FORECASTING OF RAINFALL IN METEOROLOGICAL SUBDIVISIONS OF KARNATAKA USING NON-LINEAR STATISTICAL MODELS

KODANDARAMA, S. R.

PALB 8192

DEPARTMENT OF AGRICULTURAL STATISTICS, APPLIED MATHEMATICS AND COMPUTER SCIENCE
UNIVERSITY OF AGRICULTURAL SCIENCES
BANGALORE

2020

FORECASTING OF RAINFALL IN METEOROLOGICAL SUBDIVISIONS OF KARNATAKA USING NON-LINEAR STATISTICAL MODELS

KODANDARAMA, S. R.

PALB 8192

Thesis submitted to the

UNIVERSITY OF AGRICULTURAL SCIENCES, BANGALORE

In partial fulfillment of the requirements

for the award of the Degree of

MASTER OF SCIENCE (Agriculture)

in

AGRICULTURAL STATISTICS

BENGALURU **NOVEMBER, 2020**


Wholeheartedly
Dedicated to My
Parents, Brother and
My Guide

DEPARTMENT OF AGRICULTURAL STATISTICS,
APPLIED MATHEMATICS AND COMPUTER SCIENCE
UNIVERSITY OF AGRICULTURAL SCIENCES, BANGALORE

CERTIFICATE

This is to certify that the thesis entitled **"FORECASTING OF RAINFALL IN METEOROLOGICAL SUBDIVISIONS OF KARNATAKA USING NON-LINEAR STATISTICAL MODELS"** submitted by **Mr. KODANDARAMA, S. R., ID No. PALB 8192** in partial fulfilment of the requirement for the degree of **MASTER OF SCIENCE (Agriculture)** in **AGRICULTURAL STATISTICS** to the University of Agricultural Sciences, Bangalore is recorded of *bona fide* research work done by him during the period of his study in this University under my guidance and supervision and no part of the thesis has been submitted for the award of any other degree, diploma, associateship, fellowship or any other similar titles.

BENGALURU
NOVEMBER, 2020

(MOHAN KUMAR, T. L.)
MAJOR ADVISOR

APPROVED BY:

Chairperson: _____
(MOHAN KUMAR, T. L.)

Members: 1. _____
(D. M. GOWDA)

2. _____
(H. S. SHIVARAMU)

3. _____
(T. V. KRISHNA)

ACKNOWLEDGEMENT

First and foremost, I would to like to express my deepest sense of gratitude to the almighty "GOD" for blessing me enough patience, endurance and strength to overcome all the hurdles that crossed my path in the accomplishment of this endeavor.

I would like to extend my whole hearted gratitude and sincere thanks to **Dr. Mohan Kumar, T. L.,** *Assistant Professor, Agricultural Statistics, UAS, GKVK, Bengaluru and chairperson of my advisory committee for his constant encouragement, valuable suggestions, affectionate behavior, useful discussions and constructive criticisms during the course of investigation and preparation of the manuscript.*

I avail myself of this opportunity to convey my heartfelt thanks to my Advisory Committee members **Dr. D. M. Gowda,** *Retired Professor, Agricultural Statistics, UAS, GKVK, Bengaluru,* **Dr. H. S. Shivaramu,** *Professor and Head, AICRP on Agrometeorology, UAS, GKVK, Bengaluru and* **Dr. T. V. Krishna,** *Assistant Professor (GPB), AINP on Arid Legumes, ZARS, UAS, GKVK, Bengaluru for providing necessary facilities, valuable help and encouragement during the course of this investigation.*

I wish to express my sincere gratitude to **Mr. V. Manjunath,** *Associate Professor, Department of Agricultural Statistics,* **Dr. K. N. Krishnamurthy,** *Professor and Head, Department of Agricultural Statistics,* **Dr. S. N. Megeri,** *Retired Professor, Department of Agricultural Statistics,* **Dr. K. B. Murthy,** *Professor, Department of Agricultural Statistics,* **Ms. Vimala, M.,** *Assistant Professor, Department of Agricultural Statistics,* **Mr. Rajesh, A.,** *Contract Teacher, Department of Agricultural Statistics,* **Dr. Yogesh,** *Assistant Professor, Department of Agricultural Statistics,* **Ms. Latha,** *Contract Teacher, Department of Agricultural Statistics, and* **Dr. M. Gopinath Rao,** *Retired Professor, Department of Agricultural Statistics, UAS, GKVK, Bangalore for providing me the necessary guidance and encouragement throughout this work.*

Words fail me to express my love towards my Anna **Rudrappa** *and Avva* **Sowbhagy** *for bringing me up in the best of ways and for all their boundless affection, moral support, eternal love, deep concern, prayers and personal sacrifices, which sustains peace in my life. This thesis is dedicated to my Tamma* **Dhananjaiah.**

I express my sincere thanks to my seniors **Namratha, Vinay, Raaga, Harsha, Govind, Nagaraj, Vaibhav, Prabath, Aparna, Bhagya, Kalai,** *and* **Sushma** *for their valuable guidelines.*

I am thankful to all my classmates **Vinay, Avinash, Joydeep, Prem, Madhu, Veershetty, Megha, Varalakshmi, Nanditha, Kavitha** *and* **Jhanavi** *for their priceless co-operation, constant help and encouragement.*

I thank my junior friends **Harish, Rajesh, Channabasava, Karthik, Harshith, Ravi, Suman, Apoorva, Sahana, Anusha, Vimala, Bhuvana, Madhurya** *and* **Divya**.

I thank the non-teaching faculty **Miss. Sudha, Mrs. Mallika Banu** *and* **Mr. Manjunath** *of Department of Agricultural Statistics, UAS, GKVK, Bangalore.*

I express my sincere gratitude to University of Agricultural Sciences, Bengaluru for providing an opportunity for completing my master degree programme.

Finally, I would like to thank everybody who was imperative to the successful realization of this thesis, as well as articulate my apology that I could not mention personally one by one.

Bengaluru (Kodandarama, S. R.)

November, 2020

FORECASTING OF RAINFALL IN METEOROLOGICAL SUBDIVISIONS OF KARNATAKA USING NON-LINEAR STATISTICAL MODELS

KODANDARAMA, S. R.

ABSTRACT

Indian agriculture is mainly dependent on the timely arrival of rainfall and its distribution pattern. The analysis of rainfall distribution pattern and its forecast were essential for planning and management of water resources for agriculture and allied activities. Therefore, the present study was undertaken to analyze the trend, shifting pattern and forecasting of rainfall in four meteorological subdivisions of Karnataka namely North Interior Karnataka (NIK), South Interior Karnataka (SIK), Malnad and Coastal subdivisions using sixty years of monthly rainfall data (1960- 2019) collected from AICRP (Agro-Meteorology), GKVK and KSNDMC, Bengaluru. To analyze the trend in rainfall, Mann-Kendal and Modified Mann-Kendall tests were employed. For annual rainfall data, results of Mann-Kendal test revealed that no significant trend in all the subdivisions. However, Modified Mann-Kendall test showed monotonic increasing trend in SIK (1.47) and Coastal (5.10) subdivisions, monotonic decreasing trend in NIK (-1.40) subdivision, and no monotonic trend in Malnad subdivision. Likelihood Ratio test was used to assess the shifting pattern. Results revealed that NIK subdivision had decreased rainfall distribution after shifting year 2010 whereas, in SIK, Malnad and Coastal subdivisions increased rainfall distribution was observed after shifting years 1968, 2016 and 1966 respectively. To forecast monthly rainfall, H-WES, SARIMA, ARCH, GARCH and ANN time-series models were employed. For all the subdivisions, the ANN model was performed better than other models on both training and testing data on the basis of lowest RMSE value. Hence, ANN model can be used for forecasting monthly rainfall data in all the meteorological subdivisions of Karnataka.

November, 2020

Department of Agricultural Statistics,
Applied Mathematics and Computer Science
UAS, GKVK, Bangalore.

MOHAN KUMAR, T. L.
Major Advisor

ಕರ್ನಾಟಕದ ಹವಾಮಾನ ಉಪವಿಭಾಗಗಳಲ್ಲಿ ರೇಖಾತ್ಮಕವಲ್ಲದ ಸಂಖ್ಯಾಶಾಸ್ತ್ರೀಯ ಮಾದರಿಗಳನ್ನು ಬಳಸಿಕೊಂಡು ಮಳೆಯ ಮುನ್ಸೂಚನೆ ನೀಡುವುದು

ಕೋದಂಡರಾಮ, ಎಸ್. ಆರ್.

ಸಾರಾಂಶ

ಭಾರತದಲ್ಲಿ ಕೃಷಿಯು ಮುಖ್ಯವಾಗಿ ಮಳೆಯ ಸಮಯೋಚಿತ ಆಗಮನ ಹಾಗೂ ಅದರ ವಿತರಣಾ ಮಾದರಿಯನ್ನು ಅವಲಂಬಿಸಿದೆ. ಮಳೆ ವಿತರಣಾ ಮಾದರಿಯ ವಿಶ್ಲೇಷಣೆ ಹಾಗೂ ಅದರ ಮುನ್ಸೂಚನೆಯು ಕೃಷಿ ಮತ್ತು ಸಂಬಂಧಿತ ಚಟುವಟಿಕೆಗಳಿಗೆ ಬೇಕಿರುವ ಜಲ ಸಂಪನ್ಮೂಲಗಳ ಯೋಜನೆ ಮತ್ತು ನಿರ್ವಹಣೆಗೆ ಅವಶ್ಯಕವಾಗಿರುತ್ತದೆ. ಆದ್ದರಿಂದ, ಬೆಂಗಳೂರಿನ AICRP (ಕೃಷಿ–ಹವಾಮಾನ), ಜಿಕೆವಿಕೆ ಹಾಗೂ KSNDMC ಯಿಂದ ಸಂಗ್ರಹಿಸಲಾದ ಅರವತ್ತು ವರ್ಷಗಳ (೧೯೬೦–೨೦೧೯) ಮಾಸಿಕ ಮಳೆಯ ದತ್ತಾಂಶವನ್ನು ಬಳಸಿಕೊಂಡು ಕರ್ನಾಟಕದ ನಾಲ್ಕು ಹವಾಮಾನ ಉಪವಿಭಾಗಗಳಾದ ಉತ್ತರ ಒಳನಾಡು, ದಕ್ಷಿಣ ಒಳನಾಡು, ಮಲೆನಾಡು ಮತ್ತು ಕರಾವಳಿ ಉಪವಿಭಾಗಗಳಲ್ಲಿನ ಮಳೆಯ ಪ್ರವೃತ್ತಿ, ಅದರ ವರ್ಗಾವಣೆಯ ಮಾದರಿ ಮತ್ತು ಮಳೆಯ ಮುನ್ಸೂಚನೆಯನ್ನು ವಿಶ್ಲೇಷಿಸಲು ಪ್ರಸ್ತುತ ಅಧ್ಯಯನವನ್ನು ಕೈಗೊಳ್ಳಲಾಗಿದೆ. ಮಳೆಯ ಪ್ರವೃತ್ತಿಯನ್ನು ವಿಶ್ಲೇಷಿಸಲು, ಮ್ಯಾನ್–ಕೆಂಡಾಲ್ ಹಾಗೂ ಮಾರ್ಪಡಿಸಿದ ಮ್ಯಾನ್–ಕೆಂಡಾಲ್ ಪರೀಕ್ಷೆಗಳನ್ನು ಬಳಸಲಾಗಿದೆ. ವಾರ್ಷಿಕ ಮಳೆ ಮಾಹಿತಿಗೆ, ಮ್ಯಾನ್–ಕೆಂಡಾಲ್ ಪರೀಕ್ಷೆಯ ಫಲಿತಾಂಶಗಳು ಎಲ್ಲಾ ಉಪವಿಭಾಗಗಳಲ್ಲಿ ಯಾವುದೇ ಮಹತ್ವದ ಪ್ರವೃತ್ತಿಯನ್ನು ಹೊಂದಿಲ್ಲ ಎಂದು ಬಹಿರಂಗಪಡಿಸಿದೆ. ಆದರೆ, ಮಾರ್ಪಡಿಸಿದ ಮ್ಯಾನ್–ಕೆಂಡಾಲ್ ಪರೀಕ್ಷೆಯು ದಕ್ಷಿಣ ಒಳನಾಡು (೧.೪೭) ಹಾಗೂ ಕರಾವಳಿ (೫.೧೦) ಉಪವಿಭಾಗಗಳಲ್ಲಿ ಮಳೆಯು ಹೆಚ್ಚುತ್ತಿರುವ ಪ್ರವೃತ್ತಿಯನ್ನು ತೋರಿಸಿದರೆ, ಉತ್ತರ ಒಳನಾಡು (–೧.೪೦) ಉಪವಿಭಾಗದಲ್ಲಿ ಕಡಿಮೆಯಾಗುತ್ತಿರುವ ಪ್ರವೃತ್ತಿ ಮತ್ತು ಮಲೆನಾಡು ಉಪವಿಭಾಗದಲ್ಲಿ ಯಾವುದೇ ಪ್ರವೃತ್ತಿ ಇಲ್ಲ ಎಂದು ಸೂಚಿಸಿರುತ್ತದೆ. ಮಳೆಯ ವರ್ಗಾವಣೆ ಮಾದರಿಯನ್ನು ನಿರ್ಣಯಿಸಲು ಸಂಬಾವ್ಯತೆ ಅನುಪಾತ ಪರೀಕ್ಷೆಯನ್ನು ಬಳಸಲಾಗಿದೆ. ೨೦೧೦ ರ ನಂತರ ಉತ್ತರ ಒಳನಾಡು ಉಪವಿಭಾಗದಲ್ಲಿ ಮಳೆಯು ತನ್ನ ವರ್ಗವನ್ನು ಬದಲಾಯಿಸಿದ ನಂತರ ಮಳೆ ವಿತರಣೆ ಕಡಿಮೆಯಾಗುತ್ತಿದ್ದು, ದಕ್ಷಿಣ ಒಳನಾಡು, ಮಲೆನಾಡು ಹಾಗೂ ಕರಾವಳಿ ಉಪವಿಭಾಗಗಳಲ್ಲಿ ಕ್ರಮವಾಗಿ ೧೯೬೮, ೨೦೧೬ ಹಾಗೂ ೧೯೬೬ ರಲ್ಲಿ ವರ್ಗವನ್ನು ಬದಲಾಯಿಸಿದ ನಂತರ ಹೆಚ್ಚು ಮಳೆ ವಿತರಣೆಯನ್ನು ಗಮನಿಸಲಾಗಿದೆ. ಮಾಸಿಕ ಮಳೆಯ ಮುನ್ಸೂಚನೆಗಾಗಿ, ಹಾಲ್ಟ್–ವಿಂಟರ್ಸ್ ಘಾತೀಯ ಸಾಗಿಸುವಿಕೆ (H-WES), ಸರಿಮಾ (SARIMA), ಆರ್ಚ್ (ARCH), ಗಾರ್ಚ್ (GARCH) ಮತ್ತು ಎಎನ್ಎನ್ (ANN) ಸಮಯ–ಸರಣಿಯ ಮಾದರಿಗಳನ್ನು ಬಳಸಿಕೊಳ್ಳಲಾಗಿದೆ. ಎಲ್ಲಾ ಉಪವಿಭಾಗಗಳಲ್ಲೂ, ಕಡಿಮೆ ಆರ್ಎಮ್ಎಸ್ಇ (RMSE) ಮೌಲ್ಯದ ಆಧಾರದ ಮೇಲೆ ಎಎನ್ಎನ್ ಮಾದರಿಯು ತರಬೇತಿ ಮತ್ತು ಪರೀಕ್ಷಾ ದತ್ತಾಂಶಗಳ ಮೇಲೆ ಇತರ ಮಾದರಿಗಳಿಗಿಂತ ಉತ್ತಮವಾಗಿ ಕಾರ್ಯ ನಿರ್ವಹಿಸುತ್ತಿದೆ. ಆದ್ದರಿಂದ, ಕರ್ನಾಟಕದ ಎಲ್ಲಾ ನಾಲ್ಕು ಹವಾಮಾನ ಉಪವಿಭಾಗಗಳಲ್ಲಿ ಮಾಸಿಕ ಮಳೆಯ ದತ್ತಾಂಶವನ್ನು ಮುನ್ಸೂಚಿಸಲು ಎಎನ್ಎನ್ ಮಾದರಿಯನ್ನು ಬಳಸಬಹುದಾಗಿದೆ.

ಡಿಸೆಂಬರ್, ೨೦೨೦ **(ಮೋಹನ್ ಕುಮಾರ್, ಟಿ.ಎಲ್.)**

ಕೃಷಿ ಸಂಖ್ಯಾಶಾಸ್ತ್ರ, ಆನ್ವಯಿಕ ಗಣಿತ ಪ್ರಧಾನ ಸಲಹೆಗಾರರು
ಮತ್ತು ಗಣಕಯಂತ್ರ ವಿಜ್ಞಾನ ವಿಭಾಗ
ಕೃಷಿ ವಿಶ್ವವಿದ್ಯಾನಿಲಯ, ಗಾಂಧಿ ಕೃಷಿ ವಿಜ್ಞಾನ ಕೇಂದ್ರ
ಬೆಂಗಳೂರು–೬೫

Forecasting of Rainfall in Meteorological Subdivisions of Karnataka using Non-Linear Statistical Models

KODANDARAMA, S. R., PALB 8192

Department of Agricultural Statistics, Applied Mathematics and Computer Science

University of Agricultural Sciences, GKVK, Bengaluru-65.

INTRODUCTION

In general, study on rainfall is an important task of researcher because changes in occurrence of rainfall can have serious and damaging effects on human society and infrastructure as well as on ecosystems and wildlife. It is of great importance for India's economy as agriculture is the primary source of the livelihood for about 58 per cent of India's population. The average annual rainfall in India and Karnataka are respectively 1194 mm and 1248 mm.

Based on the common rainfall distribution pattern, India Meteorological Department (IMD) has classified India into 36 meteorological subdivisions and Karnataka into four meteorological subdivisions *viz.*, North Interior Karnataka (NIK), South Interior Karnataka (SIK) and Coastal Karnataka. Karnataka State Natural Disaster Monitoring Center (KSNDMC) classified Karnataka into four meteorological subdivisions *viz.*, NIK, SIK, Coastal and Malnad. The average annual rainfall of NIK, SIK, Malnad and Coastal is 731, 1126, 2160 and 3456 mm respectively. Coastal Karnataka receives the highest amount of rainfall among four subdivisions.

The study of rainfall trends is critically important for a country like India whose food security and economy are dependent on the timely occurrence of rainfall. The success or the failure of crops particularly under rainfed conditions depends upon the amount and extent of rainfall. Therefore, it is important to study the trend of rainfall by fitting a appropriate statistical models to the past rainfall data. Hence, the statistical analysis of trend of rainfall plays more important role in agricultural planning, production and policy making.

OBJECTIVE

To analyse the trend in the rainfall pattern in meteorological subdivisions of Karnataka.

MATERIAL AND METHODS

- ✓ The secondary data of monthly rainfall over a period of 60 years (1960-2019) for four meteorological subdivisions of Karnataka was collected from AICRP, Agro-Meteorology, UAS, GKVK, Bengaluru and Karnataka State Natural Disaster Monitoring Center (KSNDMC), Yelahanka, Bengaluru.
- ✓ Collected monthly rainfall data for each subdivision were analysed using two non-parametric tests namely Mann-Kendall (M-K) and Modified Mann-Kendall (MM-K) tests, to analyse possible monotonic trend in monthly rainfall data.

Test	Tau (τ)	Z-value
Mann-Kendall (M-K) test	$\tau = \frac{S}{n(n-1)/2}$	$z = \begin{cases} \frac{S-1}{\sqrt{V(S)}} & if \ S>0 \\ 0 & if \ S=0 \\ \frac{S+1}{\sqrt{V(S)}} & if \ S<0 \end{cases}$
Modified Mann-Kendall (MM-K) test	$\tau = \frac{S}{n(n-1)/2}$	$z = \begin{cases} \frac{S-1}{\sqrt{V^*(S)}} & if \ S>0 \\ 0 & if \ S=0 \\ \frac{S+1}{\sqrt{V^*(S)}} & if \ S<0 \end{cases}$

where, S = Sum of signs of differences of consecutive observation; n = Total number of observations; $V(S)$ = Variance of S, $V^*(S)$ = Corrected variance of S = $\frac{n^*}{n}$ and n^* is Effective Sample Size (ESS)

RESULTS

Table 1: M-K and MM-K test statistic (Tau), p-value and nature of trend for all four meteorological subdivision of Karnataka for monthly rainfall (mm) data

Period	NIK subdivision					SIK subdivision					Malnad subdivision					Coastal subdivision				
	M-K test		MM-K test		Sen's slope	M-K test		MM-K test		Sen's slope	M-K test		MM-K test		Sen's slope	M-K test		MM-K test		Sen's slope
	Tau	NoT	Tau	NoT		Tau	NoT	Tau	NoT		Tau	NoT	Tau	NoT		Tau	NoT	Tau	NoT	
Winter	0.05^{NS}	-	0.05*	↑	0.03	0.07^{NS}	-	0.07**	↑	0.02	0.07^{NS}	-	0.07^{NS}	-	0.01	0.14^{NS}	-	0.14**	↑	0.06
Pre-monsoon	-0.06^{NS}	-	-0.06^{NS}	-	-0.19	0.15^{NS}	-	0.15**	↑	0.58	-0.01^{NS}	-	-0.01^{NS}	-	-0.11	0.03^{NS}	-	0.03^{NS}	-	0.18
Monsoon	-0.09^{NS}	-	-0.09**	↓	-0.64	0.11^{NS}	-	0.11**	↑	0.92	0.03^{NS}	-	0.03^{NS}	-	0.56	-0.03^{NS}	-	-0.03^{NS}	-	-0.74
Post-monsoon	-0.12^{NS}	-	-0.12**	↓	-0.69	-0.02^{NS}	-	-0.02^{NS}	-	-0.12	-0.03^{NS}	-	-0.03^{NS}	-	-0.25	0.02^{NS}	-	0.02^{NS}	-	0.16
Annual	-0.12^{NS}	-	-0.12**	↓	-1.40	0.11^{NS}	-	0.11**	↑	1.47	0.08^{NS}	-	0.08^{NS}	-	2.30	0.12^{NS}	-	0.12*	↑	5.10

NS : Non-significant; * & ** : Significant at 5 & 1 % level; NoT : Nature of Trend; - : No trend; ↑ : Monotonic increasing trend; ↓ : Monotonic decreasing trend

For the purpose of trend analysis collected monthly rainfall (mm) data was converted into seasonal and annual rainfall (mm) data by adding up respective months rainfall for all meteorological subdivisions. For the seasonal and annual rainfall data the above mentioned non-parametric tests were applied using 'modifiedmk' package of R software to know the possible trends in rainfall occurrence.

For seasonal and annual rainfall data of each subdivision, Tau value (M-K and MM-K test), Sen's slope estimator and nature of trend are tabulated in Table 1. The results in Table 1 shows that M-K test (Tau) values for seasonal and annual rainfall data are found to be non-significant for all the subdivisions, which indicates that there is no monotonic trend in seasonal and annual rainfall data. This may be due to the influence of serial correlation in the seasonal and annual rainfall data. Therefore, to overcome the effect of serial correlation, if any, MM-K test was employed.

MM-K test values for NIK subdivision indicates monotonic increasing trend in winter season, monotonic decreasing trend in monsoon and post-monsoon seasons and no monotonic trend in pre-monsoon season rainfall data. Sen's slope estimator values towards zero indicates decreasing or increasing rate of change in seasonal rainfall data is negligible. Further, MM-K test value indicates monotonic decreasing trend in annual rainfall data. Larger negative Sen's slope value of -1.40 indicates high decreasing rate of change of annual rainfall data.

MM-K test values for SIK subdivision indicates monotonic increasing trend in winter pre-monsoon and monsoon seasons and no monotonic trend in post-monsoon season rainfall data. Sen's slope estimator values towards zero indicates increasing rate of change in seasonal rainfall data is negligible. Further, MM-K test value indicates monotonic increasing trend in annual rainfall data. Larger positive Sen's slope value of 1.47 indicates high increasing rate of change of annual rainfall data.

MM-K test values for Malnad subdivision indicates no monotonic trend in seasonal and annual rainfall data.

MM-K test values for Coastal subdivision indicates monotonic increasing trend in winter season and no trend in pre-monsoon, monsoon and post-monsoon seasons rainfall data. Sen's slope estimator values (0.06) towards zero indicates increasing rate of change in winter season rainfall data is negligible. Further, MM-K test value indicates monotonic increasing trend in annual rainfall data. Larger positive Sen's slope value of 5.10 indicates high increasing rate of change of annual rainfall data.

DISCUSSION

Even though M-K test is most commonly used test for detecting trend in rainfall data, it assumes that sample data is serially independent. The existence of positive autocorrelation in the data increases the probability of detecting trends when actually none exists, and vice versa. Therefore, MM-K test, which eliminate the effect of serial correlation present in the time-series data on the M-K test statistic by correcting the variance using ESS is employed. Pal and Al-Tabbaa (2011) and Sridhara *et al.* (2020) obtained similar results.

SUMMARY

Analysis of trend in seasonal and annual rainfall data of four meteorological subdivisions of Karnataka for the period of 60 years (1960-2019) is done by using M-K and MM-K test. M-K test indicated no significant trend whereas, MM-K revealed significant monotonic increasing and decreasing trends in seasonal and annual rainfall data.

ADVISORY COMMITEE

Chairperson: Dr. Mohan Kumar, T.L.

