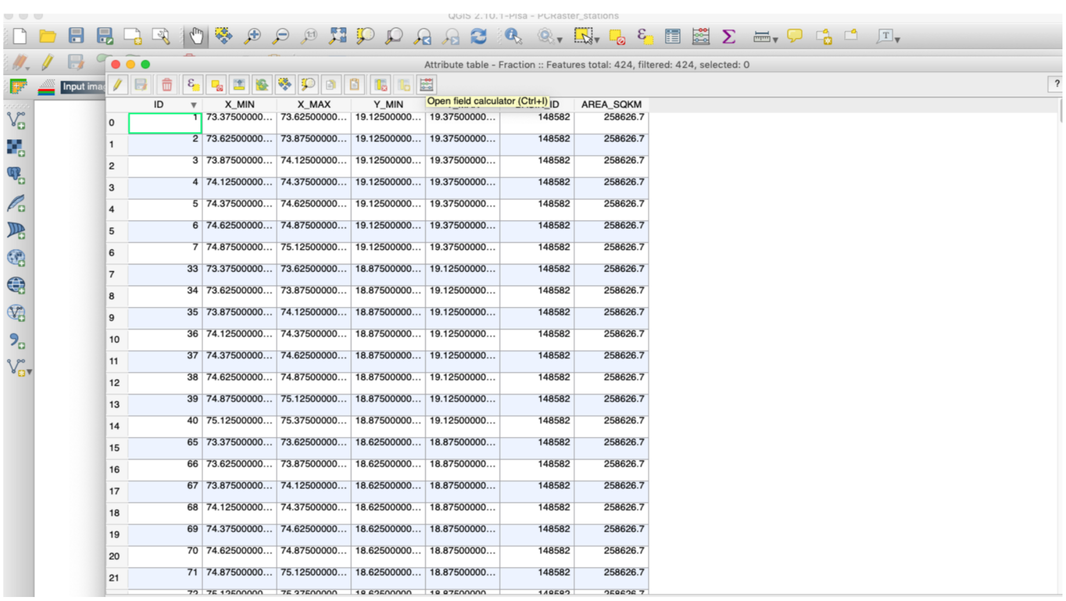


Figure A.8: Create field in attribute table

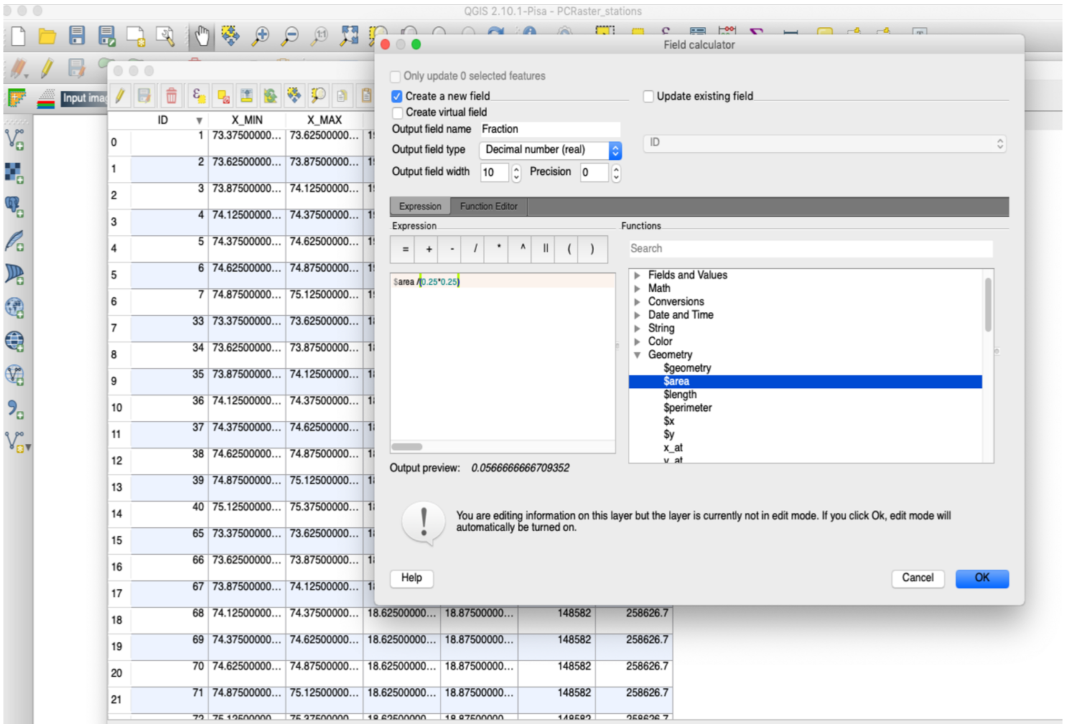