Members : Dr. D. M. Gowda

Dr. H. S. Shivaramu

Dr. T. V. Krishna

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Sl. No.	ABBREVIATIONS	EXPANSIONS
1	LPA	Long Period Average
2	IMD	India Meteorological Department
3	NIK	North Interior Karnataka
4	SIK	South Interior Karnataka
5	ARMA	Autoregressive Moving Average
6	ARCH	Autoregressive Conditional Heteroscedasticity
7	GARCH	Generalized Autoregressive Conditional Heteroscedasticity
8	M-K	Mann-Kendal
9	MM-K	Modified Mann-Kendall
10	MLP-NN	Multi-Layer Perceptron Neural Network
11	ANN	Artificial Neural Network
12	SARIMA	Seasonal Autoregressive Integrated Moving Average
13	H-WES	Holt-Winters Exponential Smoothing
14	AR(1)	Autoregression of order one
15	MA(1)	Moving Average of order one
16	ESS	Effective Sample Size
17	MAE	Mean Absolute Error
18	MSE	Mean Square Error
19	RMSE	Root Mean Squared Error
20	MAPE	Mean Absolute Percentage Error
21	AIC	Akaike’s Information Criterion

Sl. No.	ABBREVIATIONS	EXPANSIONS
22	BIC	Bayesian Information Criterion
23	MMF	Morgan Mercer Flodin
24	PELT	Pruned Exact Linear Time
25	TM	Mean Temperature
26	TMAX	Maximum Temperature
27	TMIN	Minimum Temperature
28	SQM-K	Sequential Mann-Kendall test
29	SNH	Standard Normal Homogeneity
30	ARIMA	Autoregressive Integrated Moving Average
31	USA	United States of America
32	RNNS	Recurrent Artificial Neural Networks
33	SVMs	Support Vector Machines
34	RSVM	Recurrent Support Vector Machines
35	CPSO	Chaotic Particle Swarm Optimization Algorithm
36	ARNN	Autoregressive Neural Network
37	L-M	Lagrange-Multiplier
38	PAR	Periodic Autoregressive
39	MPARCH	Mixture Periodic ARCH
40	MSPE	Mean Square Prediction Error
41	RMAPE	Relative Mean Absolute Percentage Error
42	ADF	Augmented Dickey-Fuller test
43	SAM	Simple Average Method
44	MLR	Multiple Linear Regressions

Sl. No.	ABBREVIATIONS	EXPANSIONS
45	GLM	Generalized Linear Model
46	GEP-ARCH	Gene Expression Programming-Autoregressive Conditional Heteroscedasticity
47	ANN-ARCH	Artificial Neural Networks-Autoregressive Conditional Heteroscedasticity
48	M Ha	Million Hectares
49	AICRP	All India Coordinated Research Project
50	KSNDMC	Karnataka State Natural Disaster Monitoring Center
51	BR	Buishand range
52	C.F	Correction Factor
53	MLEs	Maximum Likelihood Estimators
54	ACF	Autocorrelation Functions
55	MLP	Multi-Layer Perceptron

I INTRODUCTION

India has total geographical area of about 328.7 million hectares, of which 140.1 million hectares is the reported net sown area and 198.4 million hectares is the gross cropped area with a cropping intensity of 142 per cent. The net sown area works out to be 43 per cent of the total geographical area. The net irrigated area is 68.4 million hectares (Anon. 2019). India is influenced by different climatic conditions in different parts of the country. Therefore, there is significant variability in rainfall in both space and time. Southwest monsoon season normally brings almost 80 per cent of the rainfall, while Northeast monsoon season brings a significant portion of the rest, and summer and winter contribute almost no rainfall (Vignesh *et al.*, 2019).

Agriculture is backbone of India's economy and is largely dependent on the monsoon. Rainfall plays a dominant role in agricultural production and productivity, despite advancement in many technological fronts. The South-west monsoon in the year 2018 over the country was below normal with 91 per cent of its long period average (LPA) and nine meteorological subdivisions of the country witnessed deficit rainfall (varies from 20 to 37% deficit). The second half of the season (August-September) witnessed below average rainfall (86% of LPA) leading to mid-season drought in many parts of the country. Seasonal rainfall was normal (98% of LPA) in both North-west India and South Peninsula, whereas below normal (93% of LPA) in Central India, and deficient (76% of LPA) in East and North-east India. The intra-annual, inter-seasonal and intra-seasonal rainfall variations affect agriculture from land preparation to realization of potential yield of crops (Vijaya Kumar *et al*., 2019).

Rainfall serves as a key input for many agriculture and allied activities. Therefore, an adequate understanding of the spatial and temporal dynamics of rainfall is important for the planning and management of agriculture related activities. However, it is often difficult, because rainfall occurs because of complex and nonlinear interactions among numerous atmospheric, land and ocean processes (Ghosh *et al.*, 2010). The benefit of studying the characteristic of rainfall is to provide insight into rainfall variability and to help in improving the range of rainfall forecast.

An accurate understanding of rainfall characteristics and soil variability is critical to optimizing farm production and to precision farming. The rains are highly variable in time, space, amount and duration and water is the most important limiting factor for biological and agricultural activities. Seasonal changes in rainfall patterns may alter the cropping pattern and cropping system (Delitala *et al.*, 2000).

Forecast of rainfall is essential for planning and management of water resources especially in an agriculture-based country like India. About 65 per cent of the total cultivated land in India is under the influence of the rainfed agriculture system. Especially, monthly and seasonal rainfall forecasts provide useful information for water resource management, agricultural planning and its associated crop insurance application (Garbrecht *et al.*, 2010).

The subject area of climate change is vast; the changing pattern of rainfall is a topic within this field that deserves urgent and systematic attention, since it affects both the availability of fresh water and food production (Dore, 2005). Higher the rainfall or lower the rainfall or changes in its spatial and seasonal distribution would influence the spatial and temporal distribution of runoff, soil moisture and groundwater reserves and would affect the frequency of droughts and floods. Further, the temporal change in rainfall distribution will affect cropping patterns and productivity.

According to India Meteorological Department (IMD), Pune, depending on the spatial and temporal scales of atmospheric systems and the details of the accuracy desired, the weather forecasts are divided into the following categories (Tyagi, 2008).

(i) **Now casting:** In this method, the details about the current weather and forecasts up to a few hours ahead are given.

(ii) **Short-range forecasts** (1 to 3 days): In this method, the weather (mainly rainfall) in each successive 24 hours intervals is predicted up to 3 days. This forecast range is mainly concerned with the weather systems observed in the latest weather charts, although the generation of new systems is also considered.

(iii) **Medium range forecasts** (4 to 10 days): Average weather conditions and the weather on each day may be prescribed with progressively lesser details and accuracy than that for short-range forecasts.

(iv) **Long-range or Extended Range forecasts** (more than 10 days to a season): In this method, there is no rigid definition for Long-Range Forecasting, which may range from a monthly to a seasonal forecast.

The rainfall in India has large spatial and temporal variations. The average amount of rainfall that India receives annually is 1194 mm (Tadesse and Dinka, 2017). The highest amount of rainfall in the country is witnessed in Mawsynram of Meghalaya (11,873 mm) followed by Chirapunji (11,619 mm) and Agumbe (7691 mm). The lowest rainfall is witnessed in Jaisalmer of Rajasthan (186.2 mm) (Anon., 2017).

Based on the common rainfall distribution pattern, the Indian Meteorological Department (IMD) classified the country into 36 meteorological subdivisions. Out of the total 36 meteorological subdivisions, monsoon rainfall was excess in Kerala, normal in 23 and deficient in 12 subdivisions in the year 2018. Among the meteorological subdivisions, annually the district of Meghalaya receives the highest average rainfall (7679.8 mm) and Ferozepur district of Punjab gets low average rainfall (94.2 mm). Meteorological subdivision as a whole, annually, Konkan and Goa receive the highest average rainfall (3443.4 mm) and West Rajasthan gets the lowest average rainfall (408.3 mm). In South-West monsoon season, Konkan and Goa receive the highest average rainfall, whereas Haryana, Punjab, and Chattisgarh receive the lowest average rainfall (Kumar *et al.*, 2018).

Karnataka State has humid to sub-humid monsoonal climate on the West Coast and the Western Ghats and semi-arid to arid (very warm) climate in central and northern districts of the plateau region (Srinivasareddy *et al.*, 2019). The average annual rainfall of Karnataka is 1248 mm, Agumbe and Hulikal receive the highest and Bagalkot receives the lowest annual rainfall in Karnataka.

Need for the study

Karnataka state is facing frequent drought conditions due to uncertainty in rainfall and spatio-temporal variation in different regions. Due to these variations in rainfall distribution, the state is experiencing extreme weather conditions like drought or flood in different regions. Frequent drought and floods have caused heavy loss to crops and lives, while society loses its livelihood and rhythm of life and people are forced to migrate for their sustenance.

As rainfall is an unpredictable natural phenomenon, statistical analysis of its distribution patterns and quantum of rainfall plays an important role (Karim *et al.,* 2018). The study on the statistical analysis of rainfall may help the government in agricultural planning and policymaking, it helps farmers in contingency planning of the crop, adopting the farm production practices, management of the farm production etc. (Husak *et al.*, 2007).

Unlike irrigated agriculture, rainfed farming is usually diverse and risk prone. Small variations in the timing and the quantity of rainfall have a potential impact on agricultural output (Bhowmik *et al.,* 2017). Since there is a direct relationship between rainfall and agriculture, any small amount of change in the distribution pattern of rainfall would adversely affect the production and productivity of crops (Adarsh and Reddy, 2015). Hence, change in the amount, the pattern of rainfall, and its influence on agriculture have become the most important concern.

Shifting pattern analysis is an efficient tool to understand the fundamental information in meteorological data such as rainfall, discharge, temperature etc., especially the result of a reasonable shifting point detection method can be effectively used in the prediction of flood and drought because it provides a key to resolve the non-stationary or inhomogeneous problem by climate change (Lee and Kim, 2016).

The forecasting of rainfall is an important part of the rainfall study as it helps in managing the rainfall issues. Heavy and low rainfalls are two important changes in the rainfall pattern. The heavy rainfall leads to natural calamities like flood submergence of land under agriculture. Low rainfall results in difficulty to meet the water requirements.

The proper study of rainfall issues may overcome these problems. Proper planning to avoid natural calamities like floods is very much essential. Managing the low rainfall through an alternative technique called irrigation and to conserve the water through water conservators, it requires the study of rainfall for forecasting future scenarios. Therefore, modelling and forecasting of rainfall are very much required.

Considering the above points, the following objectives are laid out for the present study:

Objectives of the study

1. To analyze the trend in the rainfall pattern in meteorological subdivisions of Karnataka
2. To study the shift in rainfall distribution pattern
3. To evaluate a suitable non-linear time series models for forecasting rainfall

To know the pattern of the trend, month-wise, seasonal and annual rainfall data were analyzed by employing non-parametric Mann-Kendal (M-K), Modified Mann-Kendall (MM-K) tests and Sen's slope estimator. To know the shift in rainfall distribution pattern Likelihood Ratio test was employed. To evaluate suitable non-linear time series models for forecasting monthly rainfall in different meteorological subdivisions of Karnataka, three nonlinear time-series models namely Autoregressive Conditional Heteroscedasticity (ARCH), Generalized ARCH (GARCH) and Multi-Layer Perceptron Neural Network (MLP-NN) or Artificial Neural Network (ANN) were employed along with two linear models namely Seasonal Autoregressive Integrated Moving Average (SARIMA) and Holt-Winters Exponential Smoothing (H-WES) to compare the performance of these models.

II REVIEW OF LITERATURE

In this chapter, an effort has been made to collect the review of literature of past research work related to this study. The reviews are exhibited under the following headings:

2.1 To analyze the trend in the rainfall pattern in meteorological subdivisions of Karnataka

2.2 To study the shift in rainfall distribution pattern

2.3 To evaluate a suitable non-linear time-series models for forecasting rainfall

2.1 To analyze the trend in the rainfall pattern in meteorological subdivisions of Karnataka

Kulkarni and Storch (1995) examined the effect of serial correlation on the performance of the Mann-Kendall (M-K) test by using Monte-Carlo simulation and concluded that even moderate serial correlation makes the test liberal so that it signals inaccurately the presence of significant trend more often than permitted. They proposed a simple "prewhitening" procedure for serially correlated data.

Hamed and Rao (1998) concluded that the existence of positive autocorrelation in the data increases the probability of detecting trend when no trend exists in the data, and vice versa. They discussed the effect of autocorrelation on the variance of the M-K trend test statistic and derived the theoretical relationship to calculate the variance of the M-K test statistic for autocorrelated data. The special cases of AR (1) and MA (1) dependence were discussed as examples. Based on the modified value of the variance of the M-K test statistic, a modified non-parametric trend test that is suitable for autocorrelated data was proposed. They found that the accuracy of the modified test in terms of its empirical significance level was superior to that of the original M-K trend test without any loss of power. They applied the modified test to rainfall and streamflow data to demonstrate its performance as compared to the original M-K test.

Yue *et al*. (2002) studied the interaction between linear trend and lag-one autoregressive *i.e.,* AR (1) process when they exist in a time-series using Monte Carlo

simulation. Simulation experiments demonstrated that the existence of serial correlation alters the variance of the estimate of the M-K test statistic, and the presence of a trend alters the estimate of the magnitude of serial correlation. Also, they showed that removal of positive serial correlation component from time-series by pre-whitening resulted in reduction in the magnitude of the existing trend, and the removal of trend component from a time-series as a first step before pre-whitening to eliminates the influence of the trend on the serial correlation and not to affect seriously the estimate of the true AR (1). Results indicate that the commonly used pre-whitening procedure for eliminating the effect of serial correlation on the M-K test leads to potentially inaccurate assessments of the significance of trend.

Yue and Wang (2004) used M-K test to assess the significance of trend in rainfall time-series. They quoted that "the presence of serial correlation in time-series will affect the ability of the test to correctly assess the significance of trend". They proposed new Effective Sample Size (ESS) for variance correction, which modifies the M-K test statistic and eliminate the effect of serial correlation on it. The superiority of this proposed method was illustrated by Monte-Carlo simulation study. Their simulation results revealed that when no trend exists within time-series, ESS could effectively limit the effect of serial correlation on the M-K test.

Bayazit and Onoz (2007) showed that the M-K test used to detect a trend in a time-series yields an incorrect (too large) rejection rate when applied to an autocorrelated series with no trend. To overcome this problem, they employed a pre-whitening procedure, but it reduces the power of the test when there is trend in the data. Through simulation studies, they found that in general, pre-whitening should be avoided when the test has a high power *i.e.*, when the coefficient of variation is very low, the slope of trend is high, and the sample size is large. They also concluded that pre-whitening would prevent the false detection of non-existing trend, without significant power loss in identifying trend that exists in the data.

Modarres and Silva (2007) analyzed the annual rainfall, the number of rainy days per year and monthly rainfall of 20 stations to assess climate variability in arid and semi-

arid regions of Iran by employing M-K test. Their results showed mixed trend *i.e.*, increasing and decreasing rainfall for various regions of study. Increasing and decreasing monthly rainfall trend were found over large continuous areas in the study region. These trends were statistically significant during the winter and spring seasons, suggesting a seasonal movement of rainfall concentration. Their results also showed that there is no significant climate variability in the arid and semi-arid environments of Iran.

Hamed (2008) stated that climatic variability will adversely affect trend results. He modified the M-K test statistic, which is widely used to detect trend in rainfall data to account for the effect of scaling. He derived the exact expressions for the mean and variance of the test statistic under the scaling hypothesis, and they showed that the normal distribution is the reasonable approximation. A procedure for estimating the modified variance from observed data was also outlined. Then the modified test was applied to group of 57 worldwide total annual river flow time-series data from the database of the Global Runoff Data Centre in Koblenz, Germany. His results showed considerable reduction in the number of stations with significant trend when the effect of scaling was taken into account, although for the same data the earlier studies exhibited significant trend in annual maximum flow. His results indicate that the evidence of real trend in rainfall data is even weaker than suggested by earlier studies and highly significant increasing trend seem to be more common than negative ones.

Kumar *et al*. (2010) studied monthly, seasonal and annual trend of rainfall using M-K test and Sen's slope estimator. They analyzed monthly rainfall data of 135 years (1871-2005) for 30 sub-divisions of India. Half of the subdivisions showed an increasing trend in annual rainfall, but for only three subdivisions *viz.* Haryana, Punjab and Coastal Karnataka trend was statistically significant. Similarly, only one subdivision (Chattisgarh) indicated statistically significant decreasing trend out of the 15 subdivisions showing a decreasing trend in annual rainfall. For the whole of India, no significant trend was detected for annual, seasonal and monthly rainfall data.

Krishnakumar *et al.* (2009) attempted to study temporal variation in monthly, seasonal and annual rainfall over Kerala, India, for the period from 1871 to 2005. Long

term changes in rainfall determined by M-K test statistic and linear trend model. Their analysis revealed that significant decrease in southwest monsoon rainfall while an increase in post-monsoon season rainfall over the state of Kerala. Rainfall during winter and summer seasons showed non-significant increasing trend. Rainfall during June and July month showed significant decreasing trend and for January, February and April months showed an increasing trend.

Mohan Kumar *et al*. (2012) developed nonlinear statistical growth model to provide an analytical approach to describe trend in area under coffee production in India. Six nonlinear statistical growth models were fitted. The parameters of each model were estimated using the Levenberg-Marquardt iterative method. The main assumptions of 'independence' and 'normality' of error were examined by using the 'Run-test' and 'Shapiro-Wilk test', respectively. The two best-fitted models were selected based on the performance of several goodness of fit criteria *viz*. Mean Absolute Error (MAE), Mean Square Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC). Logistic and Morgan Mercer Flodin (MMF) models were found to be quite successful in describing the trend in area under coffee production. Further, the study revealed that the series had an upward trend in area under coffee production. Forecast values were also computed using two best-fitted models; result revealed that forecasted values of the Logistic model are slightly larger than the MMF model.

Mondal *et al.* (2012) employed daily rainfall data of 40 years from 1971-2010 to study the monthly variability of rainfall in a river basin of Orissa near the coastal region using M-K test, MM-K test along with the Sen's Slope estimator for the determination of trend and slope magnitude. Their results showed that there were increasing and decreasing trend in the monthly rainfall of their study area.

Mapurisa and Chikodzi (2014) analyzed trend in the proportional contribution to the seasonal rainfall of each month for the meteorological stations of Buffalo Range, Masvingo airport and Zaka in South-Eastern Zimbabwe. Their result indicates the existence of some trend for all the stations and months. However, when subjected to M-K

trend analysis, all the trends were found to be statistically non-significant for all instances except for the month of October at the Zaka station, which showed significant trend of increasing proportional rainfall.

Bari *et al.* (2016) investigated 50 years (1964-2013) of seasonal and annual rainfall trend and their fluctuation in northern Bangladesh. They tested for the autocorrelation and used non-parametric M-K test statistic along with Sen's Slope estimator to examine rainfall trend and their magnitudes. They also used the sequential M-K test statistic to identify any fluctuations in the trend over time and to detect the possible points of change in the rainfall series. They found that pre-monsoon and post-monsoon rainfall is increasing in most of the rainfall stations. However, the sequential M-K test statistic detected decreasing trend in pre-monsoon rainfall after early 1990's. Monsoon rainfall showed decreasing trend in the majority of the area studied. The sequential M-K test detected several non-significant points of change for seasonal and annual rainfall for most of the stations. Finally, they concluded that there were decreasing seasonal rainfall trend after the early 1990's for the majority of the stations.

Gajbhiye *et al.* (2016) reported that the study of trend in rainfall is critically important for a country like India whose food security and economy are dependent on the timely availability of water. The rainfall data for the periods 1901-2002 and 1942-2002 of the Sindh river basin, India were analyzed for trend in monthly, seasonal and annual data. After the removal of the effect of all significant autocorrelation coefficients in the time-series data, the M-K, MM-K tests and Sen's slope estimator were employed to identify the rate of trend. Their results suggested significant increase in rainfall for both seasonal and annual series.

Verma *et al.* (2016) studied the impact of climate change in terms of anomalies in rainfall has risen as a major challenge in the world, as it may lead to havoc by intense floods or severe droughts. Their study indicated trend in annual and monthly rainfall data collected from 39 stations for 32 years (1981-2012) for the Seonath river basin, Chhattisgarh, India using M-K test and Sen's slope estimator test. Their results revealed that few stations had significant change and the analysis for the overall Seonath river basin

showed an increasing trend of rainfall in the monsoon season and decreasing trend of rainfall in the post-monsoon season. Their study indicated that the annual rainfall had an increasing trend for the Seonath river basin

Isioma *et al.* (2018) used M-K test statistic to detect and estimate the magnitude of trend associated with rainfall data from Warri and Benin City, Nigeria. Monthly rainfall data for 36 years from 1980-2016 were used for analysis. Missing records of rainfall were estimated with the aid of the expectation-maximization algorithm. Preprocessing of the rainfall data was done by using three validation tests namely, test of homogeneity, test of normality and outlier detection. Results of the analysis revealed that the rainfall data from Warri and Benin City are statistically homogeneous. The observations did not contain outliers and they were not normally distributed as expected for most of the climatic variables. The M-K test revealed that Benin City had an increasing trend while Warri city had decreasing trend.

Hu *et al.* (2020) proposed an improved M-K test for rainfall data with both positive and negative autocorrelation. They modified the variance and distribution patterns of *S* under the scaling hypothesis. They compared different estimators of the scaling coefficient and concluded that the generalized normal distribution is the best approximation to fit the exact distribution of *S* with both positive and negative autocorrelation. They demonstrated that the centered detrending moving average is the best estimator of the scaling coefficient for series with trend. Their improved M-K test was proven more advantageous than other trend detection methods for the simulated time-series and observed rainfall data obtained from nine hydrometric stations located in China.

2.2 To study the shift in rainfall distribution pattern

Gombay and Horvath (1994) derived maximum likelihood type of statistic for testing sequence of observations for no change in the parameter against possible change. They also proved that the limit distribution of the suitably normalized and centralized statistic is double exponential, *i.e.*, Gumbel distribution under the null hypothesis.

Liu *et al.* (2008) proposed nonparametric method based on the empirical likelihood to detect the change-point in the coefficient of linear regression models. They proved that the empirical likelihood ratio test statistic had the same asymptotic null distribution as that with classical parametric likelihood. They showed that under some mild conditions, the maximum empirical likelihood change-point estimator is consistent. The sensitivity and robustness of the proposed approach were illustrated with simulated and real datasets to compare the effectiveness of this method.

Alvarez and Dey (2009) developed general approach to Bayesian isotonic change point problems. Their isotonic change point analysis includes trend and other constraint problems and it captures linear, non-smooth as well as abrupt changes. Desired marginal posterior densities are obtained using a Markov chain Monte Carlo method. This methodology was examined using one simulated and two real data sets, concluded that their proposed Bayesian approach captures the qualitative conclusion about the shape of the trend change.

Dachian (2010) reported that different change-point type models encountered in parametric statistical inference give rise to different limiting likelihood ratio processes. In his paper, he considers two such likelihood ratios. The first one is an exponential function of two-sided Poisson process driven by some parameter and the second one is an exponential function of two-sided Brownian motion. He established that for sufficiently small values of the parameter, the Poisson type likelihood ratio can be approximated by the Brownian type one. Consequently, several statistically interesting quantities (such as limiting variances of different estimators) related to the first likelihood ratio can also be approximated by those related to the second one. Finally, discussed the asymptotic for large values of the parameter and illustrated the results with numerical simulations.

Jandhyala *et al.* (2013) reviewed methods of inference for single and multiple change-points in time-series when data are of a retrospective (off-line) type. The inferential methods reviewed for single change-point in time-series include likelihood, Bayes, Bayes-type and some relevant non-parametric methods. Inference for multiple change-points requires methods that can handle large data sets and can be implemented efficiently for

estimating the number of change-points as well as their locations. They had given greater emphasis on multivariate data while reviewing inferential methods for single change-point in time-series. In their paper, more attention was paid to the estimation of unknown change-point(s) in time-series. Some data sets for which change-point modelling was carried out in the previous studies were provided as illustrative examples under both single and multiple change-point scenarios.

Killick and Eckley (2014) attempted to tackle the challenges in ability of change point analysis to detect multiple changes within a given time-series or sequence. They developed change point package to provide users with a choice of multiple change-point search methods to use in conjunction with a given change point method and in particular, provides an implementation of the recently proposed Pruned Exact Linear Time (PELT) algorithm.

Zarenistanak *et al.* (2014) conducted trend analysis and change point detection of annual and seasonal rainfall, and Mean Temperature (TM), Maximum Temperature (TMAX) and Minimum Temperature (TMIN) time-series of 50 rainfall stations and 39 temperature stations of southwest Iran for the period 1950–2007. Three statistical tests including Pettitt's test, Sequential M-K test (SQM-K test) and M-K test were used for the analysis. The results obtained for the rainfall series indicated that most stations showed insignificant trend in annual and seasonal series. The trend analysis for temperature revealed that significant increase in temperature during the summer and spring seasons. The results of change-point detection indicated that most of the positive significant mutation points in TM, TMAX and TMIN began in the 1990's.

Jaiswal *et al.* (2015) carried out an assessment of change detection and trend in monthly, seasonal and annual historical series of different climatic variables of Raipur, Chhattisgarh, India. The change detection analysis had been conceded using Pettitt's test, von Neumann ratio test, Buishand's range test and Standard Normal Homogeneity (SNH) test, M-K and Spearman rho tests were used for trend analysis. Their results showed that the annual series of minimum temperature, wind speed, sunshine hour showed significant change points, while evaporation indicated doubtful case and maximum temperature

confirmed the homogeneity at 95 per cent confidence level. The results of change point analysis for meteorological variables indicated different change points from year 1990 to 2000, with maximum change points in and around 1995. They obtained an increasing trend in the annual minimum, maximum temperatures and relative humidity and decreasing trend in pan evaporation and sunshine hour series.

Yau and Zhao (2016) proposed likelihood ratio scan method for estimating multiple change points in piecewise stationary processes. Scan statistics was employed to reduce the computationally infeasible global multiple-change-point estimation problem to number of single-change-point detection problems in various local windows. They also established consistency for the estimated numbers and locations of the change points. They also developed a procedure for constructing confidence intervals for each of the change points. Simulation experiments and real data analysis were conducted to illustrate the efficiency of the likelihood ratio scan method.

Joseph and Tamilmani (2017) analyzed rainfall data for the period 1981-2011 of Coimbatore, Tamil Nadu using the Markov Chain Probability model to describe the long-term behavior of wet or dry weather spells. The model considered less than 20 mm rainfall in a week as a dry week and 20 mm or more as a wet week to estimate the dry and wet spell probability. Their results revealed that the earliest and delayed weeks of onset and withdrawal of rainy season in Coimbatore, and they suggested that sowing of crops should be started in the 27^{th} week.

Rahman *et al.* (2018) studied the regional variation of changing patterns of rainfall in Bangladesh using wavelet transform. The duration of the study period was 69 years for Dhaka, 64 years for Cox's Bazar, 40 years for Rajshahi, 54 years for Bogra and 55 years for Sylhet. The results of the wavelet analysis reveal that in Rajshahi the amount of rainfall is decreasing at a significant rate among the other study regions.

Dette and Gosmann (2020) proposed a new approach for sequential monitoring of a general class of parameters of a *d*-dimensional time-series, which can be estimated by approximately linear functional of the empirical distribution function. They considered a

closed-end method, which was motivated by the likelihood ratio test principle and compared the new method with two alternative procedures. They proved that for large class of testing problems the new detection scheme has asymptotic level α and is consistent. They illustrated asymptotic theory for monitoring change in the mean, variance and correlation. By means of simulation study, they demonstrated that the new test performs better than the currently available procedures for change-point problems.