Figure A.9: Create field in attribute table

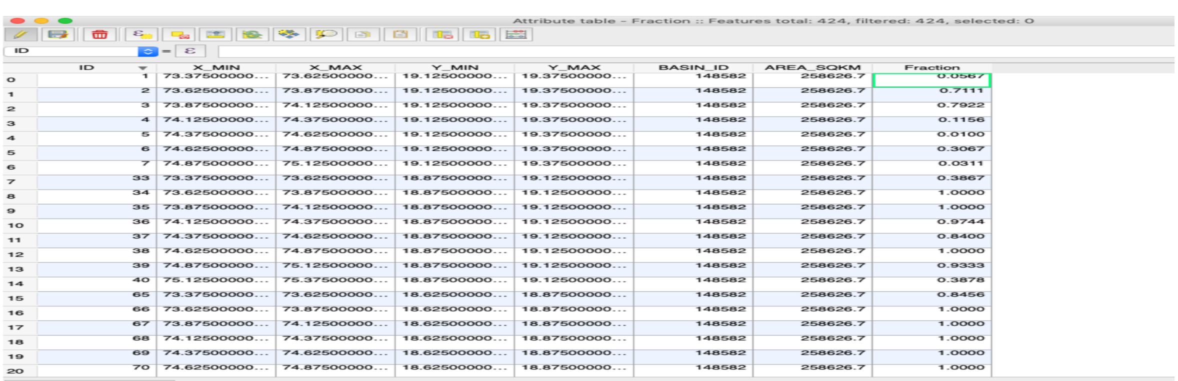


Figure A.10: Create field in attribute table

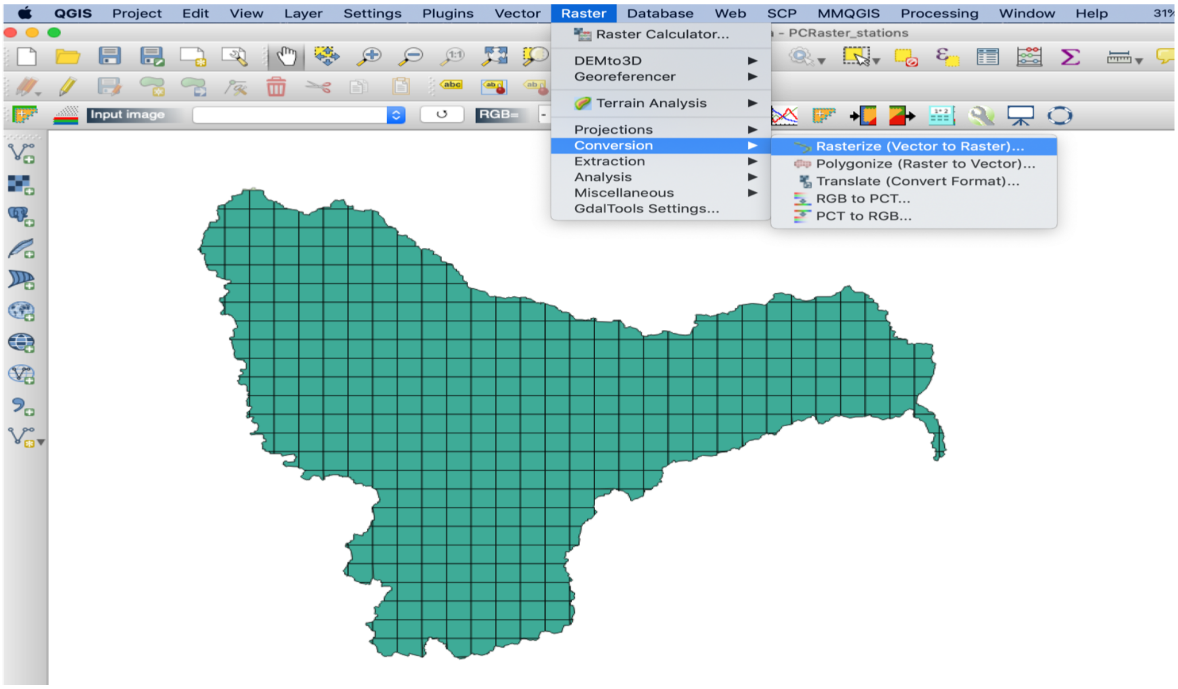