2.3 To evaluate suitable non-linear time-series models for forecasting rainfall

Toth *et al.* (2000) utilized Artificial Neural Networks (ANNs) to forecast rainfall intensity fields in space and time. Training was conducted using backpropagation where the input and output rainfall fields are presented to the neural network as a series of learning sets. Result indicates that the neural network is capable of learning the complex relationship describing the space-time evolution of rainfall that is inherent in complex rainfall simulation model. One-hour ahead forecasts were produced for training data set and compared with actual. Then built ANN model was used to forecast over testing data. The fitted ANN model performed well when a relatively large number of hidden nodes are utilized.

Luk *et. al.* (2001) highlighted the importance of rainfall forecasting for many catchment management applications, in particular for flood warning systems. They used ANN model, which performs non-linear mapping between inputs and outputs. Three alternative types of ANNs namely multilayer feed forward neural networks, partial recurrent neural networks, and time-delay neural networks were fitted. The data requirement and the accuracy obtainable from these three alternative types of ANNs were also discussed.

Somvanshi *et al.* (2006) employed Autoregressive Integrated Moving Average (ARIMA) and ANN models to evaluate the prediction efficiency of rainfall. For this purpose, they used 104 years of mean annual rainfall data from year 1901-2003 of Hyderabad, India. The models were trained with 93 years of mean annual rainfall data. The performance of the model was evaluated by using remaining 10 years of data. Their study

revealed that ANN model outperformed the ARIMA model in prediction and forecasting the rainfall.

Jain and Kumar (2007) explained the need for increased accuracy in time-series forecasting and things that motivate the researchers to develop innovative models. Their proposed approach consists of an overall modeling framework, which is a combination of conventional and ANN techniques. The proposed hybrid approach for forecasting was tested using the monthly stream flow data at Colorado River at Lees Ferry, USA. The results obtained from the study suggest that the approach of combining the strengths of the conventional and ANN techniques provides robust modelling framework capable of capturing the non-linear nature of the complex time-series and thus producing more accurate forecasts.

Chattopadhyay and Chattopadhyay (2008) identified non-linear methodologies to forecast average summer-monsoon rainfall over India. Three advanced back propagation neural network learning rules namely momentum learning, conjugate gradient descent learning and Levenberg-Marquardt learning, and statistical methodology in the form of asymptotic regression were implemented for this purpose. Monsoon rainfall data of the years from 1871 to 1999 was used. After thorough comparison using statistical procedures, they reported that the greater potential of conjugate gradient descent as learning algorithm for the back propagation neural network.

Hong (2008) investigated the feasibility of a hybrid model of Recurrent Artificial Neural Networks (RNNS) and Support Vector Machines (SVMs), namely Recurrent Support Vector Machines (RSVR), to forecast rainfall depth values. Chaotic Particle Swarm Optimization Algorithm (CPSO) was employed to choose the parameters of an SVR model. Subsequently, an example of rainfall values during typhoon periods from Northern Taiwan was used to illustrate the proposed RSVRCPSO model. Their empirical results revealed that the proposed model performs best for rainfall forecasting.

Hung *et al.* (2009) developed new approach using an ANN's technique to improve rainfall forecast performance. They used 4 years of hourly data from 75 rain gauge stations

in Bangkok area to develop the ANN model. Forecasts by the ANN model had compared to the conventional approach *viz.* simple persistent method. Results showed that superiority of ANN model over the one obtained by the persistent model, and rainfall forecasts for Bangkok from one to three hours ahead were highly satisfactory.

Zhao *et al.* (2009) proposed a novel approach for flood forecasting using the GARCH model, which is powerful in capturing the dynamics of time-series and handling high volatility. River discharge data for the period 2003-2005 was used for model building. They concluded that the GARCH model is effective in river stage forecasting and it outperforms two widely used time-series methods namely ARMA and Support Vector Machine (SVM) based on MSE, RMSE and MAPE values. They also concluded that the GARCH model is effective in providing 1-hour ahead prediction intervals.

Chattopadhyay and Chattopadhyay (2010) employed ARIMA and Autoregressive Neural Network (ARNN) to study the supremacy of models. Summer monsoon (June-August) rainfall of India for the period 1871-1999 was used for this purpose. ARIMA (0,1,1) had been identified as best ARIMA model. ARNN model with sigmoid non-linearity has been used while training the network. Finally, supremacy of ARNN has been established over ARIMA (0,1,1) with statistical evidences.

Ghosh *et al.* (2010) employed GARCH model to describe data sets depicting volatility. Lagrange-Multiplier (L-M) test was used for testing the presence of ARCH effects. Monthly rainfall data for the period 1990-2006 of Sub-Himalayan West Bengal meteorological subdivision, India, was used for modelling and forecasting. GARCH model was employed on the residuals obtained after carrying out Periodic Autoregressive (PAR) modeling of the seasonal variation. Further, the Mixture Periodic ARCH (MPARCH) model, which is an extension of the GARCH model was also applied to zero conditional mean residual series to identify time-varying volatility in the data set. Mean Square Prediction Error (MSPE), MAPE and Relative MAPE (RMAPE) were employed to select the best model. They concluded that the PAR model with AR-GARCH errors has performed better than the SARIMA model for modeling as well as forecasting.

Mohan Kumar *et al*. (2011) compared the forecasting performances of H-WES and SARIMA time-series methodologies for forecasting onion prices in Bangalore market. The monthly average price of onion for the period of April 1999 to November 2010 were used to forecast the future price. The forecast accuracy criteria such as MSE, RMSE, MAE, MAPE and Theil's U coefficient were used to determine the best forecasting model. The study revealed that the time-series data were influenced by upward trend and existence of seasonal factors. Moreover, the ACF and the Augmented Dickey-Fuller (ADF) tests have shown that the time-series data were non-stationary in nature but became stationary after the first order of the difference process was carried out. The results showed that ARIMA model outperformed the H-WES model for forecasting onion price.

Ramana *et al.* (2013) attempted to find an alternative method for rainfall prediction by combining the wavelet technique with ANN. The wavelet and ANN models were applied to monthly rainfall data of Darjeeling rain gauge station. The calibration and validation performance of the models were computed with appropriate statistical methods. The results of monthly rainfall series modelling indicate that the performances of wavelet neural network models are more effective than the ANN models.

Yusof *et al*. (2013) fitted ARIMA and GARCH models to the daily rainfall data sets from Ipoh and Alorsetar stations in Peninsular Malaysia for the period from 1968 to 2003. McLeod-Li test and test based on the L-M principle were applied to the squared residuals from ARIMA models to evaluate performance of models. They concluded that the composite ARIMA-GARCH model captures the dynamics of the daily rainfall series more precisely and the SARIMA model is the most suitable model for the monthly average rainfall series for the study locations.

Azad *et al.* (2015) analyzed monthly and annual rainfall data to predict the Indian monsoon rainfall. The periodic structure of the time-series data was extracted using wavelets, and the random residual part was separately modeled using ANN. Result shows that the estimated periodic and random components respectively comprise 30 and 15 per cent of the variance of the total rainfall in case of annual data, whereas, the model explains

93 per cent of the variance in the case of monthly data. It indicated that the prediction is more accurate when periodic and random parts are treated separately.

Osman and Abdellatif (2016) employed traditional ANN, Simple Average Method (SAM) and combining these models with Multiple Linear Regressions (MLR) and Generalized Linear Model (GLM)) to study rainfall site in North-western England. The model performance criteria of each of the primary and combining models were evaluated. The obtained results indicate that different downscaling methods can gain diverse usefulness and weakness in simulating various rainfall characteristics under different circumstances. The combining ANN model showed more adaptability by acquiring better overall performance.

Mehdizadeh *et al.* (2018) developed two novel types of hybrid models namely Gene Expression Programming-Autoregressive Conditional Heteroscedasticity (GEP-ARCH) and Artificial Neural Networks-Autoregressive Conditional Heteroscedasticity (ANN-ARCH) to estimate monthly rainfall time-series in Iran. For this, five stations with various climatic conditions were selected. The lagged monthly rainfall data were utilized to develop different GEP and ANN scenarios. The performance of proposed hybrid models were compared with individual GEP and ANN models using RMSE and coefficient of determination (R^2). The results showed that the proposed GEP-ARCH and ANN-ARCH models perform much better than the GEP and ANN for all of the stations under study. Furthermore, the ANN-ARCH model performed better in comparison with the GEP-ARCH model.

Namratha (2019) evaluated linear statistical models namely H-WES and SARIMA models for forecasting rainfall in meteorological subdivisions of Karnataka by using 50 years of monthly rainfall data from 1968 to 2017. Based on the lowest RMSE value revealed that H-WES model performed better than the SARIMA model in forecasting monthly rainfall of studied regions.

Pandey *et al.* (2019) employed hybrid model by combing SARIMA and GARCH models to eliminate conditional variance of the SARIMA model in two different climatic

environments (Agartala: humid, and Jodhpur: arid). The residuals from SARIMA models were tested for heteroscedasticity using the McLeod-Li test. Then, the rainfall time-series was transformed (differencing and Box-Cox) to eliminate heteroscedasticity. They suggest that the performance of SARIMA models was improved by using appropriate transformation (Box-Cox) along with the GARCH model residuals for highly skewed rainfall time-series from both climatic environments. They concluded that hybrid (SARIMA-GARCH) models fitted to transformed time-series rainfall data performed best in the humid as well as the arid regions.

III MATERIAL AND METHODS

This chapter consists of the details of the materials and the methods employed during the course of this investigation. To facilitate systematic presentation, this chapter is divided into the following three sections.

3.1 Description of the study area

3.2 Nature and source of data

3.3 Methodologies employed for analysis of the data

3.1 Description of the study area

Karnataka has tropical monsoon climate. The main characteristics of this climate are hot and moist summers, and cool and dry winters. However, there is diversity in the climate of the state. The main reasons for this diversity are the influence of geographical location, oceans, physical features, vegetation, and monsoon winds. Hence, the annual temperature and distribution of rain are not the same all over the state. Karnataka state is located between 11.40° to 18.27° N latitudes and 74.25° to 78.50° E longitudes and covers an area of 19.1 M Ha that accounts for 5.8 per cent of the total area of the country. It has a net cultivated area of 10.5 M Ha, net irrigated area of 3.49 M Ha and net rainfed area of 7.01M Ha. It has a 350 km long coastline, which forms the western boundary. The state is composed of 30 districts divided into 176 Taluk. These 176 Taluks are further divided into 747 Hobali. According to the 2011 Census, Karnataka has population of 6.11 Crores, an increase from the figure of 5.29 Crore in the 2001 census.

Based on the common rainfall distribution pattern, the Karnataka state is classified into four meteorological subdivisions as given below, which are depicted in the Fig. 3.1:

1. North Interior Karnataka (NIK)

2. South Interior Karnataka (SIK)

3. Coastal subdivision

4. Malnad subdivision

North Interior Karnataka subdivision comprises of 13 districts *viz*. Bagalkote, Ballari, Bidar, Kalburgi, Koppala, Bidar, Raichur, Vijayapura, Belagavi, Dharwad, Gadag, Haveri, and Yadgiri. This subdivision has the average annual rainfall about 731 mm.

South Interior Karnataka subdivision comprises 11 districts *viz*. Bengaluru Urban, Bengaluru rural, Chamarajanagara, Kolar, Mandya, Ramanagara, Chikkaballapura, Chitradurga, Davanagere, Mysore, and Tumkur. In this subdivision, the average annual rainfall is about 1286 mm.

Coastal subdivision comprises three districts Udupi, Uttar Kannada, and Dakshina Kannada. The average annual rainfall in this subdivision is about 3456 mm.

Malnad subdivision comprises four districts Chikmagalur, Hassan, Kodagu, and Shivamogga. The average yearly rainfall in this subdivision is about 2160 mm.

Karnataka experiences all the four types of climates *viz.* pre-monsoon (March to May), monsoon (June to September), post-monsoon (October to December) and winter (January to February). Karnataka receives highest amount of rainfall during the monsoon season.

Pre-monsoon: The state as a whole recorded about 194 mm rainfall during this season for the year 2018. The normal amount of rainfall for the season is 129 mm. It constitutes about 11 per cent of the annual rainfall. Fifty one per cent of rainfall departure from the actual rainfall was noticed. Hence, the state as a whole for the pre-monsoon season is considered under the excess rainfall category. The rainfall for the season varies from 67.30 mm in Raichur to 270 mm in Kodagu (Anon., 2017)

Monsoon: The state as a whole, rainfall recorded 804 mm during this season for the year 2018. The normal amount of rainfall for the season is 839 mm. The percentage of departure from actual rainfall is minus four per cent. The state as a whole for monsoon season is under the normal rainfall category. The highest amount of rainfall was received in the Udupi district (4071 mm) and the lowest amount of rainfall was observed in Chitradurga (276 mm). India receives about 80 per cent of rainfall only from Southwest monsoon.

Fig. 3.1: Map showing districts under four meteorological subdivisions of Karnataka

Post-Monsoon: The state as a whole for this season during 2018 recorded 96 mm. The normal amount of rainfall for the season is 188 mm. The percentage of departure from actual rainfall is -49 per cent (Anon., 2017). The state as a whole for monsoon season is under deficient rainfall category.

3.2 Nature and the source of data

The secondary data of monthly rainfall (mm) of four meteorological subdivisions of Karnataka v*iz.* North Interior Karnataka (NIK), South Interior Karnataka (SIK), Coastal and Malnad were collected from All India Coordinated Research Project (AICRP), Agro-Meteorology, UAS, GKVK, Bengaluru and Karnataka State Natural Disaster Monitoring Center (KSNDMC), Yelahanka, Bengaluru for the period of 60 years from 1960-2019. For time-series modelling and forecasting of rainfall, monthly rainfall data were used. For trend analysis and shifting pattern detection, monthly rainfall data was converted in to the following three different data sets:

1. **Month-wise data:** Month-wise rainfall of each month for the period of 1960-2019
2. **Seasonal data:** Total rainfall of each season *i.e.,*
    a. Winter (January and February)
    b. Pre-monsoon (March to May)
    c. Monsoon (June to September)
    d. Post-monsoon (October to December)
3. **Annual data:** Total rainfall occurred in all the months of a year *i.e.,* from January to December.

3.3 Methodologies employed for analysis of the data

3.3.1 To analyze the trend in the rainfall pattern in meteorological subdivisions

Trend analysis is one of the most commonly used tools for detecting changes in climatic and rainfall time-series. The subject of trend detection in rainfall data has received a great deal of attention at the later stage, especially in connection with the anticipated

changes in the global climate. However, climatic variability, which is reflected in rainfall data can adversely affect rainfall trend.

For trend analysis, parametric tests are more commonly used, which assumes that the data should be independent and normally distributed. However, real-world data may rarely satisfy these assumptions. To this end, more powerful nonparametric tests are employed to analyze the trend in the time-series data specially rainfall data, which does not require stringent assumption of parametric test, and can also tolerate the effect of outliers in the data. Several nonparametric tests are available in literature to analyze the trend. Mann-Kendall's (M-K) test is one of the most commonly and widely used nonparametric test for estimating the trend in rainfall data, test procedure is given below:

3.3.1.1 Mann-Kendall (M-K) test

Mann-Kendall trend (M-K) test is a nonparametric test, which is an alternative method to the parametric method of trend analysis. It is the most suitable test for detecting the trend for rainfall data. Since there will be high fluctuation present in the weather parameters, the non-parametric M-K test is useful as its statistic is based on the sign of differences, not directly on the values of random variables and therefore trend determined is less affected by the fluctuations and extreme rainfall values.

Test procedure:

Null Hypothesis (H_0) = There is no presence of a monotonic trend in the rainfall data.

Alternative Hypotheses (H_1) = There is a presence of a monotonic trend (either decreasing or increasing) in the rainfall data.

The Mann-Kendall statistics *'S'* is given as

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} sign(x_j - x_k) \qquad (1)$$

where

$$sign(x_j - x_k) = \begin{cases} +1 & if \quad (x_j - x_k) > 0 \\ 0 & if \quad (x_j - x_k) = 0 \\ -1 & if \quad (x_j - x_k) < 0 \end{cases}$$

and x_k and x_j are the sequential data values and for all $k, j \leq n$ with $k \neq j$ and n is the length of the rainfall data set.

The mean and variance of S are given by:

$$\mathrm{E(S)} = 0 \text{ and}$$

$$\mathrm{Var(S)} = \frac{1}{18}[n(n-1)(2n+5)]$$

If ties are present in the data, then the variance of S is as follows

$$Var(S) = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{i=1}^{n} t_i(i)(i-1)(2i+5)\right]$$

Where, t_i is the number of ties of extending i (Sonali and Kumar, 2013).

Tau measures the strength of the monotonic relationship between x_i and x_{i+1}. Kendall's tau correlation coefficient is given by,

$$\tau = \frac{S}{n(n-1)/2}$$

If sample size $n > 10$, the standard normal variate z is computed:

$$z = \begin{cases} \dfrac{S-1}{\sqrt{Var(S)}} & if \quad S > 0 \\ 0 & if \quad S = 0 \\ \dfrac{S+1}{\sqrt{Var(S)}} & if \quad S < 0 \end{cases}$$

In a two-sided test for trend, H_0 is accepted if $|z| \leq z_\alpha/2$ at the α level of significance. A positive value of S indicates an 'upward trend', likewise a negative value of S indicates 'downward trend' and a zero value of S indicates 'no trend'.

3.3.1.2 Modified Mann-Kendall (MM-K) test by Yue and Wang

Even though M-K test is most commonly used test for detecting trend in rainfall data, it assumes that sample data should be serially independent. However, it is well known that from many previous studies, most of rainfall time-series data exhibit serial correlation. The presence of serial correlation in time-series will alter the variance of the M-K test statistic which in turn will affect the ability of the test to assess the significance of the trend correctly because the M-K test and the Theil Sen tests are unable to consider the effect of AR(1) process of the time-series (Azam *et al.*, 2018). The presence of positive autocorrelation in the data increases the probability of detecting trend even though actual data have no trend, and vice versa. To this end, Yue and Wang (2004) developed Modified Mann-Kendall (MM-K) test, which eliminates the effect of serial correlation present in the time-series data on the M-K test statistic by correcting the variance using Effective Sample Size (ESS). The accuracy of the modified test in terms of its empirical significance level was found to be superior to that of the original Mann-Kendall trend test without any loss of power.

Therefore, the modified variance $V^*(S)$ using ESS is given by:

$$V^*(S) = V(S) \cdot \frac{n}{n^*}$$

Where *n* is the Actual Sample Size (ASS) of data, n/n^* is termed the Correction Factor (C.F) and n^* is the ESS, proposed by Lettenmaier (1976) computed by:

$$n^* = \frac{n}{1 + 2 \cdot \sum_{k=1}^{n-1}(1 - \frac{k}{n}) \cdot \rho_k}$$

Where ρ_k is the lag-*k* serial correlation coefficient, which can be estimated by the sample lag-*k* serial correlation coefficient (r_k) given by:

$$r_k = \frac{\frac{1}{n-k}\sum_{t=1}^{n-k}(x_t - \bar{x}_t)(x_{t+k} - \bar{x}_t)}{\frac{1}{n}\sum_{t=1}^{n}(x_t - \bar{x}_t)^2}$$

$$\bar{x}_t = \frac{1}{n}\sum_{t=1}^{n} x_t$$

Next variance of M-K test is replaced by modified variance and proceed with the M-K test procedure.

3.3.1.3 Sen's Slope Estimator

Sen's slope estimator has been widely used for determining the magnitude of a trend. The Sen's slope estimator is a linear slope estimator that works most effectively on monotonic data. Unlike linear regression, it is not greatly affected by data errors, outliers, or missing data. Here, the slope (T_i) of all data pairs can be computed by:

$$T_i = \frac{x_j - x_k}{j - k} \text{ for } i = 1,2, \dots, n$$

where x_j and x_k are considered as data values at time j and k ($j > k$) correspondingly. The median of these n values of T_i is represented as Sen's estimator of the slope, which is given as:

$$Q_{Med} = \begin{cases} T_{\frac{N+1}{2}} & if\ N\ is\ odd \\ \frac{1}{2}\left(T_{\frac{N}{2}} + T_{\frac{N+2}{2}}\right) & if\ N\ is\ even \end{cases}$$

Sen's estimator is computed as $Q_{med} = T_{(N+1)/2}$ if N appears odd, and it is considered as $Q_{med} = \left[T_{N/2} + T_{(N+2)/2}\right]/2$ if N appears even. In the end, T_{med} is computed by a two-sided test at $100(1-\alpha)\%$ confidence interval and then a true slope can be obtained by this non-parametric test.

Positive value of Q_{Med} indicates an upward or increasing trend and a negative value of Q_{Med} gives a downward or decreasing trend in the time-series (Mondal *et al*, 2012).

3.3.2 To study the shifting pattern in the rainfall distribution

Many time-series are characterized by abrupt shift in structure, such as sudden jumps in level or volatility. We consider change/shift point to be those time points that divide data set into distinct homogeneous segments. In practice, the number of shifting points will not be known. The ability to detect shifting point is important for both methodological and practical reasons including the validation of an untested scientific

hypothesis, monitoring and assessment of safety-critical processes and the validation of modeling assumptions.

Few non-parametric tests like Standard Normal Homogeneity Test (SNHT), Buishand Range (BR) test and Pettitt test can be used to test the null hypothesis, values Y_i of the testing variables Y are independent and identically distributed, and the series is considered as homogeneous. Meanwhile, under the alternative hypothesis, SNHT, BR test and Pettitt test assume that the series consist of break in the mean and considered as heterogeneous. These three tests are capable to detect the year where the break occurs. As an alternative to these non-parametric tests many parametric tests have been proposed by different authors *i.e.,* Bayesian approaches, Penalizing methods, Likelihood-Ratio estimate etc. Among those methods, the Likelihood-Ratio method is a popular and most powerful one.

The potential for using likelihood-based approach to detect shifting point was first proposed by Hinkley (1970) and he derived the asymptotic distribution of the likelihood ratio test statistic for change in the mean within the sequence of normally distributed observations.

3.3.2.1 Likelihood-Ratio Test

Let us consider that each x_i is normally distributed with a mean μ_i and common variance $\sigma^2, i = 1,2, \dots, n$. The interest here is about the mean shift. The hypothesis of stability or no shift point (the null hypothesis) is defined as:

$$H_0: \mu_1 = \mu_2 = \cdots = \mu_n = \mu$$

v/s

$$H_1: \mu_1 = \cdots = \mu_k \neq \mu_{k+1} = \cdots = \mu_n$$

where k is the unknown location of the shift point. The testing procedure depends on whether the nuisance parameter σ^2 is known or unknown (Chen and Gupta, 2011).

i) If Variance Known:

Without loss of generality, assume that $\sigma^2 = 1$. Under H_0, the likelihood function is

$$L_0(\mu) = \frac{1}{(\sqrt{2\pi})^n} e^{-\sum_{i=1}^{n}(x_i-\mu)^2/2}$$

and the Maximum Likelihood Estimator (MLE) of μ is

$$\hat{\mu} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Under H_1, the likelihood function is

$$L_1(\mu_1, \mu_n) = \frac{1}{(\sqrt{2\pi})^n} e^{-\sum_{i=1}^{k}(x_i-\mu_1)^2+\sum_{i=k+1}^{n}(x_i-\mu_n)^2/2}$$

and the MLEs of μ_1 and μ_n are respectively,

$$\hat{\mu}_1 = \bar{x}_k = \tfrac{1}{k}\textstyle\sum_{i=1}^{k} x_i \quad \text{and} \quad \hat{\mu}_n = \bar{x}_{n-k} = \tfrac{1}{n-k}\textstyle\sum_{i=k+1}^{n} x_i.$$

Let

$$S_k = \sum_{i=1}^{k}(x_i - \bar{x}_k)^2 + \sum_{i=k+1}^{n}(x_i - \bar{x}_{n-k})^2,$$

$$V_k = k(\bar{x}_k - \bar{x})^2 + (n-k)(\bar{x}_{n-k} - \bar{x})^2,$$

and $S = \sum_{i=1}^{n}(x_i - \bar{x})^2$; then $V_k = S - S_k$(Chen and Gupta, 2011). Simple algebra leads to $U^2 = V_{k*} = \max_{1\le k\le n-1} V_k$ is the likelihood procedure test statistic for testing H_0 against H_1.

Hawkins (1977) derived the exact and asymptotic null distributions of the test statistic *U*. The following is based on his work. First, simple algebraic computation gives an alternative expression for V_k as

$$V_k = \tfrac{n}{k(n-k)}\left[\textstyle\sum_{i=1}^{k}(x_i - \bar{x})\right]^2.$$

Let,

$$T_k = \sqrt{\frac{n}{k(n-k)}}\left[\sum_{i=1}^{k}(x_i - \bar{x})\right];$$

Then $V_k = T_k^2$ or $|V_k| = \sqrt{V_k}$. Therefore,

$$U = \sqrt{V_{k*}} = \min_{1\le k\le n-1} \sqrt{V_k} = \max_{1\le k\le n-1} |T_k|$$

is the equivalent likelihood-based test statistic for testing H_0 against H_1.

The asymptotic null distribution of U is a Gumbel distribution (Yao and Davis, 1986) and $T_k \sim N(0,1)$.

Then the null hypothesis (that there is no shift point) is rejected in favor of the alternative (there is a single shift point) if $U > c$ where c is some threshold value.

This threshold value, c, is usually obtained such that $\alpha = \Pr(U > c | H_0 \; is \; true)$ for some significance level α.

ii) If Variance Unknown:

Under H_0, the likelihood function is

$$L_0(\mu, \sigma^2) = \frac{1}{(\sqrt{2\pi}\sigma)^n} e^{-\sum_{i=1}^{n}(x_i-\mu)^2/2\sigma^2}$$

and the Maximum Likelihood Estimators (MLEs) of μ and σ^2 are

$$\hat{\mu} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \text{and} \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2,$$

respectively. Under H_1, the likelihood function is

$$L_1(\mu_1, \mu_n, \sigma^2) = \frac{1}{(\sqrt{2\pi})^n} e^{-\sum_{i=1}^{k}(x_i-\mu_1)^2/2\sigma_1^2 - \sum_{i=k+1}^{n}(x_i-\mu_n)^2/2\sigma_1^2}$$

and the MLEs of μ_1, μ_n and σ_1^2 are,

$$\hat{\mu}_1 = \bar{x}_k = \frac{1}{k}\sum_{i=1}^{k} x_i, \qquad \hat{\mu}_n = \bar{x}_{n-k} = \frac{1}{n-k}\sum_{i=k+1}^{n} x_i \text{ and}$$

$$\hat{\sigma}_1^2 = \frac{1}{n}\left[\sum_{i=1}^{k}(x_i - \bar{x}_k)^2 + \sum_{i=k+1}^{n}(x_i - \bar{x}_{n-k})^2\right],$$

respectively.

Let,

$$S = \sum_{i=1}^{n}(x_i - \bar{x})^2 \text{ and } T_k^2 = \frac{k(n-k)}{n}(\bar{x}_k - \bar{x}_{n-k})^2.$$

The likelihood procedure-based test statistic is then given by

$$V = \max_{1 \le k \le n-1} \frac{|T_k|}{S}$$

Under $H_0, T_k \sim N(0, \sigma^2)$, $S_k^2 = \sum_{i=1}^{k}(x_i - \bar{x}_k)^2 + \sum_{i=k+1}^{n}(x_i - \bar{x}_{n-k})^2$ is distributed as $\sigma^2\chi_{n-2}^2$, and T_k is independent of S_k; then $T_k/[S_k/\sqrt{n-2}] \sim \chi_{n-2}^2$. However, simple algebra shows that $V = T_k/S \sim t_{n-2}$. Asymptotic distribution of V is t-distribution with $n-2$ degrees of freedom as $n \to \infty$ (Hawkins, 1977).

Then the null hypothesis (that there is no shift point) is rejected in favor of the alternative (there is a single shift point) if $V > c$ where c is some threshold value.

This threshold value 'c' is usually obtained such that $\alpha = \Pr(V > c | H_0 \; is \; true)$ for some significance level α. If the null hypothesis is rejected at some point *k*, then that point *k* is considered as rainfall pattern shifting point (year) and mean rainfall for before and after shifting year is calculated.

3.3.3 To evaluate a suitable non-linear time-series models for forecasting rainfall.

3.3.3.1 Exponential smoothing method of forecasting

The monthly rainfall data of all four meteorological subdivisions are used to evaluate statistical models for forecasting rainfall data. The time-series comprises of

successive elements that are dependent linearly on the past observations and the error component. The study of time-series just does not help only in understanding the phenomenon that is represented by a series of observations but also helps in predicting the future. Exponential smoothing is such a type of approach in time-series analysis for forecasting. This model gives more weight to recent observations when compared to past observations.

Forecasting procedures based on exponential smoothing have become popular since they are easy to implement and can be quite effective. They can be implemented without respect to a properly defined statistical model. Exponential Smoothing is an averaging technique that uses unequal weights; however, the weights applied to past observations decline exponentially.

3.3.3.1.1 Simple exponential smoothing

When the rainfall data has no trend and seasonality, the single exponential smoothing method can be used. In this model, the value of forecast for the time $t+1$ depends on the value of rainfall and the forecast value at time *t*. The model to forecast for the time $t+1$ is given below;

The method of single exponential forecasting takes the forecast for the previous period and adjusts it using the forecast error $(Y_t - F_t)$.

$$F_{t+1} = F_t + \alpha(Y_t - F_t) \quad \text{or} \quad F_{t+1} = \alpha Y_t + (1-\alpha)F_t$$

where, Y_t = Actual rainfall at time *t*

F_t = Forecasted value at time *t*

α = Smoothing parameter ranges from 0 to 1

t = The current time period

Choosing α has a great impact on the prediction. It can be done from the grid value $(\alpha = 0.1, 0.2, \ldots, 0.9)$ such that it yields minimum RMSE value. If the value of α is small it pays more attention to past observation and, if the value of α is close to 1 it means that it pays more attention to the recent observations.

Knowing the value of α and past observation (Y_t), to find the first forecast value F_1 the initialization process can be done by equating $Y_1 = F_1$.

3.3.3.1.2 Double Exponential Smoothing (Holt's linear method)

In 1957 Holt extended single exponential smoothing to linear exponential smoothing to allow forecasting of data with trend *i.e.,* when the data contains trend then the Double Exponential Smoothing is preferable. The forecast for Holt's linear exponential smoothing is found using two smoothing constants, α and β (with values between 0 and 1), and three equations one for the level, one for trend and one for seasonality. The equations for the models are given as follows

$$L_t = \alpha Y_t + (1-\alpha)[L_{t-1} + b_{t-1}] \quad (3.1)$$

$$b_t = \beta[L_t - L_{t-1}] + (1-\beta)b_{t-1} \quad (3.2)$$

$$F_{t+m} = L_t + mb_t \quad (3.3)$$

where,

L_t = Level at time *t*

b_t =Trend at time *t*

F_{t+m} = Forecast value for *m* period ahead

α, β = Smoothing parameters ranging from 0 to 1. The combination of these parameters is selected based on minimum RMSE value.

Here, L_t denotes an estimate of the level of the series at time *t* and b_t denotes an estimate of the slope of the series at time *t*. Equation (3.1) adjusts (L_t) directly for the trend of the previous period, b_{t-1}, by adding it to the last smoothed value, L_{t-1}. This helps to eliminate the lag and brings L_t to the approximate level of the current data value. Equation (3.2) then updates the trend, which is expressed as the difference between the last two smoothed values. This is an appropriate difference between the last two smoothed values. This is appropriate because if there is a trend in the data, new values should be higher or lower than previous ones. Since there may be some randomness remaining, the trend is modified by smoothing with β the trend in the last period $(L_t - L_{t-1})$, and adding that to

the previous estimate of the trend multiplied by $(1-\beta)$. Thus, equation (3.2) is similar to the basic form of single smoothing but applied to the updating of the trend. Finally, equation (3.3) is used to forecast. The trend, b_t, is multiplied by the number of periods ahead to be forecast, *m*, and added to the base value, L_t (Makridakis *et al*., 1998).

The initialization process for Holt's linear exponential smoothing requires two estimates, one to get the first smoothed value for $L_1$and the other to get the trend b_1. One alternative is to set $L_1 = Y_1$ and $b_1 = Y_2 - Y_1$or $b_1 = (Y_4 - Y_1)/3$.Another alternative is to use least squares regression on the first few values of the series for finding L_1 and b_1.

3.3.3.1.3 Holt-Winter's Exponential Smoothing (H-WES) method

The moving average methods and exponential smoothing method deal with almost any type of data as long as such data are non-seasonal. When seasonality does exist, however, these methods are not appropriate on their own (Makridakis *et al*., 1998).

Holt's method was extended by Winter to capture seasonality directly. The Holt-Winter's method is based on three smoothing equations, one for level, one for trend, and one for seasonality. It is similar to Holt's method, with an additional equation to deal with seasonality. Holt-Winter's Exponential Smoothing (H-WES) methods are widely used when the data shows trend and seasonality. It has two types of models, based on the seasonality variations one is the Holt-Winter's additive model (additive trend, additive seasonality) and another Holt-Winter's multiplicative model (additive trend, multiplicative seasonality). If the seasonal variation is constant over time, "Holt-Winter's Additive model" is used and if the seasonality varies over time, "Holt-Winter's Multiplicative model" is used. In the present study, the Additive model is used as the seasonal variation over time is observed (not constant). The four equations for the model are given as follows:

$$L_t = \alpha(Y_t - S_t) + (1-\alpha)(L_{t-1} + b_{t-1})$$

$$b_t = \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1}$$

$$S_t = \gamma(Y_t - L_t) + (1-\gamma)S_{t-s}$$

$$F_{t+m} = L_t + mb_t + S_{t-s+m}$$

where s = the length of seasonality

L_t =Level at time t

b_t = Trend at time t

S_t = Seasonal component at time t

F_{t+m} = Forecast value for m period ahead

$\alpha, \beta \; and \; \gamma$ are level, trend and seasonal smoothing constants or the weights respectively, which lies between 0 and 1. The combination of these parameters is selected based on minimum RMSE and MAPE value.

3.3.3.2 Box-Jenkins time-series model for forecasting:

The Box-Jenkins procedure is concerned with fitting a mixed Auto-Regressive Integrated Moving Average (ARIMA) model to a given set of data. The main objective of fitting the ARIMA model is to identify the stochastic process of the time-series and predict future values accurately. Originally, ARIMA models have been studied extensively by George Box and Gwilym Jenkins and their names have been frequently used synonymously with the general ARIMA process applied to time-series analysis and forecasting. However, the optimal forecast values of given time-series are determined by the stochastic process of that series. A stochastic process is either stationary or non-stationary. The first thing to note is that most time-series are non-stationary and the ARIMA model refers only to a stationary (Makridakis *et al*., 1998).

Stationarity and non-stationarity: The term stationarity means that the process generating the data is in equilibrium around a constant value and that the variance around the mean remains constant over time. The data must be roughly horizontal along the time axis.

If mean changes over time (with some trend cycle pattern) and variance are not reasonably constant, then series is said to be non-stationary in both mean and variance. It is very difficult to find a stationary series in practical either mean shift with time because of several factors like technology, demand, population, etc., or the variance changes over

time. However, the method of differencing is used to make the series stationary. Conversely, a stationary series may be summed or integrated to give a non-stationary one.

d^{th} order difference

$$(1-B)^d X_t$$

The general $ARIMA\ (o, d, o)$ model will be

$$(1-B)^d X_t = e_t$$

where $e_t$is error term distributed normally with

$$E(e_t) = 0, V(e_t) = \sigma_t^2 \text{and } CoV(e_i, e_j) = 0 \text{ for all } t\ (i \neq j).$$

To test the stationarity, compute the Autocorrelation Functions (ACF) of differenced series (Y_t) up to 36 lags. If the ACF for first and higher differences (after 2-3 lags) drops abruptly to zero, then it indicates that the series is stationary.

Some important stationary time-series model:

3.3.3.2.1 Autoregressive model with the order p: AR (*p*)

A stochastic model that can be extremely useful in the representation of certain practically occurring series is the autoregressive model. In this model, the current value of the process is expressed as a finite linear aggregate of previous values of the process and a shock e_t.

Let us denote the values of a process at equally spaced time epochs (periods) t, t -1, t-2, … by y_t, y_{t-1}, y_{t-2}, …., then y_t can be described as

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \varepsilon_t$$

If we define an autoregressive operator of order *p* by

$$\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p$$

where B is the backshift operator such that $BY_t = Y_{t-1}$, the autoregressive model can be written as $\varphi(B)y_t = \varepsilon_t$.

3.3.3.2.2 Moving Average (MA) model of order q: MA (q)

Another kind of model of great practical importance in the representation of observed time-series is the finite moving average process. MA (q) model is defined as:

$$y_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \cdots - \theta_q\varepsilon_{t-q}$$

If we define a moving average operator of order q by

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$$

where B is the backshift operator such that $BY_t = Y_{t-1}$, moving average model can be written as $y_t = \theta(B)\varepsilon_t$.

3.3.3.2.3 Autoregressive Moving Average (ARMA) model: ARMA (p, q)

To achieve greater flexibility in the fitting of time-series data, it is sometimes advantageous to include both autoregressive and moving average processes. This leads to a mixed autoregressive-moving average model

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \cdots - \theta_q\varepsilon_{t-q}$$

$$y_t = \sum_{i=1}^{p} \varphi_i y_{t-i} - \sum_{j=1}^{q} \theta_j y_{t-j} + \varepsilon_t$$

or

$$\varphi\varepsilon_t = \theta(B)\varepsilon_t$$

and is written as ARMA(p, q). In practice, it is quite often an adequate representation of actually occurring stationary time-series can be obtained with autoregressive, moving average, or mixed models, in which p and q are not greater than 2 and often less than 2.

3.3.3.2.4 Autoregressive Integrated Moving Average (ARIMA) model: ARIMA (*p, d, q*)

A generalization of ARMA models that incorporates a wide class of non-stationary time-series is obtained by introducing the differencing into the model. The simplest example of a non-stationary process that reduces to a stationary one after differencing is a random walk. A process $\{y_t\}$ is said to follow an Integrated ARMA model, denoted by ARIMA (*p*, *d*, *q*), if

$$\nabla^d y_t = (1-B)^d \varepsilon_t \text{ is ARMA}(p, q).$$

The model is written as

$$\varphi(B)(1-B)^d y_t = \theta(B)\varepsilon_t$$

where, $\varepsilon_t \sim WN(0, \sigma^2)$, *WN* indicating White Noise. The integration parameter *d* is a non-negative integer. When *d* = 0, ARIMA (*p*, *d*, *q*) = ARMA (*p*, *q*).

3.3.3.2.5 Seasonal Autoregressive Integrated Moving Average (SARIMA) Model:

Sometimes the series exhibit perceptible periodic pattern for instance, price and arrivals of agricultural commodities usually have a seasonal pattern process in general.

Then ARIMA notation can be extended readily to handle seasonal aspects. In its general form, the Seasonal ARIMA model is characterized by a notation as ARIMA (*p, d, q*) $(P, D, Q)_s$ is given by

$$(1-\varphi_p B)\left(1-\Phi_p B^s\right)(1-\mathrm{B})(1-B^s)y_t = (1-\theta_q B)(1-\Theta_q B^s)\varepsilon_t$$

where *B* is the backshift operator (*i.e.*, $By_t = y_{t-1}, B^2 y_t = y_{t-2}$ and so on), '*s*' the seasonal lag and 'ε_t' and '*t*' a sequence of independent normal error variables with mean 0 and variance σ^2. $\Phi's$ and $\varphi's$ are respectively the seasonal and non-seasonal autoregressive parameters. $\Theta's$ and $\theta's$ are respectively seasonal and non-seasonal moving average parameters. The order p and q are orders of non-seasonal autoregressive and moving average parameters respectively, whereas P and Q are that of the seasonal

autoregressive and moving average parameters respectively. Also '*d*' and '*D*' denotes non-seasonal and seasonal differences respectively (Makridakis *et al*., 1998).

Table 3.1: Types of ARIMA model with its parameters

ARIMA order	Name
ARIMA (0,0,0)	White noise
ARIMA (0,1,0) without constant	Random walk
ARIMA (0,1,0) with constant	Random walk with drift
ARIMA (p,0,0)	Autoregressive model
ARIMA (0,0,q)	Moving average model
ARIMA (p,0,q)	ARMA model
ARIMA (p,d,0)	Differenced order autoregressive model
ARIMA (0,d,q) without constant	Exponential smoothing
ARIMA (0,d,q) with constant	Exponential smoothing with growth

3.3.3.3 Autoregressive Conditional Heteroscedasticity (ARCH) family models

The most widely used technique for analysis of time-series data, undoubtedly, the Box-Jenkins ARIMA methodology. However, it is based on some crucial assumptions, like linearity, stationarity and homoscedastic errors. Further, time-series data quite often exhibit features that cannot be explained by the ARIMA model, which is 'linear'. Many weather data show periods of stability followed by unstable periods with high volatility. The loss in continuing to use the ARIMA methodology is that this type of behavior cannot be explained satisfactorily, and so 'nonlinear time-series models' are usually needed to describe the time-series data in which variance changes through time, thus study uses ARCH family models. If there is a volatility or ARCH effect in the time-series data, we can run the ARCH family models. To check the ARCH effect in the times series data, the ARCH LM Test was used.

The ARCH LM Test

Let $\varepsilon_t = y_t - \theta y_{t-1}$ be the residual series. The conditional heteroscedasticity, also known as ARCH effects, is tested on the squared series $\{\varepsilon_t^2\}$. Engle (1982) gave a test called the Lagrange Multiplier test to assess the significance of the ARCH effects. This test is equivalent to usual F-statistics for testing H_0: $a_i = 0, i = 1,2, \dots, q$ in the linear regression

$$\varepsilon_t^2 = a_0 + a_1\varepsilon_{t-1}^2 + \cdots + a_q\varepsilon_{t-q}^2 + e_t, t = q+1, \dots, T$$

Where e_t denotes the error term, q is the pre-specified positive integer, and T is the sample size.

Let,

$$SSR_0 = \sum_{t=q=1}^{T} (\varepsilon_t^2 - \bar{\omega})^2$$

where

$$\bar{\omega} = \sum_{t=q+1}^{T} \varepsilon_t^2 /_T$$

is the sample mean of $\{\varepsilon_t^2\}$, and

$$SSR_1 = \sum_{t=q+1}^{T} \hat{e}_t^2$$

where, $\hat{e}_t^2$ is the least-squares of the residuals of the regression equation. Then, under H_0,

$$F = \frac{(SSR_0 - SSR_1)/q}{SSR_1(T-q-1)}$$

is asymptomatically distributed as chi-squared distribution with q degrees of freedom. The decision rule is to reject H_0 if $F > \chi_q^2(\alpha)$ alternatively, the p-value of F is less than α.

ARCH Model

The most promising parametric nonlinear time-series model is Autoregressive Conditional Heteroscedasticity (ARCH) model, which was introduced by Engle (1982). It was one of the first models that provided a way to model conditional heteroscedasticity in

volatility. The ARCH model allows the conditional variances to change over time as a function of squared past errors leaving the unconditional variance constant.

ARCH models assume the variance of the current error term or innovation to be a function of the actual sizes of the previous time period's error terms: often the variance is related to the squares of the previous innovations (Toth *et al.*, 2000).

The ARCH (q) model for the series $\{\varepsilon_t\}$ is defined by specifying the conditional distribution of ε_t given the information available up to time *t-1*. ARCH (q) model for the series $\{\varepsilon_t\}$ is given by

$$\varepsilon_t = \sigma_t^2 \xi_t, \qquad \{\xi_t\} \sim iid(0,1)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2$$

where, $\alpha_0 > 0$ and $\alpha_i \geq 0$, for all i and $\sum_{i=1}^{q} \alpha_i < 1$ are required to be satisfied to ensure non-negative and finite unconditional variance of stationary $\{\varepsilon_t\}$ series.

3.3.3.4 GARCH Model

The ARCH model has some drawbacks. Firstly, when the order of ARCH model is very large, estimation of a very large number of parameters is required. Secondly, conditional variance of ARCH (q) model has the property that unconditional autocorrelation function of squared residuals, if exists, decays very rapidly compared to what is typically observed, unless maximum lag q is large. To overcome these difficulties, Bollerslev (1986) proposed the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model in which conditional variance is also a linear function of its own lags. This model is also a weighted average of past squared residuals, but it has declining weights that never go completely to zero. It gives parsimonious models that are easy to estimate and, even in its simplest form, has proven surprisingly successful in predicting conditional variances. The GARCH (p, q) model for the series $\{\varepsilon_t\}$ is given by

$$\varepsilon_t = \sigma_t^2 \xi_t, \qquad \{\xi_t\} \sim IID(0,1)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \sigma_{t-j}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$$

where, $\alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0, \sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1, \alpha_i = 0 \; for \; i > q \; and \; \beta_j = 0 \; for \; j > p.$

The properties of the GARCH model is quite similar to that of the ARCH but requires far less parameters to adequately model the volatility process. The GARCH model is able to model the volatility clustering but does not address the problem with the ARCH models lack of ability to model the asymmetric effect of positive and negative values. The GARCH model also imposes restrictions on the parameters to have a finite fourth moment as was the case for the ARCH model. The GARCH model is similar to the ARCH model but with the addition of lagged conditional variances, σ_{t-j}^2 as well as the lagged squared values, σ_{t-j}^2. The addition of the lagged conditional variances avoids the need for adding many lagged squared values, as was the case for the ARCH model to be able to appropriately model the volatility. This greatly reduces the number of parameters that need to be estimated. In fact, for the GARCH (1,1) model, the conditional variance can be rewritten as

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

The main stages in setting up a Box-Jenkins and volatility models for forecasting are as follows:

Step 1: Determine stationary of time-series

The series under study must be stationary. A stationary time-series has the property that its statistical properties such as the mean and variance are constant over time. Stationarity in the data can be checked by plotting the Autocorrelation Function (ACF) and Partial-Autocorrelation Function (PACF). In this study, the monthly rainfall data is seasonally differenced to make it stationary in mean.

Step 2: Identify the most suitable SARIMA model

By looking into significant lags in ACF and PACF plots of stationary rainfall data, tentative SARIMA models of seasonal and non-seasonal order parameters *p, q,* and *P, Q* are selected. The most suitable SARIMA model is selected using the minimum Akaike Information Criterion (AIC) given by

$$AIC = (n * \ln(SSR/n) + 2k)$$

where, *n* is number of observations, *SSR* is sum of squared residuals and *k* is number of parameters to be estimated.

Step 3: Estimate the model parameters and diagnostic checking

For selected best suitable SARIMA model, the parameters are estimated via MLE method to minimize overall error calculation or maximize the likelihood function and significance of estimated parameters are checked.

To check if the model assumptions about the errors are satisfied or not portmanteau test is performed. The test checks whether the best-fitted model residuals are white noise or not, the Ljung-Box statistic is computed under the Null hypothesis that there is no autocorrelation in the residuals.

The Ljung-Box statistic is given by:

$$Q = n(n+2)\sum_{k=1}^{h}(n-k)^{-1}r_k^2 \sim \chi^2_{(h-m)\ d.f}$$

where, *h* is the maximum lag, *n* is the number of observations, *m* is the number of parameters in the model. Q is distributed approximately as a Chi-square statistic with (*h-m*) degree of freedom. If the p-value associated with the Q statistic is small (p-value<0.05 or 0.01), the model is considered inadequate.

Step 4: Determination of residuals and heteroscedasticity test:

Selected best mean model (SARIMA) residuals are obtained and new variable called “r-square” by squaring the residuals is formed. Then the ACF and PACF values of

the “r-square” are determined and the lags in which these values are found to be significant are identified. The test for heteroscedasticity is done at identified significant lags. The test employed is the ARCH-LM test. If heteroscedasticity is found then the best ARCH and GARCH model is selected based on the lowest AIC criterion, among the available class of ARCH and GARCH models.

Step 5: Selecting best model for forecasting

The principal objective of developing any time-series model for a variable is to generate post sample period forecast for the same variable. The ultimate test for any model is whether it is capable of predicting future events accurately or not.

Selected best SARIMA, ARCH and GARCH models are used to predict and forecast the rainfall. For selecting best fitting and forecasting model the accuracy of prediction and forecasts was tested using Root Mean Square Error (RMSE) given by:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Y_t - \hat{Y}_t)^2}$$

where $Y_t$the actual rainfall is value and $\hat{Y}_t$ is the predicted or forecasted rainfall value. Smaller the value of RMSE more efficient is the model.

3.3.6 Artificial neural network models

Box-Jenkin’s ARIMA and its variant SARIMA models have dominated univariate time-series analysis during last several decades. However, the limitation of these models is linearity, which may not hold true in reality (Mohan Kumar and Prajneshu, 2014). Thus, in literature wide ranges of models have been proposed with different mathematical representations of non-linearity present in the data. One of such most common and widely used non-linear model is Artificial Neural Networks (ANN). The first step toward artificial neural networks came in 1943 when Warren McCulloch, neurophysiologist and Walter Pitts, mathematician wrote a paper on how neurons might work. They modeled a simple neural network with electrical circuits for the first time.

The Artificial Neural Network (ANN) is a data driven, self-adaptive, nonlinear nonparametric statistical method. ANN functions similar to the human brains. ANN models are the powerful tool for modelling, especially when the underlying data relationship is unknown. Its modelling has attracted a new technique for estimation and forecasting in many fields of study including agriculture, economics and statistics. Researchers get attracted to ANN due to its freedom from restrictive assumptions such as linearity that is often needed to make the traditional mathematical model useable. ANNs have become the focus of much attention, largely because of their wide range of applicability and the ease with which they can treat complicated problems. A very important feature of this network is its learning capability by examples.

A simple Multi-Layer Perceptron Neural Network (MLP-NN) with three input nodes, five hidden nodes and one output node are displayed in Fig 3.2.

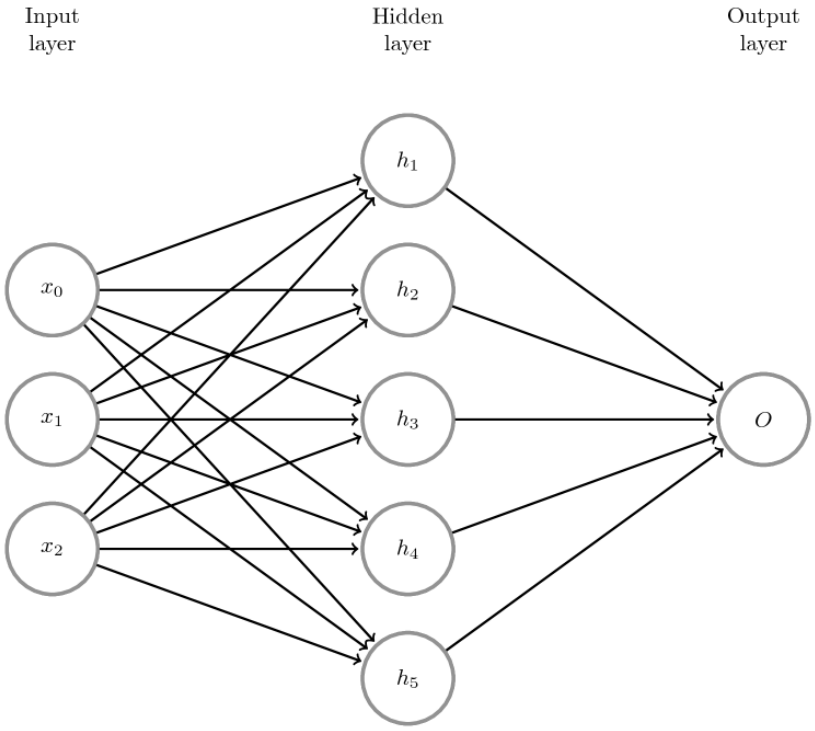

Fig. 3.2: Diagram of Multi-Layer Perceptron Neural Network (MLP-NN)

3.3.6.1 Development of an ANN model

The ANNs are generally termed as multilayer ANNs as layers of nodes are used to construct them, such that each node in a layer performs a similar task. The very first layer consists of the input nodes, which are in statistical term known as the independent variables. Last layer contains the output nodes, which are in statistical term known as the dependent or response variables. The all other remaining nodes in the model are known as the hidden nodes and constitute the hidden layers. There are two functions, which govern the behavior of a node in a particular layer, which are normally same for all nodes within the whole ANN, *i.e.*,

1. The input function
2. The output/activation function.

Input into a node is a weighted sum of outputs from nodes connected to it. Thus net input into a node *i* is

$$netinput_i = \sum_j w_{ij} * output_j + \mu_i$$

where $w_{ij}$are weights connecting neuron *j* to neuron *i*, $output_j$ is output from node *j* and μ_i is a threshold for neuron *i*. Threshold term is baseline input to a node in absence of any other inputs. If a weight $w_{ij}$is negative, it is termed inhibitory as it decreases net input; otherwise, it is known excitatory. Each node takes its net input and applies an activation function to it. For example, output of j^{th} node, also called activation value of the node, is $g(\sum w_{ij}x_i)$, where $g(.)$is activation function and $x_i$isoutput of i^{th} node connected to node *j*.