Figure A.11: Rasterize a vector shapefile

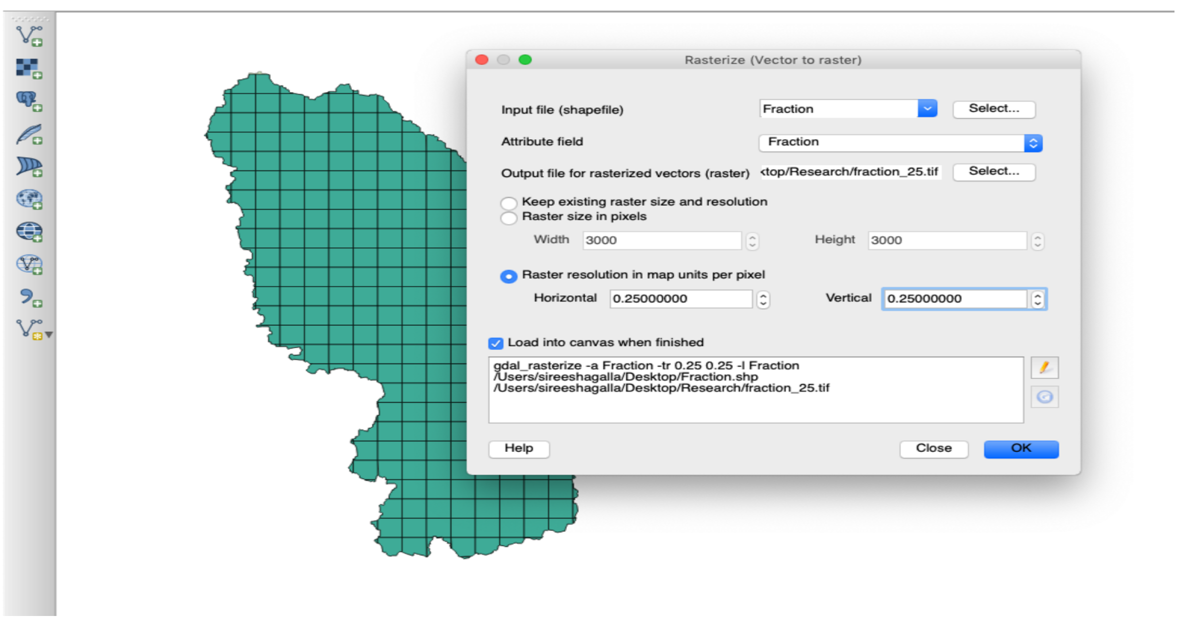


Figure A.12: Rasterize a vector shapefile - Options