A number of nonlinear functions have been used as activation functions by many researchers. The activation function exhibits a great variety and has the biggest impact on behavior and performance of the ANN. The main task of the activation function is to map the outlying values of the obtained neural input back to a bounded interval such [0, 1] or [-1, 1]. The sigmoid function is most commonly used activation function as it has some major advantages due to its differentiability within the context of finding a steepest descent

gradient for the back propagation method and moreover maps a wide domain of values in the interval [0, 1].

With no hidden nodes, an ANN can classify only linearly separable. However, it has been proved that with one hidden layer, an ANN can describe any continuous function (if there are enough hidden nodes), and with two hidden layers, it can describe any function. The weights in an ANN, similar to coefficient in a regression model, are adjusted to solve the problem presented to ANN. Learning or training is used to describe process of finding values of these weights (Achal, 2013).

In general, the relationship between the output (y_t) and the inputs (y_{t-1}, $y_{t-2}\ldots\ y_{t-i}$) in ANN model has the following mathematical representation.

$$y_t = w_0 + \sum_{j}^{q} w_j g\left(w_{0,j} + \sum_{i}^{p} w_{i,j}\, y_{t-i}\right) + \varepsilon_t; i = 1, 2, 3, \ldots p \;\&\; j = 1, 2, 3, \ldots, q$$

where, y_t = output at time t

ε_t = error term at time

p = the number of input nodes

q = the number of hidden nodes and

w_j are model parameters often called connection weights

3.3.6.2 Issues in ANN modeling for forecasting

Modeling issues that affect the performance of an ANN must be considered carefully. One of the critical decisions is to determine the appropriate architecture, that is, the number of layers and the number of nodes in each layer. Other network design decisions include selection of activation functions of the hidden, the training algorithm, training and testing data sets, and performance measures. These issues are explained under following headings:

A) The network architecture

In designing a Multi-Layer Perceptron (MLP), one must determine the following variables:

- The number of input nodes
- The number of hidden layers and hidden nodes
- The number of output nodes

Hence, the designing of an ANN is more of an art than a science (Achal, 2013).

i) The number of input nodes

The selection of number of input nodes is important as too few or too many inputs nodes can affect either the learning or the prediction capability of the network. It is quite difficult to choose the number of input nodes in a time-series, as there is no fixed procedure for it. The number of input nodes are corresponds to the number of lagged observations in non-seasonal time-series and number of seasonal lags (12 for monthly data, 4 for quarterly data and 52 for weekly data) in seasonal time-series data, to make forecasts for future values.

ii) The number of hidden layers and nodes

The hidden nodes and the hidden layer helps the neural networks for detection of the feature, to capture the pattern present in the data and helps to establish complicated nonlinear mapping between input and output variables. In absence of the hidden nodes, the simple perceptron's with linearly connected output nodes are same as linear statistical forecasting models. Literature reveals the fact that ANNs can sufficiently approximate any complex nonlinear function with any desired accuracy by a single hidden layer. Sometimes, it is observed that two hidden layers may give better results for some specific problems; especially when one hidden layer network is overloaded with too many hidden nodes to give satisfactory results. In general, networks having fewer hidden nodes are preferable as they usually have better generalization ability and less over fitting problem. However, networks with too few hidden nodes may not have enough power to model and learn the data.

iii) The number of output nodes

Neural networks with multiple outputs, especially when these outputs are widely spaced, will give inferior results as compared to a network with single output. For a time-series forecasting problem, the number of output nodes are directly related to the forecasting horizon and it will be usually one node for many time-series data.

B) Activation function

The activation function is also known as the transfer function. It solely determines the relationship between input and outputs of a node and a network. The activation function is responsible for introducing a degree of nonlinearity that is valuable for most ANN applications. Roughly speaking, any differentiable function can be an activation function. In practice, only a small number of "well behaved" (bounded, monotonically increasing, and differentiable) activation functions are used. They are

i. The sigmoid or logistic function:

$$f(y) = \frac{1}{1+e^{-y}};$$

ii. The hyperbolic tangent (Tanh) function:

$$f(y) = \frac{exp(y)-exp\,(-y)}{exp(y)+exp\,(-y)};$$

iii. The linear (identity) function:

$$f(y) = y.$$

Among this, sigmoid transfer function is the most commonly used for all hidden and output nodes. However, for a forecasting problem which involves continuous values, it is reasonable to use a linear activation function for output nodes. It is important to note that ANNs with linear output nodes come with the limitation that they cannot model a time-series containing a trend. Hence, for this type of neural networks, pre-differencing may be needed to eliminate the trend effects.

C) Training sample and test sample

For building an ANN model, training and test sample are must. The training sample is used for ANN model development and test sample is adopted for evaluating the forecasting ability of the model. Sometimes a third one called the validation sample is also utilized to avoid the over fitting problem or to determine the stopping point of the training process. The selection of the training, validation and test sample may affect the performance of ANNs. The main question is how to divide the data into the training, validation and test sets? Although there is no definite answer to this problem, several factors such as the problem characteristics, the data type and the size of the available data should be considered while dividing the data set. Most authors select the training, validation and test sets based on the proportion rule of (70:20:10), (80:15:5), (85:10:5). The amount of data for the network training depends on the network structure, the training method and the complexity of the particular problem or the amount of noise in the data on hand. The ANN forecasting performance increases as the training sample size increases.

IV RESULTS AND DISCUSSION

In consonance with the objectives of the study, the necessary data collected were analyzed and interpreted. In this chapter, the results of the analysis along with discussion are presented under the following heads:

4.1 To analyze the trend in the rainfall pattern in meteorological subdivisions of Karnataka

4.2 To study the shift in rainfall distribution pattern

4.3 To evaluate a suitable non-linear time series models for forecasting rainfall

4.1 To analyze the trend in the rainfall pattern in meteorological subdivisions of Karnataka

To determine the nature of trend in month-wise, seasonal and annual rainfall data of four meteorological subdivisions of Karnataka, non-parametric tests *viz.* Mann-Kendall (M-K) and Modified Mann-Kendall (MM-K) tests were employed for the period of 60 years *i.e.,* from 1960-2019. Significance values of M-K and MM-K test statistic (Tau) indicates that there is presence of monotonic trend in the rainfall data. The positive value of Tau (τ) indicates an upward or increasing trend and a negative value gives a downward or decreasing trend in the rainfall. The Sen's slope estimator (Q_{Med}) was employed to get the magnitude of the trend. Sen's slope estimate value near to zero indicates negligible rate of change of rainfall, either larger positive or negative value indicates high increasing or decreasing rate rainfall.

Before trend analysis, significance of serial correlation coefficient was tested to check the presence of the serial correlation in all rainfall data series. This test revealed that most of the rainfall data series have a significant serial correlation. Therefore, to eliminate the effect of serial correlation on acceptance or rejection of Kendall Tau (τ), the Modified Mann-Kendall (MMK) test for variance correction is used. Rainfall trend is analyzed using '***modifiedmk***' package in 'R' software.

4.1.1 Analysis of trend in month-wise rainfall data

NIK subdivision

For month-wise rainfall data of NIK subdivision, M-K Tau value, Correction Factor (C.F), Sen's slope estimator (Q) and p-values for M-K and MM-K tests are presented in Table 4.1. For NIK subdivision, M-K test statistic (Tau) values in Table 4.1 are found to be non-significant at 5 per cent level of significance (p-values ≥ 0.05) for all the months, which indicates that there is no monotonic trend in month-wise rainfall data. This may be due to the influence of serial correlation in the month-wise rainfall data. Therefore, to overcome the effect of serial correlation, if any, MM-K test was employed.

MM-K test statistic (Tau) values for January (0.10), March (0.04), April (-0.05) and August (0.06) months were found to be significant at 5 per cent level of significance ($p \leq 0.05$), and February (0.12), May (-0.07), July (-0.13) and September (-0.09) months found to be significant at 1 per cent level of significance ($p \leq 0.01$), whereas remaining months namely June (0.03), October (-0.05), November (-0.04) and December (0.04) found to be non-significant. Significance and non-significance values of MM-K test statistic (Tau) respectively indicates that presence and absence of monotonic trend in the month-wise rainfall data for NIK subdivision. The positive sign of MM-K (Tau) value for January, February, March and August months revealed that there is a monotonic increasing trend and negative sign of MM-K (Tau) value for April, May, July and September months showed that monotonic decreasing trend in month-wise rainfall data. Remaining months June, October, November and December have no monotonic trend in month-wise rainfall data. Sen's slope values towards zero indicates negligible rate of change of rainfall for month-wise rainfall of NIK subdivision. The results obtained in the study are in accordance with Chakraborty *et al.* (2013).

SIK subdivision

For month-wise rainfall data of SIK subdivision, M-K Tau value, Correction Factor (C.F), Sen's slope estimator (Q) and p-values for M-K and MM-K tests are tabulated in Table 4.2. For SIK subdivision, M-K test statistic (Tau) values in Table 4.2 are found to be non-significant at 5 per cent level of significance (p-values ≥ 0.05) for all the months,

Table 4.1: M-K and MM-K test statistic (Tau) and Sen's slope estimate for NIK subdivision for monthly rainfall (mm) data

Period	M-K test			MM-K test				Sen's slope
	Tau	p	Trend	Tau	C.F	p	Trend	
January	0.10^{NS}	0.26	No Trend	**0.10***	0.20	0.01	**Increasing**	0.00
February	0.12^{NS}	0.18	No Trend	**0.12****	0.14	<0.01	**Increasing**	0.0008
March	0.04^{NS}	0.64	No Trend	**0.04***	0.43	0.05	**Increasing**	0.01
April	-0.05^{NS}	0.61	No Trend	**-0.05***	0.16	0.02	**Decreasing**	-0.06
May	-0.07^{NS}	0.44	No Trend	**-0.07****	0.08	0.01	**Decreasing**	-0.17
June	0.03^{NS}	0.77	No Trend	0.03^{NS}	0.22	0.05	No trend	0.05
July	-0.13^{NS}	0.15	No Trend	**-0.13****	0.06	<0.01	**Decreasing**	-0.50
August	0.06^{NS}	0.50	No Trend	**0.06***	0.08	0.01	**Increasing**	0.20
September	-0.09^{NS}	0.32	No Trend	**-0.09****	0.10	<0.01	**Decreasing**	-0.50
October	-0.05^{NS}	0.57	No Trend	-0.05^{NS}	0.17	0.17	No trend	-0.26
November	-0.04^{NS}	0.66	No Trend	-0.04^{NS}	0.22	0.34	No trend	-0.04
December	0.04^{NS}	0.64	No Trend	0.04^{NS}	0.33	0.41	No trend	0.001

Table 4.2: M-K and MM-K test statistic (Tau) and Sen's slope estimate for SIK subdivision for monthly rainfall (mm) data

Period	M-K test			MM-K test				Sen's slope
	Tau	p	Trend	Tau	C.F	p	Trend	
January	0.02^{NS}	0.15	No trend	0.02^{NS}	0.10	0.05	No trend	0.00
February	0.08^{NS}	0.39	No trend	**0.08****	0.10	0.01	**Increasing**	0.003
March	0.15^{NS}	0.08	No trend	**0.15****	0.37	0.01	**Increasing**	0.08
April	0.09^{NS}	0.32	No trend	**0.09***	0.19	0.02	**Increasing**	0.16
May	0.06^{NS}	0.47	No trend	0.06^{NS}	0.40	0.31	No trend	0.16
June	0.17^{NS}	0.06	No trend	**0.17****	0.20	<0.01	**Increasing**	0.37
July	0.03^{NS}	0.25	No trend	0.03^{NS}	0.06	0.16	No trend	0.05
August	0.15^{NS}	0.10	No trend	**0.15****	0.08	<0.01	**Increasing**	0.56
September	-0.07^{NS}	0.44	No trend	-0.07^{NS}	0.16	0.05	No trend	-0.38
October	0.02^{NS}	0.86	No trend	0.02^{NS}	0.19	0.70	No trend	0.08
November	-0.03^{NS}	0.75	No trend	-0.03^{NS}	0.09	0.29	No trend	-0.07
December	-0.04^{NS}	0.65	No trend	-0.04^{NS}	0.31	0.41	No trend	-0.03

p: probability value, NS: Non-Significant *Significant at 5% level of significance, **Significant at 1% level of significance

which indicates that there is no monotonic trend in month-wise rainfall data. This may be due to the influence of serial correlation in the month-wise rainfall data. Therefore, to overcome the effect of serial correlation, if any, MM-K test was employed.

MM-K test statistic (Tau) values for April (0.09) month is found to be significant at 5 per cent level of significance ($p \leq 0.05$), and February (0.08), March (0.15), June (0.17) and August (0.15) months were found to be significant at 1 per cent level of significance ($p \leq 0.01$). Whereas, remaining months namely January (0.02), May (0.06), July (0.03), September (-0.07), October (0.02), November (-0.03) and December (-0.04) found to be non-significant. Significance and non-significance values of MM-K test statistic (Tau) respectively indicates that presence and absence of monotonic trend in the month-wise rainfall data for SIK subdivision. The positive sign of MM-K (Tau) value for February, March, April, June and August months revealed that there is a monotonic increasing trend in month-wise rainfall data. Remaining January, May, July, September, October, November and December have no monotonic trend in month-wise rainfall data. Sen's slope values towards zero indicates negligible rate of change of rainfall for month-wise rainfall of NIK subdivision.

Malnad subdivision

For month-wise rainfall data of Malnad subdivision, M-K Tau value, Correction Factor (C.F), Sen's slope estimator (Q) and p-values for M-K and MM-K tests are presented in Table 4.3. For Malnad subdivision, M-K test statistic (Tau) values in Table 4.3 are found to be non-significant at 5 per cent level of significance (p-values ≥ 0.05) for all the months, which indicates that there is no monotonic trend in month-wise rainfall data. This may be due to the influence of serial correlation in the month-wise rainfall data. Therefore, to overcome the effect of serial correlation, if any, MM-K test was employed.

MM-K test statistic (Tau) values for June (0.07) month is found to be significant at 5 per cent level of significance ($p \leq 0.05$) and remaining months namely January (-0.04), February (0.08), March (0.05), April (-0.04), May (-0.08), July (-0.13), August (0.03), September (0.07), October (-0.03), November (-0.04) and December (-0.04) found to be non-significant. Significance and non-significance values of MM-K test statistic (Tau)

respectively indicates that presence and absence of monotonic trend in the month-wise rainfall data for Malnad subdivision. The positive sign of MM-K (Tau) value for June month revealed that there is a monotonic increasing trend in this month rainfall data. Remaining months namely, January, February, March, April, May, July, August, September, October, November and December have no monotonic trend in month-wise rainfall data. Sen's slope value (0.79) towards zero indicates negligible increasing rate of change of rainfall for June month rainfall of Malnad subdivision.

Coastal subdivision

For month-wise rainfall data of Coastal subdivision, M-K Tau value, Correction Factor (C.F), Sen's slope estimator (Q) and p-values for M-K and MM-K tests are organized in Table 4.4. For Coastal subdivision, M-K test statistic (Tau) value for February (0.27) month is found to be significant at 1 per cent level of significance and non-significant for remaining months namely January (0.15), March (0.15), April (0.12), May (-0.03), June (0.07), July (-0.15), August (-0.003), September (0.03), October (0.06), November (0.04) and December (-0.05), which indicates that there is no monotonic trend in month-wise rainfall data. This may be due to the influence of serial correlation in the month-wise rainfall data. Therefore, to overcome the effect of serial correlation, if any, MM-K test was employed.

MM-K test statistic (Tau) value March (0.15) month is found to be significant at 5 per cent level of significance ($p \leq 0.05$), and January (0.15), February (0.12), April (0.12) and July (-0.15) months were found to be significant at 1 per cent level of significance ($p \leq 0.01$). Whereas remaining months namely May (-0.03), June (0.07), August (-0.003), September (0.03), October (0.06), November (0.04) and December (-0.05) found to be non-significant. Significance and non-significance values of MM-K test statistic (Tau) respectively indicates that presence and absence of monotonic trend in the month-wise rainfall data for Coastal subdivision. The positive sign of MM-K (Tau) value for January, February, March and April months revealed that there is a monotonic increasing trend and negative sign of MM-K (Tau) value for July month reveled that monotonic decreasing trend in July month rainfall data.

Table 4.3: M-K and MM-K test statistic (Tau) and Sen's slope estimate for Malnad subdivision for monthly rainfall (mm) data

Period	M-K test			MM-K test				Sen's slope
	Tau	p	Trend	Tau	C.F	p	Trend	
January	-0.04 NS	0.97	No trend	-0.04 NS	0.07	0.10	No trend	0.00
February	0.08 NS	0.40	No trend	0.08 NS	0.47	0.22	No trend	0.00
March	0.05 NS	0.60	No trend	0.05 NS	0.42	0.41	No trend	0.02
April	-0.04 NS	0.69	No trend	-0.04 NS	0.19	0.35	No trend	-0.10
May	-0.08 NS	0.35	No trend	-0.08 NS	0.38	0.13	No trend	-0.41
June	0.07 NS	0.43	No trend	**0.07***	0.12	0.02	**Increasing**	0.79
July	-0.13 NS	0.13	No trend	-0.13 NS	0.76	0.09	No trend	-2.16
August	0.03 NS	0.74	No trend	0.03 NS	0.34	0.57	No trend	0.46
September	0.07 NS	0.43	No trend	0.07 NS	0.37	0.19	No trend	0.36
October	-0.03 NS	0.75	No trend	-0.03 NS	0.23	0.51	No trend	-0.13
November	-0.04 NS	0.62	No trend	-0.04 NS	0.09	0.09	No trend	-0.11
December	-0.04 NS	0.65	No trend	-0.04 NS	0.17	0.26	No trend	-0.01

Table 4.4: M-K and MM-K test statistic (Tau) and Sen's slope estimate for Coastal subdivision for monthly rainfall (mm) data

Period	M-K test			MM-K test				Sen's slope
	Tau	p	Trend	Tau	C.F	p	Trend	
January	0.15 NS	0.09	No trend	**0.15****	0.08	<0.01	**Increasing**	0.00
February	**0.27****	0.01	**Increasing**	**0.27****	0.46	<0.01	**Increasing**	0.0003
March	0.15 NS	0.09	No trend	**0.15***	0.60	0.03	**Increasing**	0.02
April	0.12 NS	0.17	No trend	**0.12****	0.07	<0.01	**Increasing**	0.02
May	-0.03 NS	0.75	No trend	-0.03 NS	0.21	0.48	No trend	-0.29
June	0.07 NS	0.43	No trend	0.07 NS	0.45	0.24	No trend	1.10
July	-0.15 NS	0.08	No trend	**-0.15****	0.14	0.00	**Decreasing**	-3.43
August	-0.003 NS	0.97	No trend	-0.00 NS	0.21	0.94	No trend	-0.03
September	0.03 NS	0.74	No trend	0.03 NS	0.34	0.56	No trend	0.36
October	0.06 NS	0.52	No trend	0.06 NS	0.28	0.23	No trend	0.42
November	0.04 NS	0.70	No trend	0.04 NS	0.14	0.30	No trend	0.14
December	-0.05 NS	0.60	No trend	-0.05 NS	0.65	0.52	No trend	-0.01

p: probability value, NS: Non-Significant *Significant at 5% level of significance, **Significant at 1% level of significance

Remaining months namely May, June, August, September, October, November and December have no monotonic trend in month-wise rainfall data. Sen's slope values towards zero indicates negligible rate of change of rainfall for month-wise rainfall of Coastal subdivision. Using MM-K test, Pal and Al-tabbaa (2011) reported similar results.

4.1.2 Analysis of trend in seasonal rainfall data

NIK subdivision

For seasonal rainfall data of NIK subdivision, M-K Tau value, Correction Factor (C.F), Sen's slope estimator (Q) and p-values for M-K and MM-K tests are presented in Table 4.5. For NIK subdivision, M-K test statistic (Tau) values for all seasons found to be non-significant at 5 per cent level of significance (p-values $\geq 0.05$for all the seasons, which indicates that there is no monotonic trend in seasonal rainfall data. This may be due to the influence of serial correlation in the seasonal rainfall data. Therefore, to overcome the effect of serial correlation, if any, MM-K test was employed.

MM-K test statistic (Tau) values for winter (0.05) season found to be significant at 5 per cent level of significance ($p \leq 0.05$), and monsoon (-0.09) and post-monsoon (-0.12) seasons found to be significant at 1 per cent level of significance ($p \leq 0.01$), whereas for pre-monsoon (-0.06) season found to be non-significant. Significance and non-significance values of MM-K test statistic (Tau) respectively indicates that presence and absence of monotonic trend in the seasonal rainfall data for NIK subdivision. The positive sign of MM-K (Tau) value for winter season revealed that there is a monotonic increasing trend and negative sign of MM-K (Tau) value for monsoon and post-monsoon revealed that monotonic decreasing trend in seasonal rainfall data. In addition, pre-monsoon season has no monotonic trend. Sen's slope values towards zero indicates negligible rate of change of rainfall for seasonal rainfall of NIK subdivision.

SIK subdivision

For seasonal rainfall data of SIK subdivision, M-K Tau value, Correction Factor (C.F), Sen's slope estimator (Q) and p-values for M-K and MM-K tests are tabulated in Table 4.6. For SIK subdivision, M-K test statistic (Tau) values for all seasons found to be

non-significant at 5 per cent level of significance (p-values ≥ 0.05) for all the seasons, which indicates that there is no monotonic trend in seasonal rainfall data. This may be due to the influence of serial correlation in the seasonal rainfall data. Therefore, to overcome the effect of serial correlation, if any, MM-K test was employed.

MM-K test statistic (Tau) values for winter (0.05) season found to be significant at 5 per cent level of significance (p ≤ 0.05) and monsoon (-0.09) and post-monsoon (-0.12) seasons found to be significant at 1 per cent level of significance (p ≤ 0.01), whereas for pre-monsoon (-0.06) season found to be non-significant. Significance and non-significance values of MM-K test statistic (Tau) respectively indicates that presence and absence of monotonic trend in the seasonal rainfall data for SIK subdivision. The positive sign of MM-K (Tau) value for winter season is revealed that there is a monotonic increasing trend and negative sign for monsoon and post-monsoon revealed that there is a monotonic decreasing trend in seasonal rainfall data. In addition, pre-monsoon season has no monotonic trend. Narayanan *et al*. (2016) obtain similar results for seasonal and annual rainfall data using MM-K test. Sen's slope values towards zero indicates negligible rate of change of rainfall for seasonal rainfall of SIK subdivision.

Malnad subdivision

For seasonal rainfall data of Malnad subdivision, M-K Tau value, Correction Factor (C.F), Sen's slope estimator (Q) and p-values for M-K and MM-K tests are formulated in Table 4.7. For Malnad subdivision, M-K and MM-K test statistic (Tau) values for all seasons found to be non-significant at 5 per cent level of significance (p-values ≥ 0.05) for all the seasons, which indicates that there is no monotonic trend in seasonal rainfall data. Test statistic (Tau) values for winter (0.07), pre-monsoon (-0.01), monsoon (0.03) and post-monsoon (-0.03) are non-significant, which indicates no trend in seasonal rainfall data of Malnad region.

Coastal subdivision

For seasonal rainfall data of Coastal subdivision, M-K Tau value, Correction Factor (C.F), Sen's slope estimator (Q) and p-values for M-K and MM-K tests are depicted in Table 4.8.

Table 4.5: M-K and MM-K test statistic (Tau) and Sen's slope estimate for NIK subdivision for seasonal rainfall (mm) data

Period	M-K test			MM-K test				Sen's slope
	Tau	p	Trend	Tau	C.F	p	Trend	
Winter	0.05 ^{NS}	0.57	No trend	**0.05***	0.07	0.03	**Increasing**	0.03
Pre-monsoon	-0.06 ^{NS}	0.53	No trend	-0.06 ^{NS}	0.33	0.28	No trend	-0.19
Monsoon	-0.09 ^{NS}	0.30	No trend	**-0.09****	0.11	<0.01	**Decreasing**	-0.64
Post-monsoon	-0.12 ^{NS}	0.19	No trend	**-0.12****	0.21	<0.01	**Decreasing**	-0.69

Table 4.6: M-K and MM-K test statistic (Tau) and Sen's slope estimate for SIK subdivision for seasonal rainfall (mm) data

Period	M-K test			MM-K test				Sen's slope
	Tau	p	Trend	Tau	C.F	p	Trend	
Winter	0.07 ^{NS}	0.47	No trend	**0.07****	0.07	0.01	**Increasing**	0.02
Pre-monsoon	0.15 ^{NS}	0.10	No trend	**0.15****	0.32	<0.01	**Increasing**	0.58
Monsoon	0.11 ^{NS}	0.23	No trend	**0.11****	0.13	0.001	**Increasing**	0.92
Post-monsoon	-0.02 ^{NS}	0.84	No trend	-0.02 ^{NS}	0.12	0.57	No trend	-0.12

Table 4.7: M-K and MM-K test statistic (Tau) and Sen's slope estimate for Malnad subdivision for seasonal rainfall (mm) data

Period	M-K test			MM-K test				Sen's slope
	Tau	p	Trend	Tau	C.F	p	Trend	
Winter	0.07 ^{NS}	0.46	No trend	0.07 ^{NS}	0.27	0.15	No trend	0.01
Pre-monsoon	-0.01 ^{NS}	0.89	No trend	-0.01 ^{NS}	0.42	0.84	No trend	-0.11
Monsoon	0.03 ^{NS}	0.75	No trend	0.03 ^{NS}	0.64	0.70	No trend	0.56
Post-monsoon	-0.03 ^{NS}	0.74	No trend	-0.03 ^{NS}	0.17	0.41	No trend	-0.25

Table 4.8: M-K and MM-K test statistic (Tau) and Sen's slope estimate for Coastal subdivision for seasonal rainfall (mm) data

Period	M-K test			THE MM-K test				Sen's slope
	Tau	p	Trend	Tau	C.F	p	Trend	
Winter	0.14 ^{NS}	0.10	No trend	**0.14****	0.14	<0.01	**Increasing**	0.06
Pre-monsoon	0.03 ^{NS}	0.75	No trend	0.03 ^{NS}	0.26	0.54	No trend	0.18
Monsoon	-0.03 ^{NS}	0.76	No trend	-0.03 ^{NS}	0.19	0.49	No trend	-0.74
Post-monsoon	0.02 ^{NS}	0.82	No trend	0.02 ^{NS}	0.20	0.67	No trend	0.16

P: probability value, NS: Non-Significant, *Significant at 5% level of significance, **Significant at 1% level of significance

For Coastal subdivision, M-K test statistic (Tau) values for all seasons found to be non-significant at 5 per cent level of significance (p-values ≥ 0.05) for all the seasons, which indicates that there is no monotonic trend in seasonal rainfall data. This may be due to the influence of serial correlation in the seasonal rainfall data. Therefore, to overcome the effect of serial correlation, if any, MM-K test was employed.

MM-K test statistic (Tau) values for winter (0.14) is found to be significant at 1 per cent level of significance ($p \leq 0.01$), whereas for pre-monsoon (0.03), monsoon (-0.03) and post-monsoon (0.02) seasons found to be non-significant. Significance and non-significance values of MM-K test statistic (Tau) respectively indicates that presence and absence of monotonic trend in the seasonal rainfall data for SIK subdivision. The positive sign of MM-K (Tau) value for winter season revealed that there is a monotonic increasing trend. In addition, pre-monsoon, monsoon and post-monsoon seasons have no monotonic trend. The present findings were in corroboration with the studies conducted by Sridhara *et al.* (2020) for seasonal rainfall data. Sen's slope values towards zero indicates negligible rate of change of rainfall for seasonal rainfall of Coastal subdivision.

4.1.3 Analysis of trend in annual rainfall data

For annual rainfall data of all four subdivisions, M-K Tau value, Correction Factor (C.F), Sen's slope estimator (Q) and p-values for M-K and MM-K tests are illustrated in Table 4.9. The M-K test statistic (Tau) values for all subdivisions were found to be non-significant at 5 per cent level of significance (p-values ≥ 0.05) for annual rainfall data, which indicates that there is no monotonic trend in seasonal rainfall data. This may be due to the influence of serial correlation in the annual rainfall data. Therefore, to overcome the effect of serial correlation, if any, MM-K test was employed.

MM-K test statistic (Tau) values for Coastal (0.12) subdivision found to be significant at 5 per cent level of significance ($p \leq 0.05$), and for NIK (-0.12) and SIK (0.11) subdivisions found to be significant at 1 per cent level of significance ($p \leq 0.01$), whereas for Malnad (0.08) subdivision found to be non-significant. Significance and non-significance values of MM-K test statistic (Tau) respectively indicates that presence and absence of monotonic trend in the annual rainfall data for each subdivision. The positive sign of MM-K (Tau) value for SIK and Coastal subdivision revealed that there is a

monotonic increasing trend and negative sign of MM-K (Tau) value for SIK subdivision indicates there is a monotonic decreasing trend. In addition, Malnad subdivision has no monotonic trend. Sharma and Saha (2017) and Chakraborty *et al.* (2013) reported decreasing trend in annual rainfall data of their study region. Larger negative Sen's slope values for NIK (-1.40) subdivision and positive Sen's slope for Malnad (2.30) and Coastal (5.10) subdivisions respectively indicates high decreasing and increasing rate of change of rainfall for annual rainfall data.

Table 4.9: M-K and MM-K test statistic (Tau) and Sen's slope estimate for all four subdivisions for annual rainfall (mm) data

Period	M-K test			MM-K test				Sen's slope
	Tau	p	Trend	Tau	C.F	p	Trend	
NIK	-0.12 ^{NS}	0.16	No trend	**-0.12****	0.10	<0.01	**Decreasing**	-1.40
SIK	0.11 ^{NS}	0.21	No trend	**0.11****	0.09	<0.01	**Increasing**	1.47
Malnad	0.08 ^{NS}	0.40	No trend	0.08 ^{NS}	0.42	0.20	No trend	2.30
Coastal	0.12 ^{NS}	0.18	No trend	**0.12***	0.40	0.03	**Increasing**	5.10

*Significant at 5% level of significance, **Significant at 1% level of significance, NS: Non-Significant

4.2 To study the shift in rainfall distribution pattern

To understand the shifting patterns (shifting point) in month-wise, seasonal and annual rainfall data of all the four meteorological subdivisions of Karnataka, the Likelihood-Ratio was computed by using '***changepoint***' package of R software at 5 per cent level of significance. Before applying Likelihood-Ratio test, normality of all data sets were examined using Shapiro-Wilk test, and it was observed that majority of data sets were following normal distribution. Therefore, Normal distribution was considered while applying Likelihood-Ratio test.

For each subdivisions, shifting point (year) in month-wise, seasonal and annual rainfall data, average rainfall (mm) before, and after shifting point (year) were obtained. Standard normal rainfall (mm) of each subdivision over different periods (month, season and annual) is compared with rainfall before and after shifting point.

4.2.1 Shifting pattern analysis for month-wise rainfall data

NIK subdivision

For each month-wise rainfall data for the period of 60 years from 1960-2019 of NIK subdivision, shifting point (year), average month-wise rainfall (mm) before and after shifting point (year) were computed are presented in Table 4.10 along with nature of shifting and normal rainfall (mm) of each month. The line graph of month-wise rainfall distribution along with rainfall shifting year, and average month-wise rainfall before and after shifting year is shown in Fig. 4.1 (a)-(b).

Shifting point (year) in month-wise rainfall was found in the month of February (shifting year- 2007), March (2013), April (1963), June (1972), August (1968), October (1975), November (1987) and December (1989), and there were no shift in rainfall was observed in the month of January, May, July and September. Results in Table 4.10 revealed that the average rainfall was increased in the month of February (1.50), March (8.03), June (16.89) and August (27.71) whereas, the average rainfall was decreased in the month of April (-22.24), October (-27.61), November (-12.27) and December (-2.20) as compared with before and after shifting point in rainfall. Further, it can also observed from the Table 4.10, the average rainfall in the month of February, March, June and August were below the normal rainfall before shifting year, but it was increased after shifting year, which was above the normal rainfall. Whereas in the month of April, October, November, and December the average rainfall was above normal rainfall before shifting year, but it was decreased after shifting year, which was below the normal rainfall except for the month of April.

SIK subdivision

For each month-wise rainfall data for the period of 60 years from 1960-2019 of SIK subdivision, shifting point (year), average month-wise rainfall (mm) before and after shifting point (year) were computed are illustrated in Table 4.11 along with nature of shifting and normal rainfall (mm) of each month. The line graph of month-wise rainfall distribution along with rainfall shifting year, and average month-wise rainfall before and after shifting year is shown in Fig. 4.2 (a)-(b).

Table 4.10: Shifting pattern of monthly rainfall (mm) data of NIK subdivision

Period	Shifting year	Average rainfall (mm)		Change in rainfall (mm)	Nature of shifting	Normal rainfall (mm)
		Before	After			
January	-	1.25	1.25	-	No shift	2.10
February	2007	0.83	2.33	1.50	Increasing	1.80
March	2013	4.71	12.74	8.03	Increasing	5.00
April	1963	47.88	25.64	-22.24	Decreasing	25.00
May	-	53.51	53.51	-	No shift	55.00
June	1972	86.51	103.40	16.89	Increasing	101.00
July	-	121.92	121.92	-	No shift	131.00
August	1968	90.88	118.59	27.71	Increasing	118.00
September	-	140.91	140.91	-	No shift	145.00
October	1975	124.76	97.15	-27.61	Decreasing	111.00
November	1987	31.15	18.88	-12.27	Decreasing	28.00
December	1989	3.33	1.13	-2.20	Decreasing	7.00

Table 4.11: Shifting pattern of monthly rainfall (mm) data of SIK subdivision

Period	Shifting year	Average rainfall (mm)		Change in rainfall (mm)	Nature of shifting	Normal rainfall (mm)
		Before	After			
January	2014	0.87	3.03	2.16	Increasing	1.80
February	1998	2.28	4.23	1.95	Increasing	3.50
March	2013	7.06	16.26	9.20	Increasing	8.00
April	1993	38.03	50.15	12.12	Increasing	41.00
May	2003	86.91	112.81	25.90	Increasing	96.00
June	1976	51.49	72.54	21.05	Increasing	64.00
July	2016	78.47	50.54	-27.93	Decreasing	79.00
August	1994	79.98	109.59	29.61	Increasing	81.00
September	1989	150.99	128.90	-22.09	Decreasing	135.00
October	1990	128.54	156.51	27.97	Increasing	146.00
November	2015	56.97	18.25	-38.72	Decreasing	50.00
December	1972	21.49	11.38	-10.11	Decreasing	14.00

Shifting point (year) in month-wise rainfall was found in the all the months for SIK subdivision *i.e.*, January (2014), February (1998), March (2013), April (1993), May (2003), June (1976), July (2016), August (1994), September (1989), October (1990), November (2015) and December (1972). Results in Table 4.11 revealed that the average rainfall was increased in the month of January (2.16), February (1.95), March (9.20), April (12.12), May (25.90), June (21.05), August (29.61) and October (27.97) whereas, the average rainfall was decreased in the month of July (-27.93), September (-22.09), November (-38.72) and December (-10.11) as compared with before and after shifting point in rainfall. Further, it can also observed from the Table 4.11, the average rainfall in the month of January, February, March, April, May, June, July, August and October were below the normal rainfall before shifting year, but it was increased after shifting year, which was above the normal rainfall except for the month of July. Whereas in the month of September, November, and December the average rainfall was above normal rainfall before shifting year, but it was decreased after shifting year, which was below the normal rainfall.

Malnad subdivision

For each month-wise rainfall data for the period of 60 years from 1960-2019 of Malnad subdivision, shifting point (year), average month-wise rainfall (mm) before and after shifting point (year) were computed are tabulated in Table 4.12 along with nature of shifting and normal rainfall (mm) of each month. The line graph of month-wise rainfall distribution along with rainfall shifting year, and average month-wise rainfall before and after shifting year is shown in Fig. 4.3 (a)-(b).

Shifting point (year) in month-wise rainfall was found in the month of February (1963), March (2012), May (1962), June (2014), July (1962), August (2016), September (2005) and December (1966)and there were no shift in rainfall was observed in the month of January, April, October and November. Results in Table 4.12 revealed that the average rainfall was increased in the month of March (8.84), June (181.19), August (323.02) and September (61.35) whereas, the average rainfall was decreased in the month of February (-2.20), May (-118.04), July (-619.92) and December (-10.44) as compared with before and after shifting point in rainfall. Further, it can also observed from the Table 4.12, the

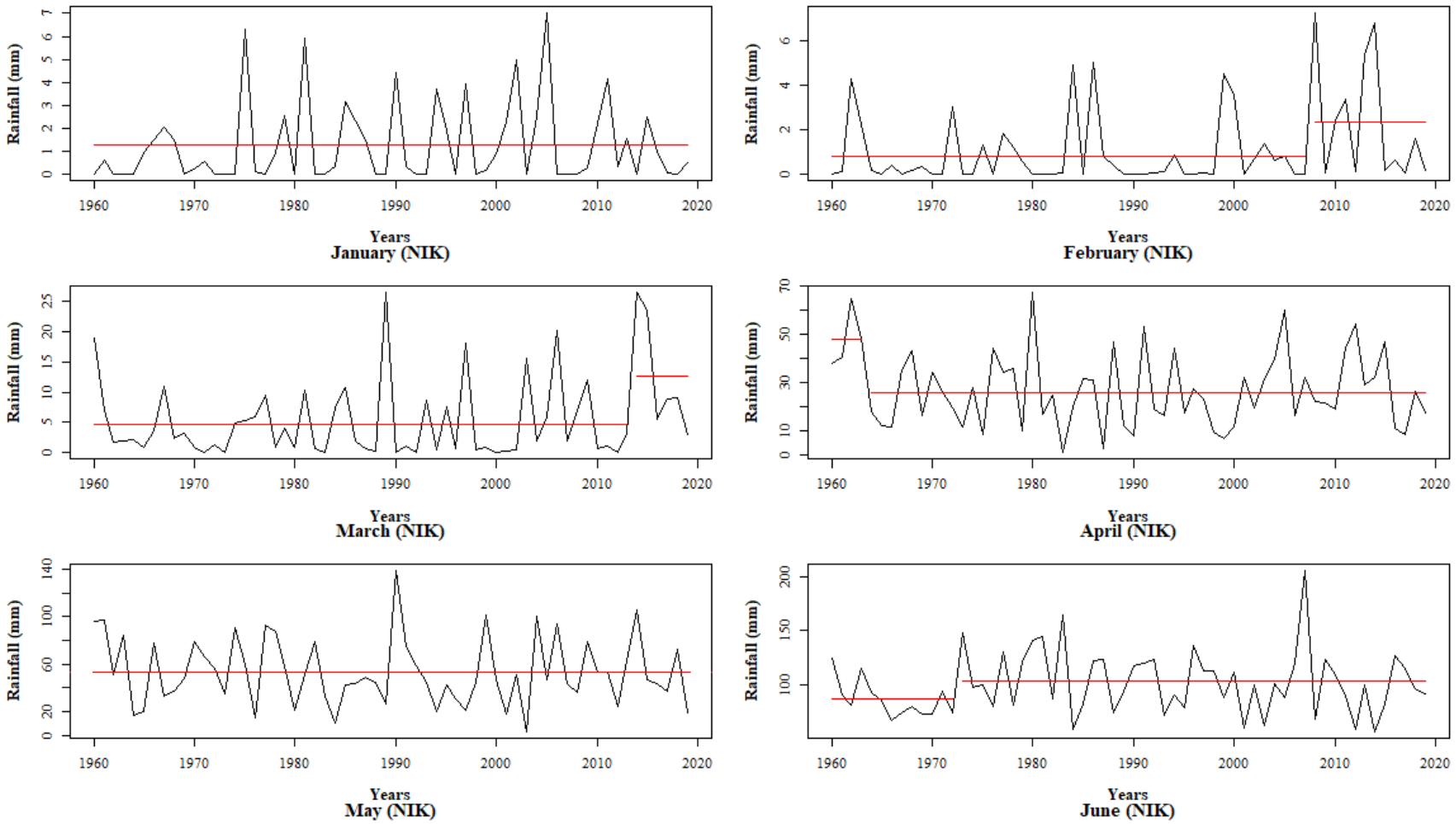

Fig. 4.1 (a): Line graphs showing rainfall distribution along with shifting pattern for January to June months for NIK subdivision

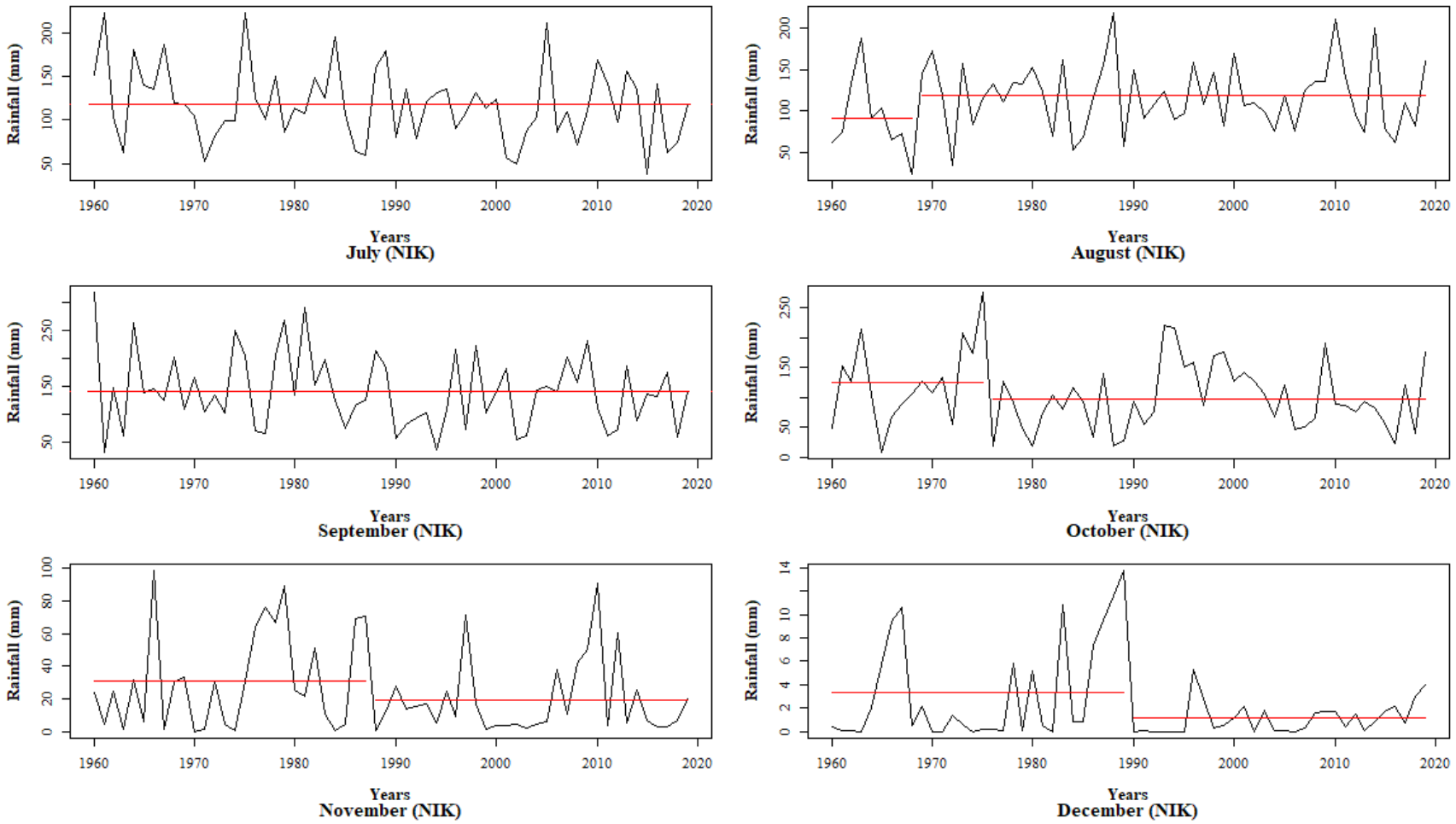

Fig. 4.1 (b): Line graphs showing rainfall distribution along with shifting pattern for July to December months for NIK subdivision

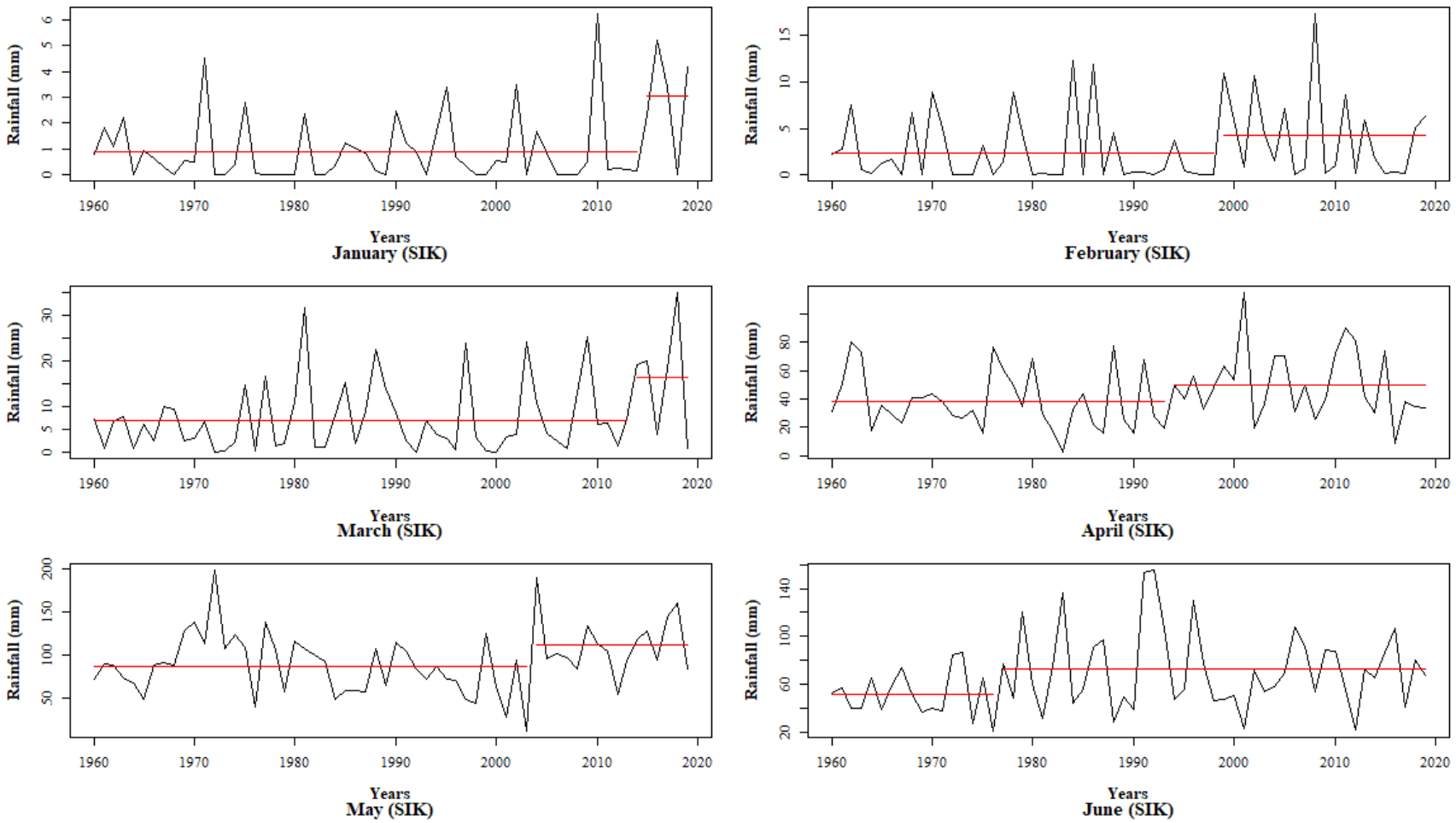

Fig. 4.2 (a): Line graphs showing rainfall distribution along with shifting pattern for January to June months for SIK subdivision

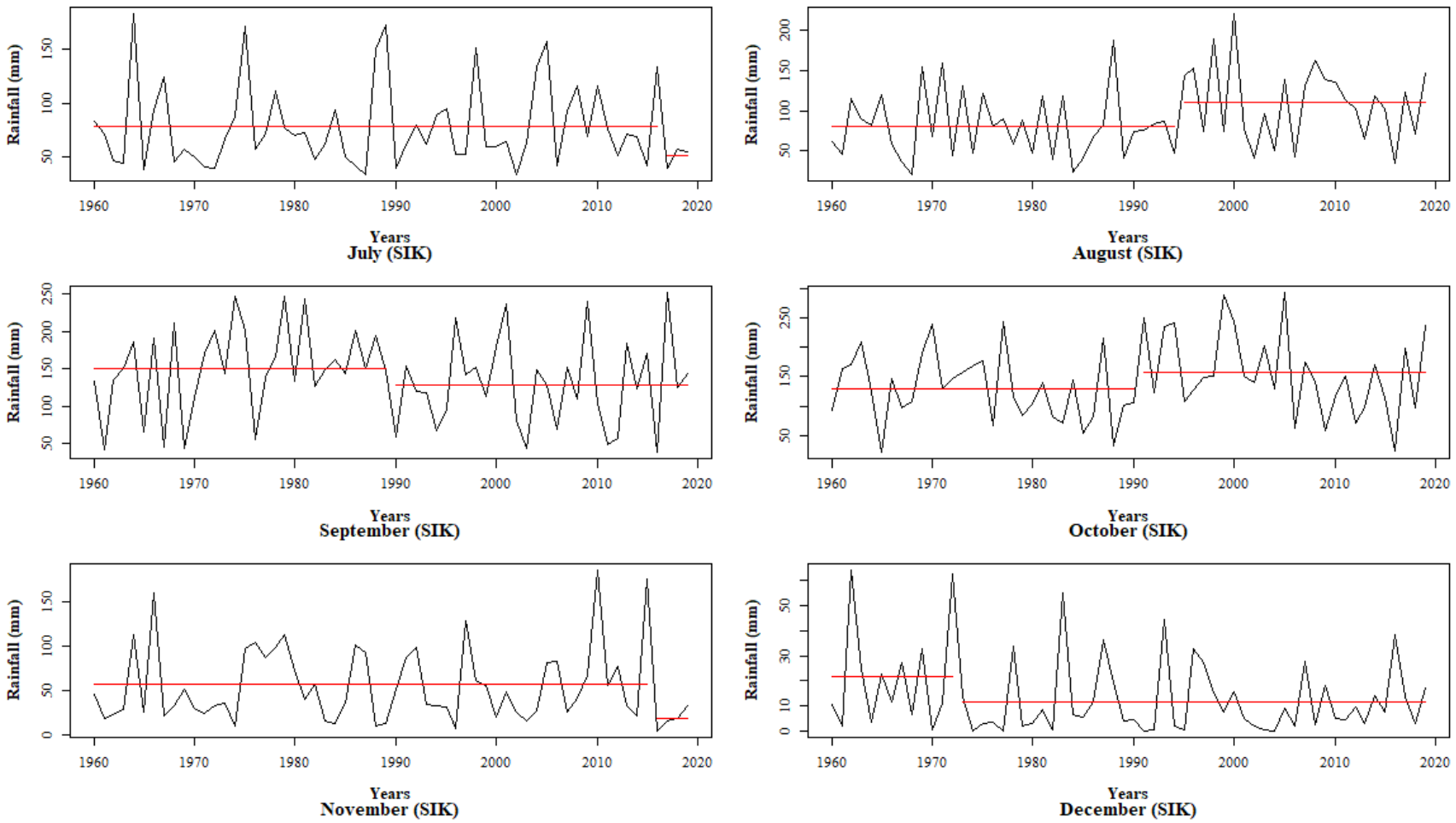

Fig. 4.2 (b): Line graphs showing rainfall distribution along with shifting pattern for July to December months for SIK subdivision

average rainfall in the month of March, June, July, August and September were below the normal rainfall before shifting year, but it was increased after shifting year, which was above the normal rainfall. Similar results are obtained by Guhathakurta *et al.* (2015). Whereas in the month of February, May, July and December the average rainfall was above normal rainfall before shifting year, but it was decreased after shifting year, which was below the normal rainfall.

Coastal subdivision

For each month-wise rainfall data for the period of 60 years from 1960-2019 of Coastal subdivision, shifting point (year), average month-wise rainfall (mm) before and after shifting point (year) were computed are displayed in Table 4.13 along with nature of shifting and normal rainfall (mm) of each month. The line graph of month-wise rainfall distribution along with rainfall shifting year, and average month-wise rainfall before and after shifting year is shown in Fig. 4.4 (a)-(b).