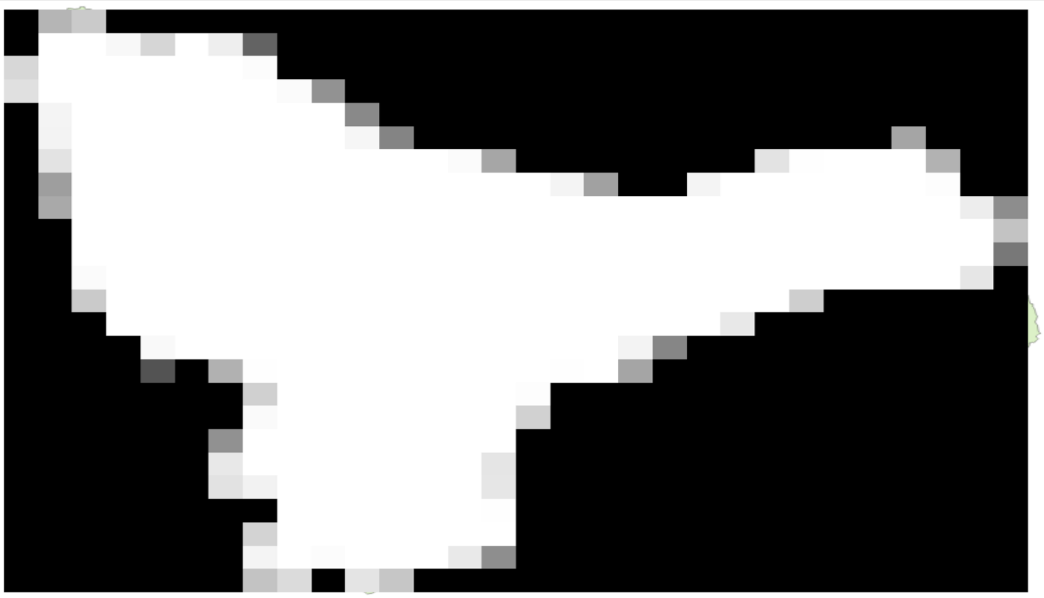


Figure A.13: Rasterized Fraction file for Krishna basin

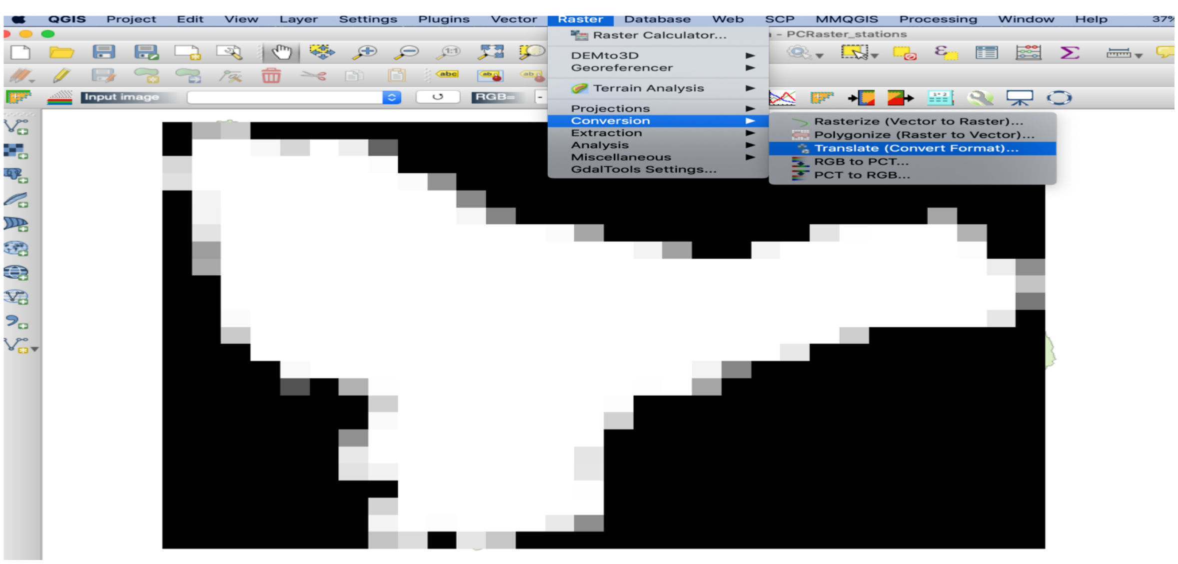


Figure A.14: Rasterized Fraction file for Krishna basin

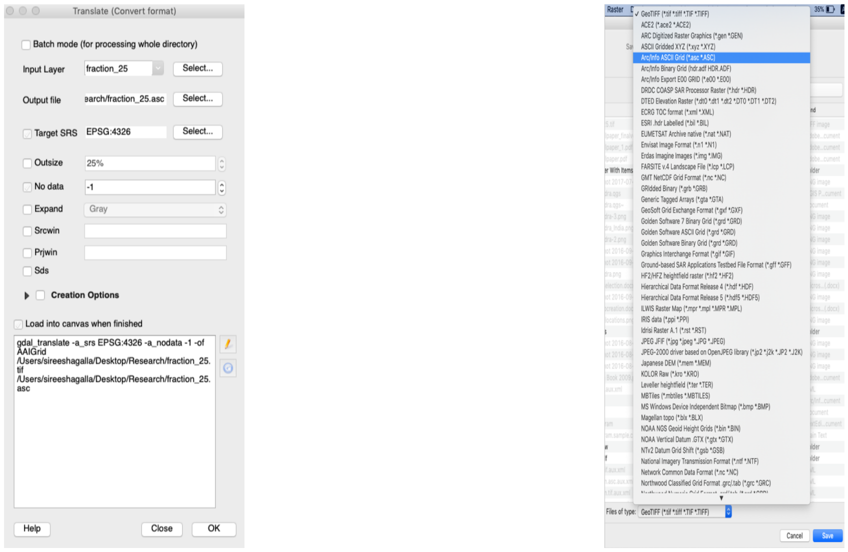
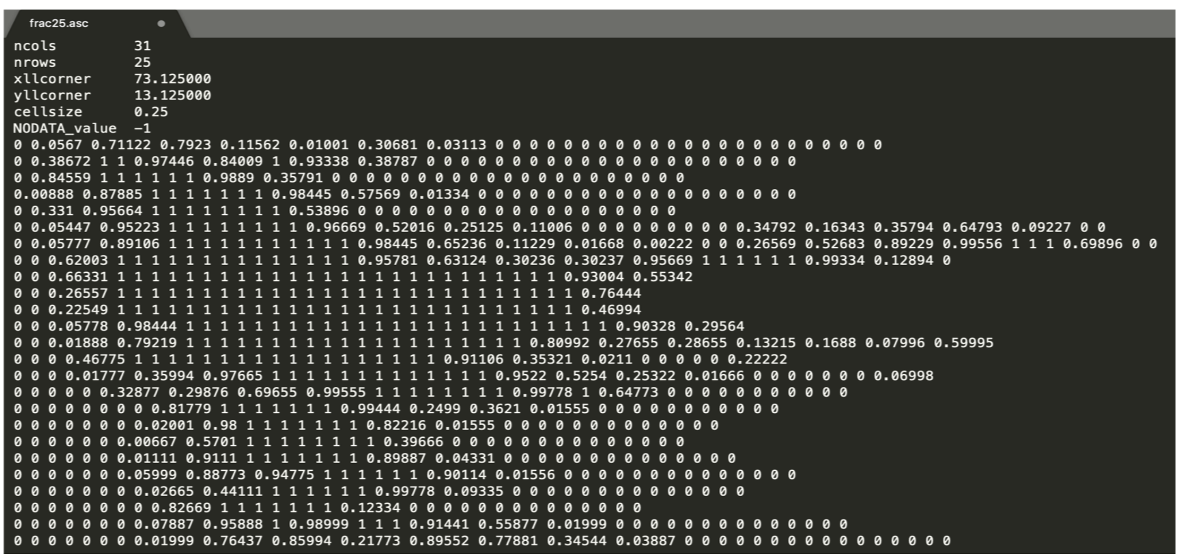