Shifting point (year) in month-wise rainfall was found in the month of March (2016), April (2000), May (1962), June (1967), July (1999), September (2004), November (1965) and December (2008) and there were no shift in rainfall was observed in the month of January, February, August and October. Results in Table 4.13 revealed that the average rainfall was increased in the month of March (11.21), April (14.36), June (195.41), September (113.67), November (31.12) and December (4.18) whereas, the average rainfall was decreased in the month of May (-235.60) and July (-181.94) as compared with before and after shifting point in rainfall. Further, it can also observed from the Table 4.13, the average rainfall in the month of March, April, June, July, September, November and December were below the normal rainfall before shifting year, but it was increased after shifting year, which was above the normal rainfall except for the month of June, July and December. Whereas in the month of May the average rainfall was above normal rainfall before shifting year, but it was decreased after shifting year, which was below the normal rainfall.

Table 4.12: Shifting pattern of monthly rainfall (mm) data of Malnad subdivision

Period	Shifting year	Average rainfall (mm)		Change in rainfall (mm)	Nature of shifting	Normal rainfall (mm)
		Before	After			
January	-	1.89	1.89	-	No shift	2.00
February	1963	3.82	1.62	-2.20	Decreasing	2.40
March	2012	7.49	16.33	8.84	Increasing	11.00
April	-	54.37	54.37	-	No shift	56.00
May	1962	212.58	94.54	-118.04	Decreasing	112.00
June	2014	289.03	470.22	181.19	Increasing	358.00
July	1962	1093.25	473.33	-619.92	Decreasing	598.00
August	2016	351.98	675.00	323.02	Increasing	382.00
September	2005	152.98	214.33	61.35	Increasing	165.00
October	-	151.85	151.85	-	No shift	161.00
November	-	54.33	54.33	-	No shift	55.00
December	1966	20.32	9.88	-10.44	Decreasing	13.00

Table 4.13: Shifting pattern of monthly rainfall (mm) data of Coastal subdivision

Period	Shifting year	Average rainfall (mm)		Change in rainfall (mm)	Nature of shifting	Normal rainfall (mm)
		Before	After			
January	-	2.39	2.39	-	No shift	1.00
February	-	1.10	1.10	-	No shift	0.40
March	2016	2.99	14.20	11.21	Increasing	5.00
April	2000	24.91	39.27	14.36	Increasing	31.00
May	1962	352.33	116.73	-235.60	Decreasing	135.00
June	1967	616.51	811.92	195.41	Increasing	813.00
July	1999	1070.83	888.89	-181.94	Decreasing	1156.00
August	-	720.38	720.38	-	No shift	760.00
September	2004	268.08	381.75	113.67	Increasing	290.00
October	-	186.94	186.94	-	No shift	187.00
November	1965	32.04	63.16	31.12	Increasing	61.00
December	2008	5.43	9.61	4.18	Increasing	13.00

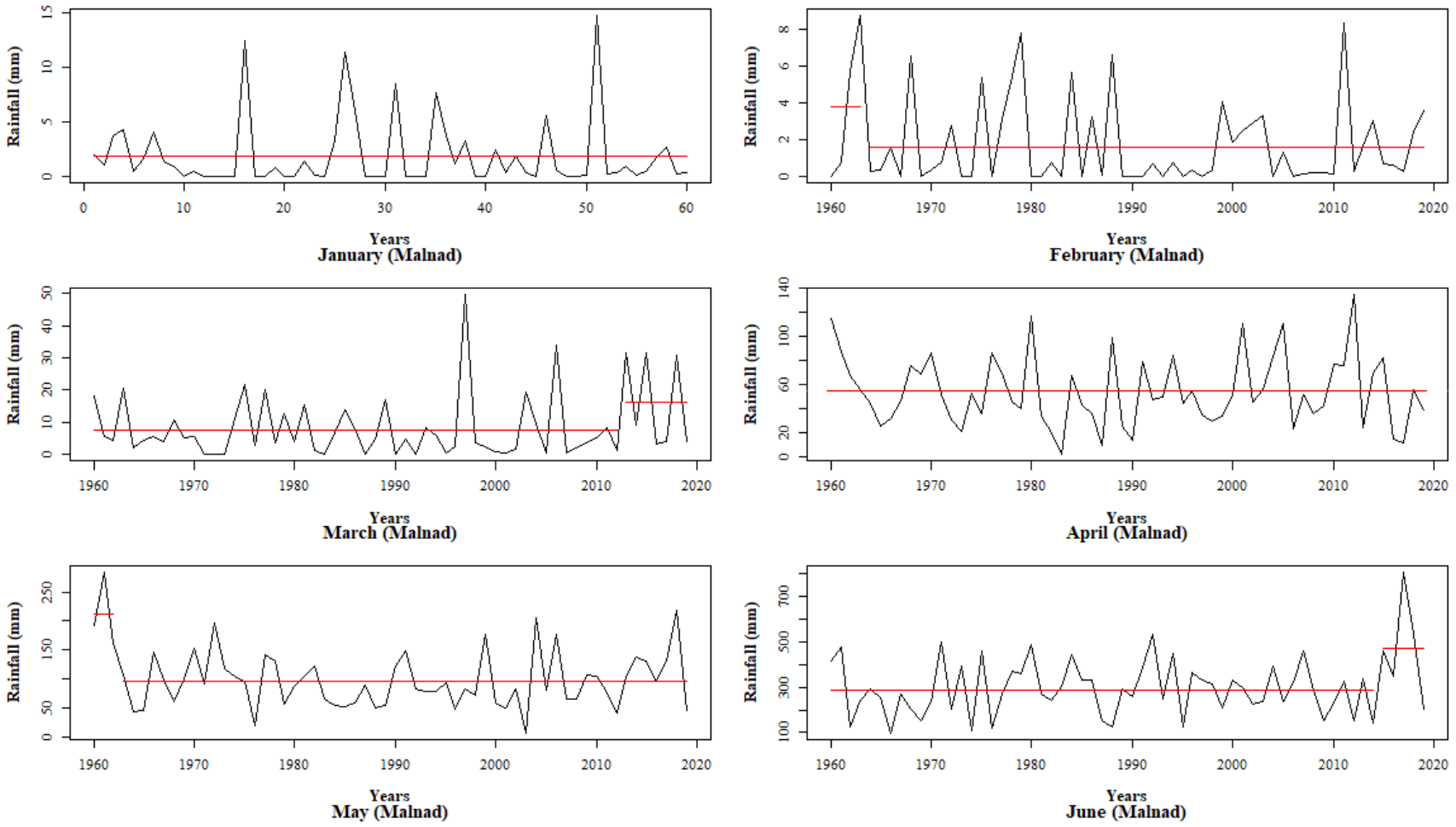

Fig. 4.3 (a): Line graphs showing rainfall distribution along with shifting pattern for January to June months for Malnad subdivision

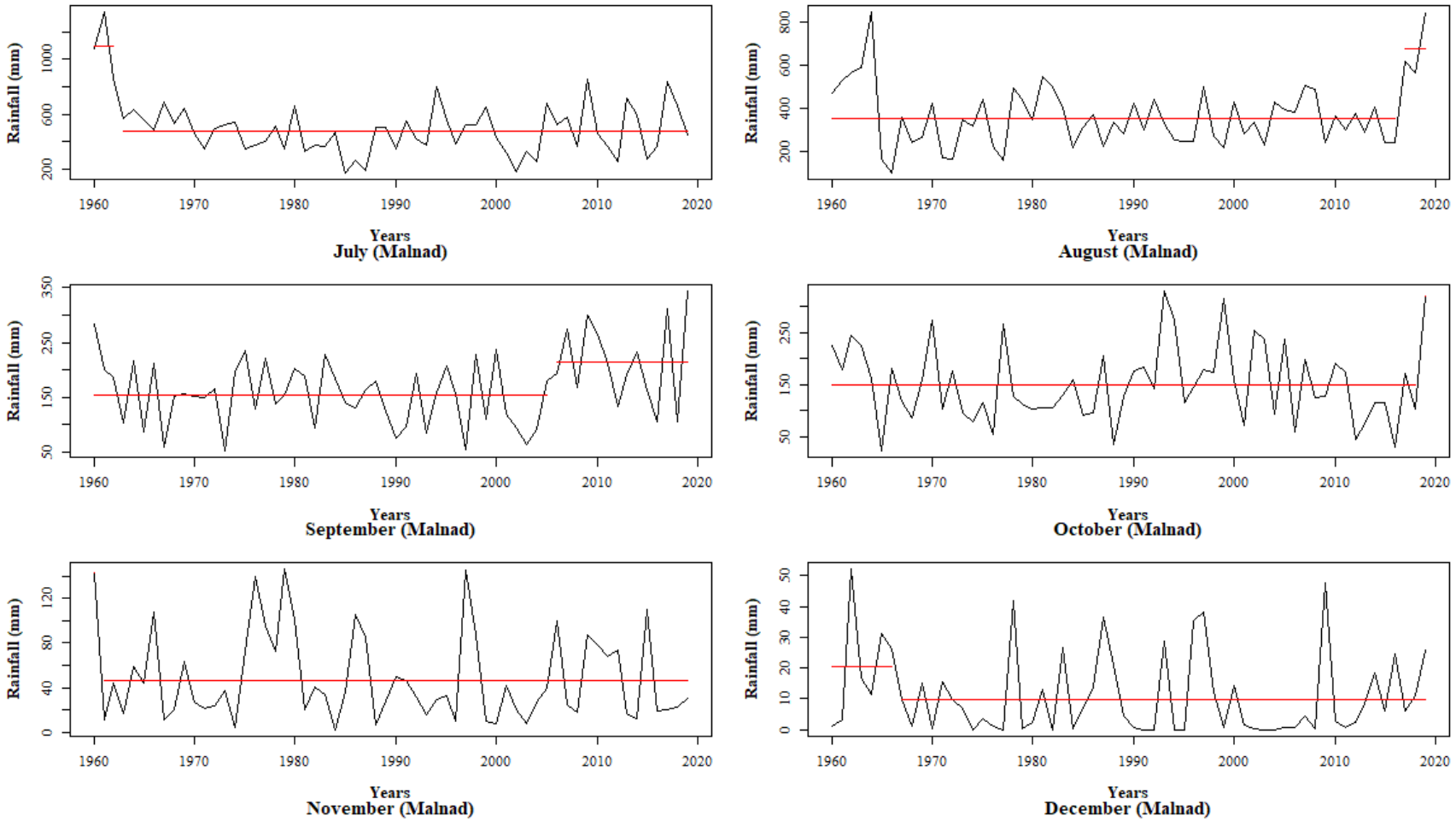

Fig. 4.3 (b): Line graphs showing rainfall distribution along with shifting pattern for July to December months for Malnad subdivision

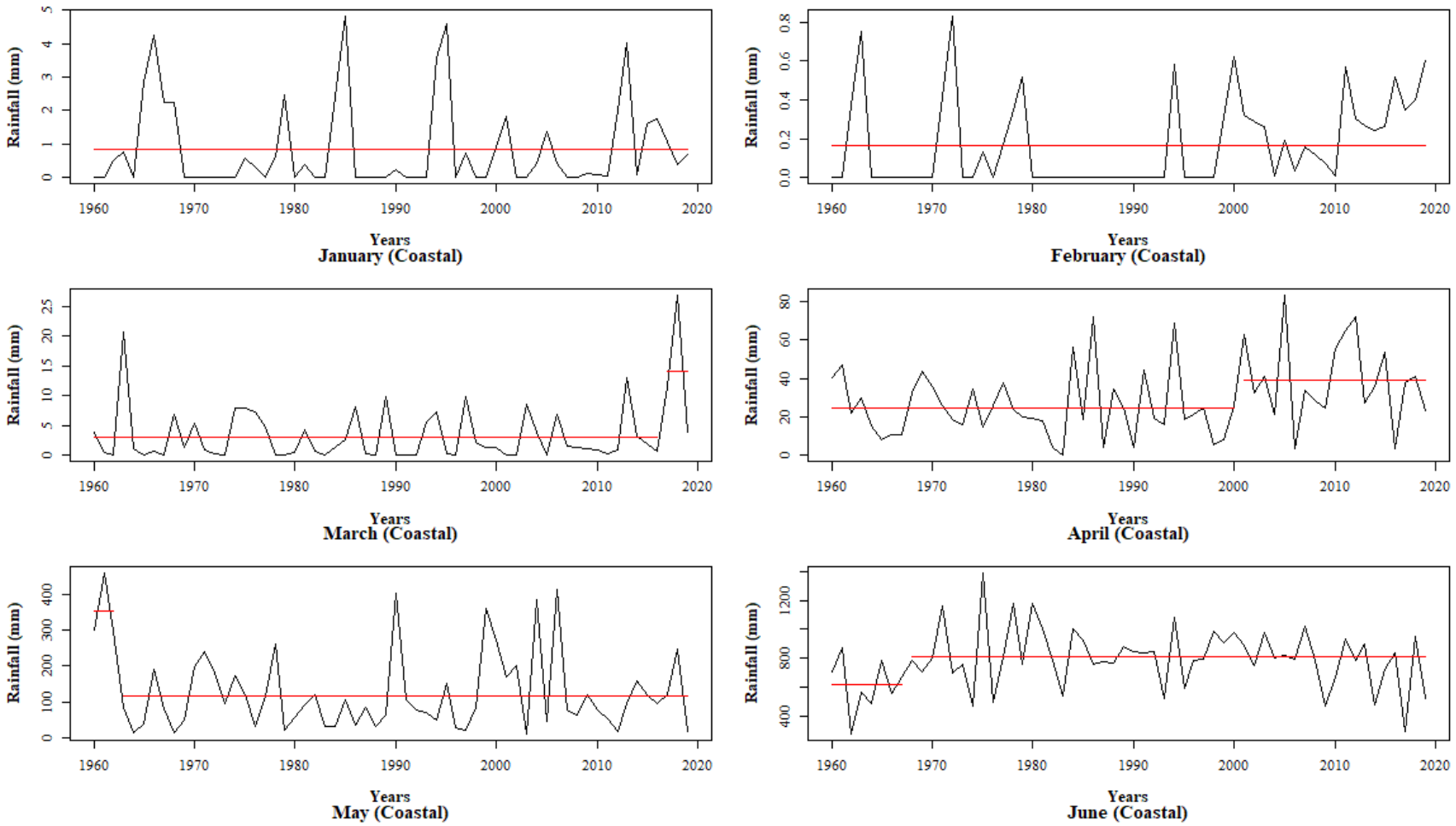

Fig. 4.4 (a): Line graphs showing rainfall distribution along with shifting pattern for January to June months for Coastal subdivision

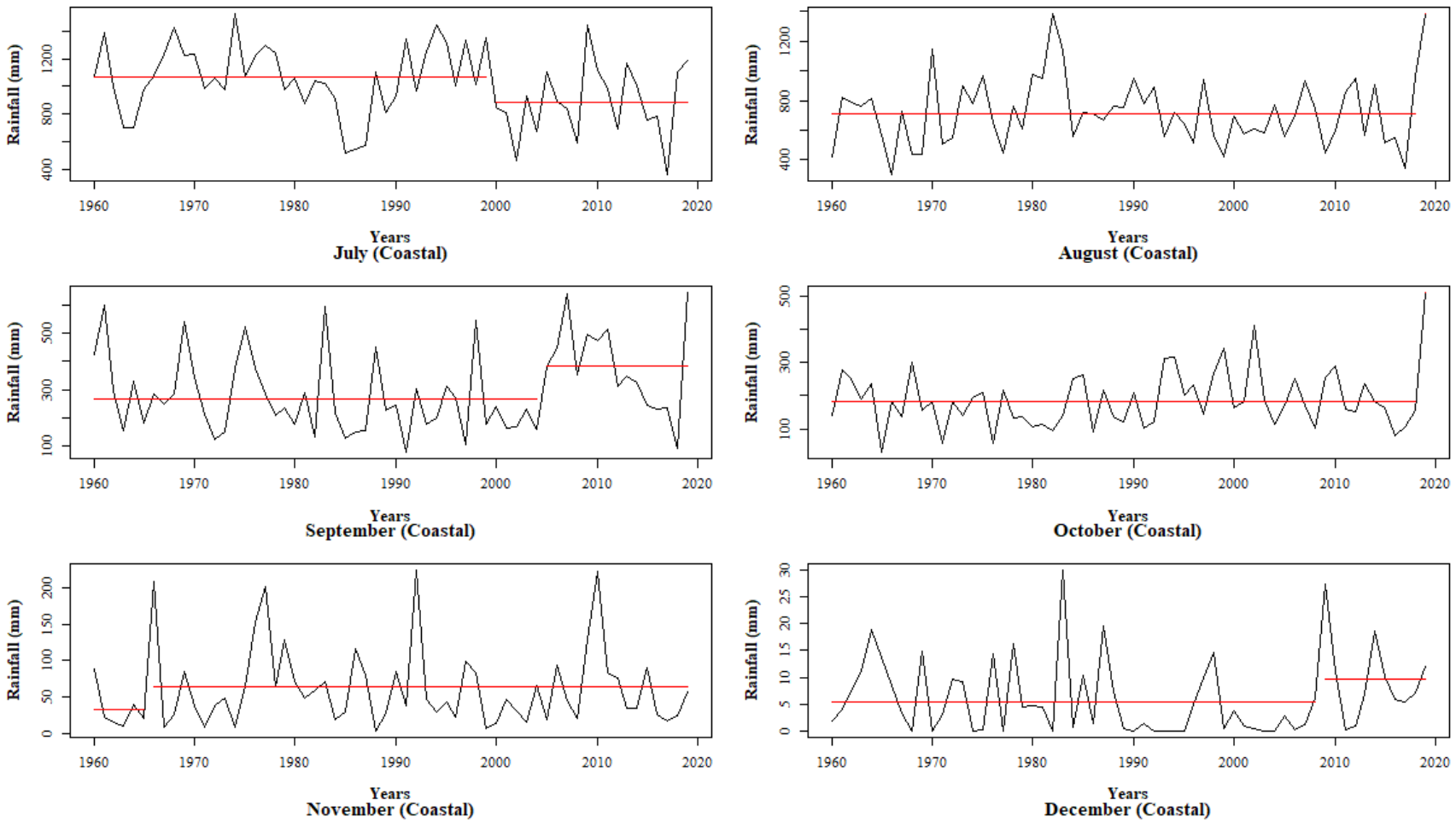

Fig. 4.4 (b): Line graphs showing rainfall distribution along with shifting pattern for July to December months for Coastal subdivision

4.2.2 Shifting pattern analysis for seasonal rainfall data

NIK subdivision

For each seasonal rainfall data for the period of 60 years from 1960-2019 of NIK subdivision, shifting point (year), average seasonal rainfall (mm) before and after shifting point (year) were computed are presented in Table 4.14 along with nature of shifting and normal rainfall (mm) of each season. The line graph of seasonal rainfall distribution along with rainfall shifting year, and average seasonal rainfall before and after shifting year is shown in Fig. 4.5.

Shifting point (year) in seasonal rainfall was found in all the seasons *viz.* winter (2015), pre-monsoon (1963), monsoon (1960) and post-monsoon (2002). Sharma and Saha (2017) obtained similar results for seasonal and annual rainfall at Damodar river basin, India. Results in Table 4.14 revealed that the average rainfall was decreased for winter (-3.03), pre-monsoon (-61.60), monsoon (-187.40) and post-monsoon (-34.14) seasons as compared with before and after shifting point in rainfall. Further, it can also observed from the Table 4.14, the average seasonal rainfall in winter, pre-monsoon, monsoon and post-monsoon was above normal rainfall before shifting year, but it was decreased after shifting year, which was below the normal rainfall.

SIK subdivision

For each seasonal rainfall data for the period of 60 years from 1960-2019 of SIK subdivision, shifting point (year), average seasonal rainfall (mm) before and after shifting point (year) were computed are presented in Table 4.15 along with nature of shifting and normal rainfall (mm) of each season. The line graph of seasonal rainfall distribution along with rainfall shifting year, and average seasonal rainfall before and after shifting year is shown in Fig. 4.6.

Shifting point (year) in seasonal rainfall was found in the all seasons *viz.* winter (1993), pre-monsoon (2003), monsoon (1970) and post-monsoon (1990). Results in Table 4.15 revealed that the average rainfall was increased for winter (3.69), pre-monsoon (45.71), monsoon (65.13) and post-monsoon (24.90) seasons as compared with before and

after shifting point in rainfall. Further, it can also observed from the Table 4.15, the average seasonal rainfall in winter, pre-monsoon, monsoon and post-monsoon was below normal rainfall before shifting year, but it was increased after shifting year, which was above the normal rainfall.

Malnad subdivision

For each seasonal rainfall data for the period of 60 years from 1960-2019 of Malnad subdivision, shifting point (year), average seasonal rainfall (mm) before and after shifting point (year) were computed are depicted in Table 4.16 along with nature of shifting and normal rainfall (mm) of each season. The line graph of seasonal rainfall distribution along with rainfall shifting year, and average seasonal rainfall before and after shifting year is shown in Fig. 4.7.

Shifting point (year) in seasonal rainfall was found for the season winter (2009), pre-monsoon (1962) and monsoon (1964) and no shifting pattern is observed for post-monsoon season. Results in Table 4.16 revealed that the average rainfall was increased for winter (2.59) season and decreased for pre-monsoon (-188.89) and monsoon (-724.00) seasons as compared with before and after shifting point in rainfall. Further, it can also observed from the Table 4.16, the average seasonal rainfall in winter season was below normal rainfall before shifting year, but it was increased after shifting year, which was above the normal rainfall. Whereas for pre-monsoon and monsoon seasons the average seasonal rainfall were above normal rainfall before shifting year, but it was decreased after shifting year, which was below the normal rainfall.

Coastal subdivision

For each seasonal rainfall data for the period of 60 years from 1960-2019 of Coastal subdivision, shifting point (year), average seasonal rainfall (mm) before and after shifting point (year) were computed are illustrated in Table 4.17 along with nature of shifting and normal rainfall (mm) of each season. The line graph of seasonal rainfall distribution along with rainfall shifting year, and average seasonal rainfall before and after shifting year is shown in Fig. 4.8.

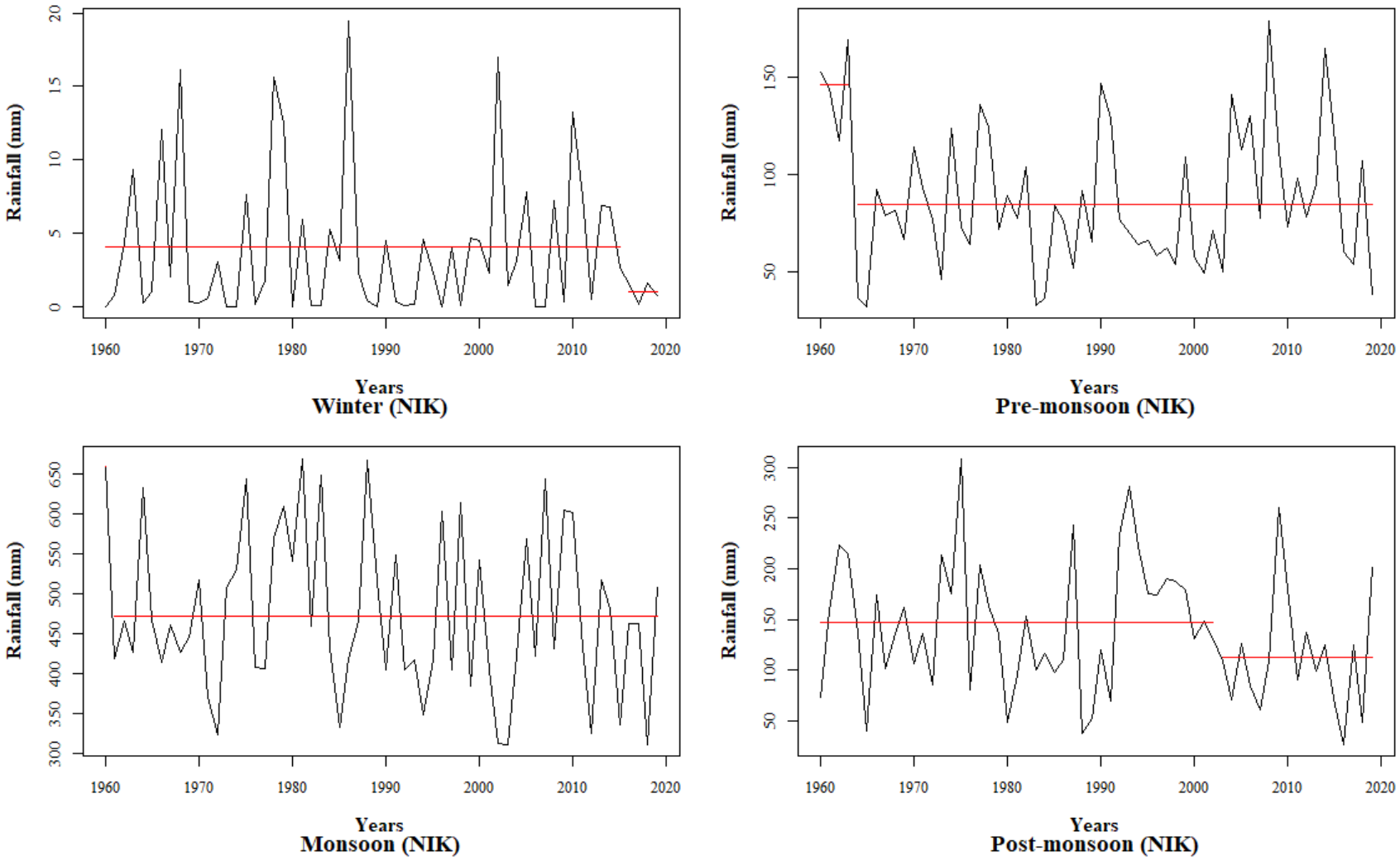

Fig. 4.5: Line graphs showing rainfall distribution along with shifting pattern for all four seasons for NIK subdivision