Figure A.15: Raster format conversion to ascii - Options

- Open the fraction file and observe the values. The first 6 lines contain the extent information whereas the fraction values are present from the next lines as shown in Figure A.16.



```
frac25.asc
ncols        31
nrows        25
xllcorner    73.125000
yllcorner    13.125000
cellsize     0.25
NODATA_value  -1
0 0.0567 0.71122 0.7923 0.11562 0.01001 0.30681 0.03113 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0.38672 1 1 0.97446 0.84009 1 0.93338 0.38787 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0.84559 1 1 1 1 1 1 0.9889 0.35791 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.00888 0.87885 1 1 1 1 1 1 1 0.98445 0.57569 0.01334 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0.331 0.95664 1 1 1 1 1 1 1 1 0.53896 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0.05447 0.95223 1 1 1 1 1 1 1 1 0.96669 0.52016 0.25125 0.11006 0 0 0 0 0 0 0 0 0 0.34792 0.16343 0.35794 0.64793 0.09227 0 0
0 0.05777 0.89106 1 1 1 1 1 1 1 1 1 1 1 0.98445 0.65236 0.11229 0.01668 0.00222 0 0 0.26569 0.52683 0.89229 0.99556 1 1 1 0.69896 0 0
0 0 0.62003 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0.95781 0.63124 0.30236 0.30237 0.95669 1 1 1 1 1 0.99334 0.12894 0
0 0 0.66331 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0.93004 0.55342
0 0 0.26557 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0.76444
0 0 0.22549 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0.46994
0 0 0.05778 0.98444 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0.90328 0.29564
0 0 0.01888 0.79219 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0.80992 0.27655 0.28655 0.13215 0.1688 0.07996 0.59995
0 0 0 0.46775 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0.91106 0.35321 0.0211 0 0 0 0 0 0.22222
0 0 0 0.01777 0.35994 0.97665 1 1 1 1 1 1 1 1 1 1 1 1 1 0.9522 0.5254 0.25322 0.01666 0 0 0 0 0 0 0 0.06998
0 0 0 0 0 0.32877 0.29876 0.69655 0.99555 1 1 1 1 1 1 1 1 0.99778 1 0.64773 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0.81779 1 1 1 1 1 1 1 0.99444 0.2499 0.3621 0.01555 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0.02001 0.98 1 1 1 1 1 1 1 0.82216 0.01555 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0.00667 0.5701 1 1 1 1 1 1 1 1 0.39666 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0.01111 0.9111 1 1 1 1 1 1 1 0.89887 0.04331 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0.05999 0.88773 0.94775 1 1 1 1 1 1 0.90114 0.01556 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0.02665 0.44111 1 1 1 1 1 1 0.99778 0.09335 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0.82669 1 1 1 1 1 1 1 0.12334 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0.07887 0.95888 1 0.98999 1 1 1 0.91441 0.55877 0.01999 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0.01999 0.76437 0.85994 0.21773 0.89552 0.77881 0.34544 0.03887 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

Figure A.16: Fraction file in ascii format

- Now to create a mask file read this fraction file in MATLAB using import data or dlmread option and pick grids whose fraction is greater than 0.25(25%) or 0.5(50%) depending on the criteria required. Script to create a mask file from ascii file is shown in the Figure A.17.

```matlab
frac = dlmread('frac25.asc',' ',6,0);   % Put the path of fraction file
[no_lats,no_longs] = size(frac);   % Find the size of basin to create mask
mask = zeros(no_lats,no_longs); % Initialise mask with zeros
% When fraction is greater than 0.5, then set mask value for that grid to 1.
% This implies we are selecting grids which have atleast 50% of their area
% in the basin
for p=1:no_lats
    for q=1:no_longs
        if(frac(p,q)>=0.5)
            mask(p,q) = 1;
        end
    end
end
save('mask_25.mat','mask')
```

Figure A.17: Script to create mask file from ascii file

Appendix B

PCRaster Hydrological Modelling - Code

```
Model for simulation of runoff at Outlet using Water Balance Equation
64 timesteps of 1 year each i.e modelling time (64 years)

#!--matrixtable
#!--lddin
#binding
RainStations=rainstationsGrids.map;    #map with location of rainstations
RainTimeSeries=RainfallAnnualKrishna1951-2014.tss;   #timeseries with
rain at rainstations
RainZones=rainzoneGrids.map;      #Grids map
SurfaceWater=rain;    #reported stack of maps with rain (mm/year)
AETTimeSeries=AETAnnualKrishna1951-2014.tss;
AETClimate=AET;    #reported stack of maps with AET Budyko (mm/year)

Dem=kbDem.map;         #digital elevation map
Ldd=ldd.map;           #reported local drain direction map
ConvConst= 1;          #conversion mm/year
SamplePlaces=samples.map;     #map with runoff sampling locations
RunoffTimeSeries=runoffAccufluxBudyko.tss; #reported timeseries with runoff
at sampling locations

timer
```

```
 1  64 1;  #1951-2014
initial
# coverage of meteorological stations for the whole area
RainZones=spreadzone(RainStations,0,1);
# generate the local drain direction map on basis of the elevation map
report Ldd=lddcreate(Dem,1e31,1e31,1e31,1e31); ˜


dynamic
#calculate maps with rainfall and AET at each timestep (mm/day)
SurfaceWater=timeinputscalar(RainTimeSeries,RainZones);
AETClimate=timeinputscalar(AETTimeSeries,RainZones);
#Calculate Runoff at each grid using the Water Balance Equation
RunoffVal=(SurfaceWater-AETClimate);
#Calculate the net runoff at each grid after accumulation
RunoffPerGrid=accuthresholdflux(Ldd,RunoffVal,0);
report RunoffTimeSeries=timeoutput(SamplePlaces,RunoffPerGrid) ˜
```

Variables are bound to the files using the binding option and initialised using the initial option. The dynamic option is used for the variables which have to be calculated in a loop using the iterations mentioned in the timer option.

Appendix C

Machine Learning Algorithms used in Statistical Downscaling

C.1 Principal Component Analysis(PCA)

PCA is applied to the standardized and bias corrected data in order to reduce the computational problems of multidimensionality as well as multicollinearity. The following paragraph contains the PCA algorithm explained in detail.

- Considering X as the predictor variables, the mean(μ) of the 2-dimensional data matrix 'X' is calculated and subtracted from each time-step across all the predictors.

  $$\mu = mean(X);$$

  $$t = X - \mu; \tag{C.1}$$

  The GCM data used is already standardized and bias corrected, hence this step might not be required for this data.

- Calculate co-variance matrix 'S' from the standardized data.

  $$S = t * t'; \tag{C.2}$$

  Where t is the standardized data from C.1 and t^T is the transpose of t.

- The Eigen vectors and Eigen values of the co-variance matrix 'S' are estimated and the first 'k' components preserving atleast 98% of the information along the direction of maximum variance are picked.

  $$[VD] = eig(S); \tag{C.3}$$

  $$v_1 = V(:,1); v_2 = V(:,2) - - - - - - v_k = V(:,k);$$

  where $v_1, v_2 - - - v_k$ are the 'k' principal components.

C.2 K-means Clustering

K-means clustering is a clustering technique that divides the 'n' observations into K clusters for which each observation belongs to the cluster with the closest mean. This technique reads the observed rainfall values for all grids in the basin in a month, clusters them, and provides a single representative value that is referred to as the state of rainfall for that particular month. The step is important in the sense that it provides the rainfall category for a particular month, and can be linked to the predictors for establishing the statistical relationship.

K-means Algorithm

- Consider X = $x1, x2, x3, ..., xn$ as the set of data points and V = μ_1, μ_2, μ_3 as the set of cluster means.
- Initialize k(3) cluster means μ_1, μ_2, μ_3 based on $25^{th}, 50^{th}$ and 75^{th} percentile value of rainfall dara.
- For each data point, the nearest mean is identified and the point is assigned to the corresponding cluster. The cluster mean is recalculated using the equation C.4

$$\mu_i = \frac{1}{c_i} \sum_{j=1}^{c_i} x_i \tag{C.4}$$

  where 'c_i' represents the number of data points in i^{th} cluster.

- The procedure is repeated until the cluster means are no longer changing by a threshold or on reaching the maximum number of iterations.

C.3 Classification and Regression Trees(CART)

In Classification and Regression Trees(CART), each principal component is given a weight which when applied collectively to the principal components in a month, gives the rainfall state for that month. The weight for each principal component is obtained by passing an observation with a missing split attribute value down every level, where the weight is proportional to the number of observations non-missing that attribute in the level.

- For every iteration, the Gini's impurity/ diversity Index [51] of the data is calculated using the formula given in the equation C.5

$$E = \sum_{i \neq j} P(w_i)P(w_j) = 1 - \sum_j P(w_j)^2 \tag{C.5}$$

  where $P(w_j)$ represents the probability of data being in j^{th} class for each value of the attribute.

- An attribute splitting that minimizes the decrease in this impurity(E) as much as possible is chosen. Ideally, E is 0 when all the patterns at node have the same class label.
- Continue until impurity(E) is less than a certain threshold(η) or on reaching the maximum number of iterations.

$$E < \eta$$

C.4 Support Vector Regression(SVR)

SVR tries to minimize the error of misclassification calculated using the formula C.6 in addition to maximizing the margin 'b' between the separating hyperplane and the support vectors. SVR using different formulations and kernels is explained in detail.

Linear SVR

Linear SVR finds a model w such that $w^T x$ + a constant 'b' is close to the target value 'y'. In mathematical terms, the objective can be written as

- **Primal Formula**

  Minimize

$$\frac{1}{2}||w||^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \tag{C.6}$$

  satisfying the set of constraints mentioned in the equation C.7

$$\begin{aligned} y_i - wx_i + b &\leq \varepsilon + \xi_i \\ wx_i + b - y_i &\leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* &\geq 0 \end{aligned} \tag{C.7}$$

  where x is the predictor data, w the set of weights and b is the bias of data, the amount of predictand(y) can be written as y=wx+b, i is the i^{th} data point and ξ_i, ξ_i^* are slack variables.

- **Dual Formula**

$$y = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)\langle x_i, x\rangle + b \tag{C.8}$$

  where $\langle x_i, x\rangle$ stands for the inner product of x_i and x.

Non-linear SVR - Kernel

Some regression problems cannot adequately be described using a linear model. In such cases, Kernel transformation is applied to transform the data into a higher dimensional space where it can be linearly separated. Here Linear kernel, Radial Basis Function(RBF) (or) Gaussian Kernel and Polynomial Kernel have been compared. It has been observed that RBF and Polynomial kernels were overfitting the data, whereas Linear kernel obtaining a good balance for the training and validation data, was selected for regression analysis.

- **Primal Formula**

$$y = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*)(\varphi x_i, \varphi x) + b$$

where $\varphi(x)$ is the transformation that maps x to a high-dimensional space.

Kernel Functions

  – Linear Kernel

$$K(x_i, x_j) = (x_i.x_j) + c$$

where $x_1.x_2$ stands for the inner product of x_1 and x_2, c is a constant

  – Polynomial Kernel

$$K(x_i, x_j) = (x_i.x_j)^d$$

where d is a power constant applied to inner product.

  – RBF Kernel

$$K(x_i, x_j) = exp\left(- \frac{||x_i - x_j||^2}{2\sigma^2} \right)$$

- **Dual Formula**

$$y = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) K(x_i, x) + b$$