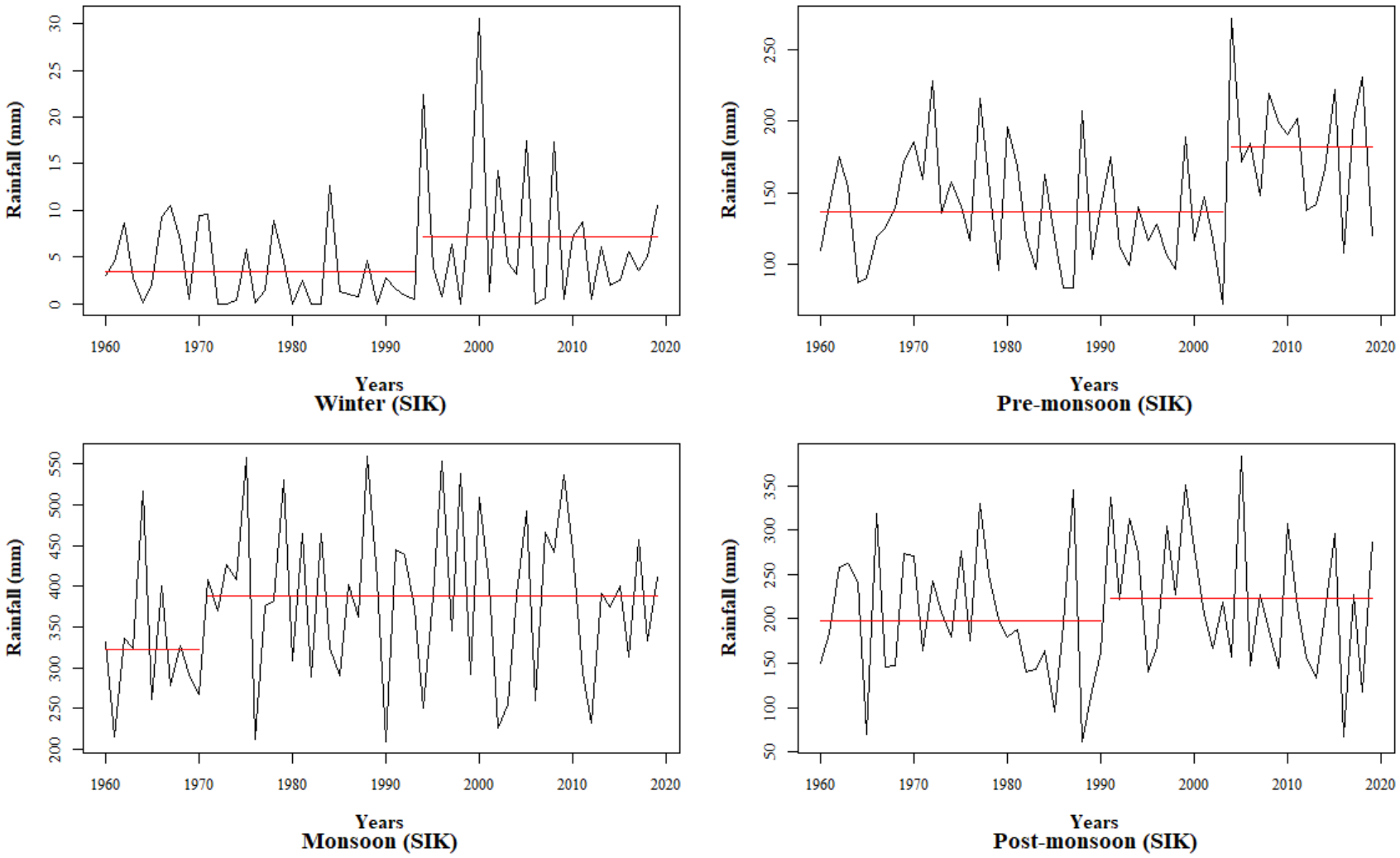

Fig.4.6: Line graphs showing rainfall distribution along with shifting pattern for all four seasons for SIK subdivision

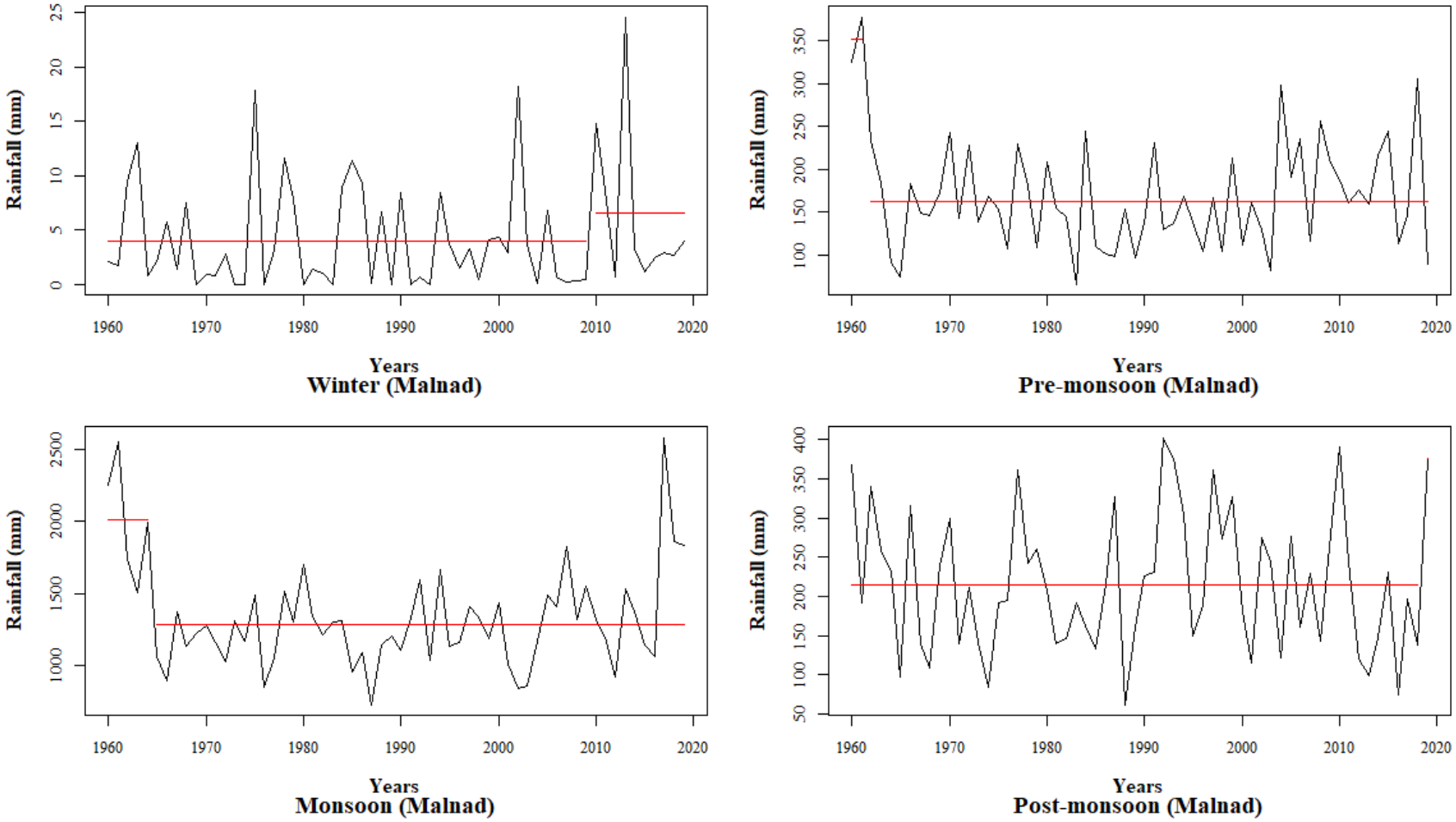

Fig. 4.7: Line graphs showing rainfall distribution along with shifting pattern for all four seasons for Malnad subdivision

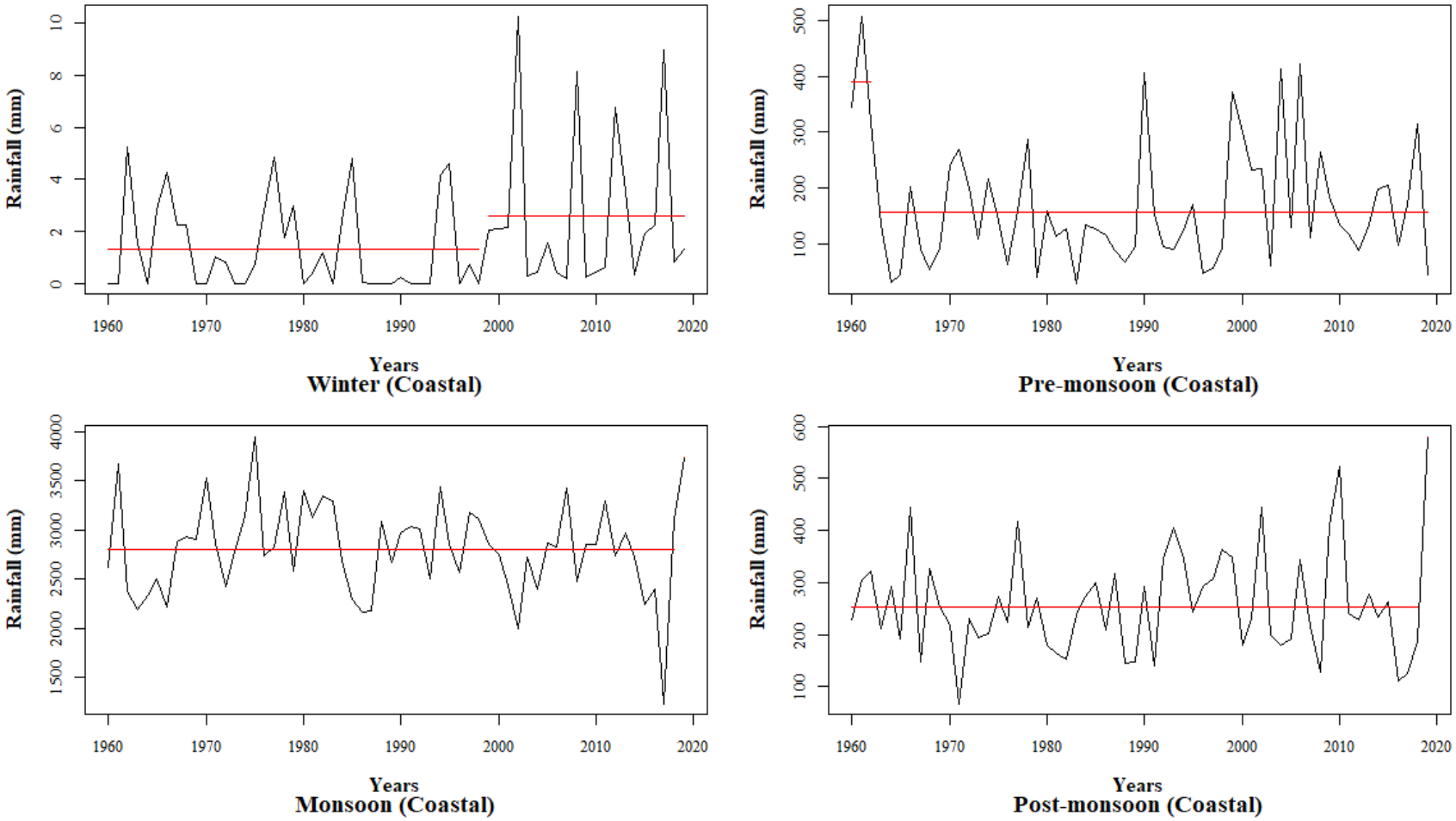

Fig. 4.8: Line graphs showing rainfall distribution along with shifting pattern for all four seasons for Coastal subdivision

Table 4.14: Shifting pattern of seasonal rainfall (mm) data of NIK subdivision

Period	Shifting year	Average rainfall (mm)		Change in rainfall (mm)	Nature of shifting	Normal rainfall (mm)
		Before	After			
Winter	2015	4.03	1.00	-3.03	Decreasing	4.00
Pre-monsoon	1963	146.13	84.53	-61.6	Decreasing	85.00
Monsoon	1960	659.20	471.80	-187.4	Decreasing	494.00
Post-monsoon	2002	147.22	113.08	-34.14	Decreasing	145.00

Table 4.15: Shifting pattern of seasonal rainfall (mm) data of SIK subdivision

Period	Shifting year	Average rainfall (mm)		Change in rainfall (mm)	Nature of shifting	Normal rainfall (mm)
		Before	After			
Winter	1993	3.45	7.14	3.69	Increasing	5.00
Pre-monsoon	2003	136.51	182.22	45.71	Increasing	145.00
Monsoon	1970	322.72	387.85	65.13	Increasing	359.00
Post-monsoon	1990	197.99	222.89	24.9	Increasing	210.00

Table 4.16: Shifting pattern of seasonal rainfall (mm) data of Malnad subdivision

Period	Shifting year	Average rainfall (mm)		Change in rainfall (mm)	Nature of shifting	Normal rainfall (mm)
		Before	After			
Winter	2009	3.90	6.49	2.59	Increasing	4.00
Pre-monsoon	1962	351.38	162.49	-188.89	Decreasing	178.00
Monsoon	1964	2007.55	1283.55	-724	Decreasing	1504.00
Post-monsoon	-	217.28	217.28	-	No shift	228.00

Table 4.17: Shifting pattern of seasonal rainfall (mm) data of Coastal subdivision

Period	Shifting year	Average rainfall (mm)		Change in rainfall (mm)	Nature of shifting	Normal rainfall (mm)
		Before	After			
Winter	1998	1.33	2.60	1.27	Increasing	1.00
Pre-monsoon	1962	390.17	155.97	-234.2	Decreasing	170.00
Monsoon	-	2812.92	2812.92	-	No shift	3019.00
Post-monsoon	-	259.27	259.27	-	No shift	261.00

Shifting point (year) in seasonal rainfall was found for the season winter (1998) and pre-monsoon (1962) and no shifting pattern is observed for monsoon and post-monsoon seasons. Narayanan *et al.* (2016) obtained similar results using Pettitt-Mann-Whitney test for seasonal and annual rainfall data for their study region. Results in Table 4.17 revealed that the average seasonal rainfall was increased for winter (1.27) season and decreased for pre-monsoon (-234.20) season as compared with before and after shifting point in rainfall. Further, it can also observed from the Table 4.17, the average seasonal rainfall in winter and pre-monsoon seasons were above normal rainfall before shifting year, but it was decreased for pre-monsoon after shifting year, which was below the normal rainfall, except for winter season which is still above normal rainfall.

4.2.3 Shifting pattern analysis for annual rainfall data

For each annual rainfall data for the period of 60 years from 1960-2019 for all four subdivisions of Karnataka, shifting point (year), average annual rainfall (mm) before and after shifting point (year) were computed are tabulated in Table 4.18 along with nature of shifting and normal rainfall (mm) of each subdivision. The line graph of annual rainfall distribution along with rainfall shifting year, and average annual rainfall before and after shifting year is shown in Fig. 4.9.

Shifting point (year) in annual rainfall was found in all subdivisions namely NIK (2010), SIK (1968), Malnad (2016) and Coastal (1966) subdivisions. Chakraborty *et al.* (2013) obtained similar results. Results in Table 4.18 revealed that the average annual rainfall was decreased for NIK (-89.21) and increased for SIK (91.21), Malnad (858.66) and Coastal (526.87) subdivisions as compared with before and after shifting point in rainfall. Further, it can also observed from the Table 4.18, the average annual rainfall in NIK and Coastal subdivisions were stayed below normal rainfall before and after shifting year. Whereas the average annual rainfall in SIK and Malnad subdivisions were below normal rainfall before shifting year and increased after shifting year, which was above the normal rainfall.

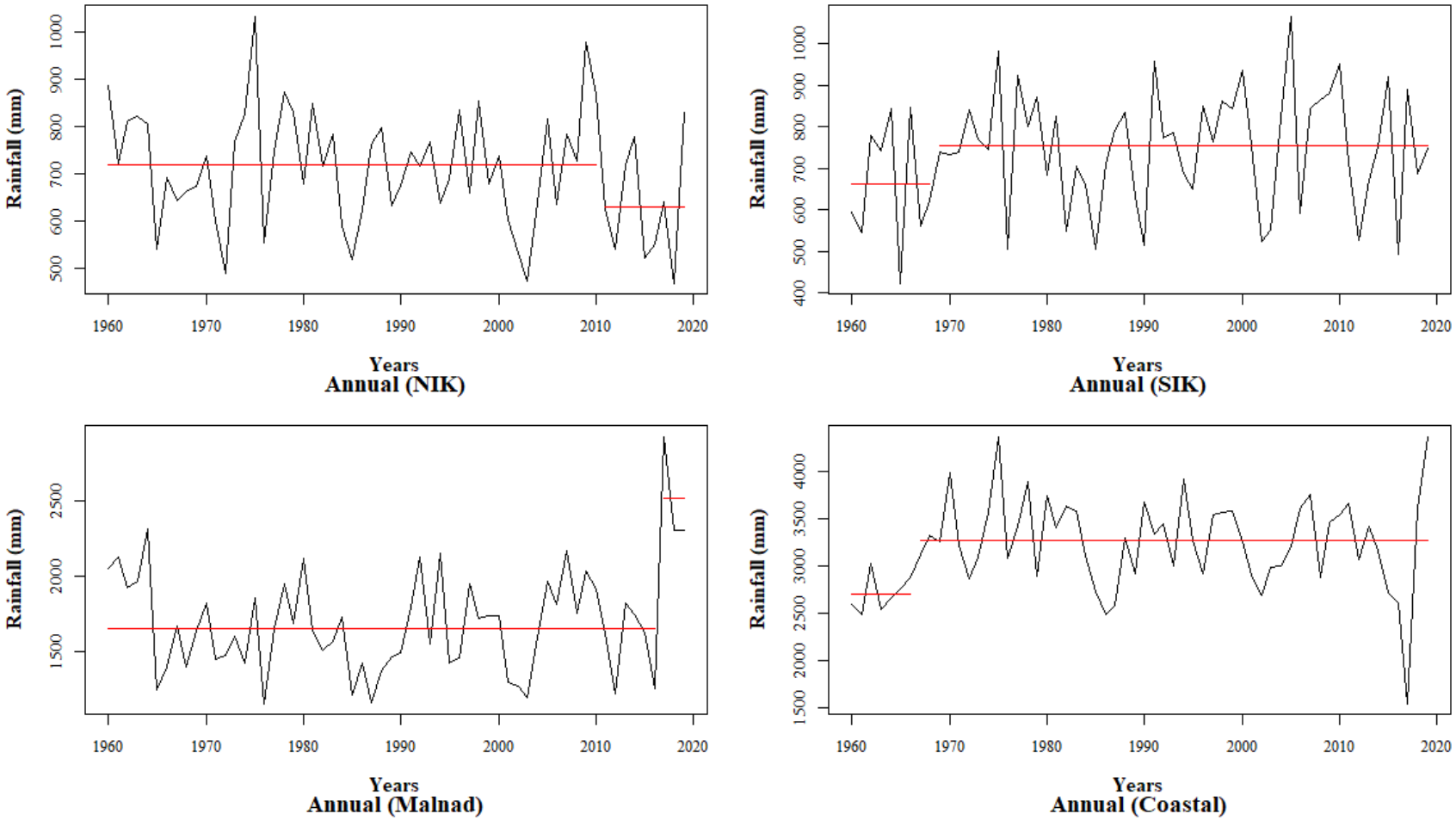

Fig. 4.9: Line graphs showing rainfall distribution along with shifting pattern for annual rainfall data for all four subdivisions

Table 4.18: Shifting pattern of annual rainfall (mm) data for all four subdivision

Period	Shifting year	Average rainfall (mm)		Change in rainfall (mm)	Nature of shifting	Normal rainfall (mm)
		Before	After			
NIK	2010	720.12	630.91	-89.21	Decreasing	728.00
SIK	1968	662.01	753.22	91.21	Increasing	719.00
Malnad	2016	1653.58	2512.24	858.66	Increasing	1914.00
Coastal	1966	2702.83	3265.70	562.87	Increasing	3451.00

4.3 To evaluate a suitable non-linear time series models for forecasting rainfall

In the present study, three non-linear time series models namely, Autoregressive Conditional Heteroscedasticity (ARCH), Generalized Autoregressive Conditional Heteroscedasticity (GARCH) and Artificial Neural Networks (ANN) models were fitted to month-wise rainfall data of four meteorological subdivisions for the period of 60 years from 1960 to 2019 to evaluate suitable nonlinear time-series model for forecasting rainfall data. In addition to these nonlinear models, two linear models *viz*. Seasonal Autoregressive Integrated Moving Average (SARIMA) and Holt's-Winters Exponential Smoothing (H-WES) were also employed to compare the performance with nonlinear models for forecasting rainfall data. First, month-wise rainfall data was divided into two set of data, the first one was training data set and second was testing data set, *i.e.*, month-wise rainfall data of 59 years from 1960 to 2018 as training data set, and one year data of 2019 as testing data set. The training data set was used to build H-WES, SARIMA, ARCH and GARCH models, and testing data set was used for evaluating forecasting performance of these models. Whereas, for fitting ANN models, month-wise rainfall data was divided into three data sets namely training, validation and testing data sets *i.e.,* month-wise rainfall data of 57 years from 1960 to 2016 as training, two years data of 2017 and 2018 as validation, and one year data of 2019 as testing data. The training data set was used to build ANN model, validation data set was used to find the optimal number of hidden nodes and transformation function, and testing data set was used for asses the forecasting performance of model. RMSE value was computed to find the best-fitted model.

H-WES method: H-WES additive model was used for forecasting month-wise rainfall data. The combination of three smoothing parameters α, β and γ were selected using grid search method which gives the lowest RMSE value. The training data set was used for H-WES model building and testing data set was used for evaluating forecasting performance of the model. The fitted model was used for back forecasting and one year ahead forecasting. The RMSE value for both training and testing data sets were calculated using back and forward forecasted values. H-WES model was fitted with the help of IBM-SPSS software Version 25.

SARIMA: This approach involves several steps in building model for forecasting rainfall. First, the time-series data shall be visualized to check the stationarity (by using ACF and PACF plots). If the data series is not stationary, then convert it to stationary by using technique of difference. Once the stationarity of series achieved, the ACF and PACF plots are examined to find out the tentative SARIMA model *i.e.,* order of seasonal (*P,D,Q*) and non-seasonal *(p,d,q)*. The tentative model parameters were estimated using the maximum likelihood estimator. The model that gives lowest AIC was considered as best fitted model for forecasting. The training data set was used for SARIMA model building and testing data set was used for evaluating forecasting performance of the model. The fitted model was used for back forecasting and one year ahead forecasting. The RMSE value for both training and testing data sets were calculated using back and forward forecasted values. The '***astsa***' and '***forecast***' packages of R-software were employed for fitting SARIMA models. The residual analysis of fitted models was carried by plotting the ACF and PACF plots and the Ljung–Box test statistic.

Volatility models: Tentative models with different order of *q* for the ARCH model and different combinations of order *p* (for error term) and *q* (for squared variance term) for the GARCH model are constructed. Parameters of tentative models were estimated using the maximum likelihood estimator. Model with lowest AIC value was considered as the best-fitted model for forecasting purpose. The training data set was used for volatility model building and testing data set was used for evaluating forecasting performance of the model. The fitted model was used for back forecasting and one year ahead forecasting. The RMSE value for both training and testing data sets were calculated using back and forward

forecasted values. For fitting ARCH and GARCH models, package '***fGarch***' of R-software was employed.

Artificial Neural Networks (ANN): Multi-Layer Perceptron Neural Network (MLP-NN) model with 12 input nodes, one hidden layer and one output node were fixed prior fitting ANN model. Suitable activation function and number of hidden nodes (h) were determined by trial and error process using validation data set. ANN models with three different activation function (*i.e.,* Logistic, Hyperbolic tangent and Linear) and different number of hidden nodes (h = 1, 2, …) were trained, one at a time. Each trained models were used to validate the model parameters using validation data set by computing RMSE value. Model with lowest RMSE on validation set was considered as best fitted model. Then, training and validation data sets were clubbed together to form new training data set, which was used to forecast for testing data set. The training data set was used for ANN model building and testing data set was used for evaluating forecasting performance of the model. The fitted model was used for back forecasting and one year ahead forecasting. The RMSE value for both training and testing data sets were calculated using back and forward forecasted values. For building ANN model, Feed-forward back propagation neural network algorithm available in '***nnfor***' package of R-software was used.

4.3.1 H-WES, SARIMA, ARCH, GARCH and ANN models for forecasting monthly rainfall for NIK subdivision

H-WES model: Different combination of smoothing parameters α, β and γ were tried by grid search method to get suitable H-WES parameters. The parameter estimates of H-WES model along with standard error for NIK meteorological subdivision are given in Table 4.19. The parameter estimates of H-WES model $\alpha = 0.01, \beta = 0.00000004$ and $\gamma = 0.001$ were found to be suitable for NIK subdivision with lowest RMSE. The p-values for α, β and γ are respectively 0.24, 1.00 and 0.93 indicates that smoothing parameters of H-WES model are non-significant. Therefore, this model is not suitable for forecasting monthly rainfall for future time period. RMSE value calculated for both training and testing data sets are respectively 35.34 and 28.70 and are illustrated in Table 4.27.

SARIMA model: Distribution pattern of monthly rainfall data along with its ACF and PACF plot for NIK subdivision are depicted in Fig. 4.10. The plots of ACF and PACF at seasonal lags revealed that decay rate was too slow, which indicates that the data was non-stationary at seasonal lags. Hence, seasonal differencing at 12^{th} lag ($D = 1$) was performed to achieve stationarity. The plots of seasonal differenced monthly rainfall data along with its ACF and PACF plots presented in the Fig. 4.11. ACF and PACF plots of seasonally differenced series indicates that data was stationary. So, by examining the plots of ACF and PACF, tentative models were identified. The list of tentatively identified models along with their AIC values and Ljung-Box statistic for residuals are tabulated in the Table 4.20. The parameters of model were estimated using maximum likelihood estimator. To check the adequacy of the model, residual analysis was carried out using Ljung-Box test statistic. Based on the lowest AIC value (6998.62), SARIMA (0,0,0) $(2,1,2)_{12}$was found to be the best-fitted model among the tentative models. Estimated model parameters with their respective standard errors, t-statistic and p-values are exhibited in the Table 4.21.

The ACF and PACF plots of residuals obtained from best-fitted model are displayed in the Fig. 4.12. The Ljung-Box test statistic (1.81) presented in Table 4.20 for best-fitted model was found to be non-significant at 5 per cent level of significance (p-value 0.19) indicates that errors are independent and the model is adequate. RMSE value computed for training and testing data sets were respectively 35.12 and 27.74 and are presented in Table 4.27.

ARCH model: The ARCH-LM test was carried out to check the volatility or ARCH effect in the residuals of SARIMA model. ARCH-LM test statistic value of 51.23 with p-value < 0.01 reveals that, there is highly significant ARCH effect in the best-fitted SARIMA model for NIK subdivision. Therefore, ARCH and GARCH models were built for NIK subdivision.

Tentative identified models along with their AIC values are tabulated in the Table 4.22. Initially, four tentative ARCH models were fitted and the model with AIC values (10.19), ARCH (1) is selected as best fitted model among the tentative models.

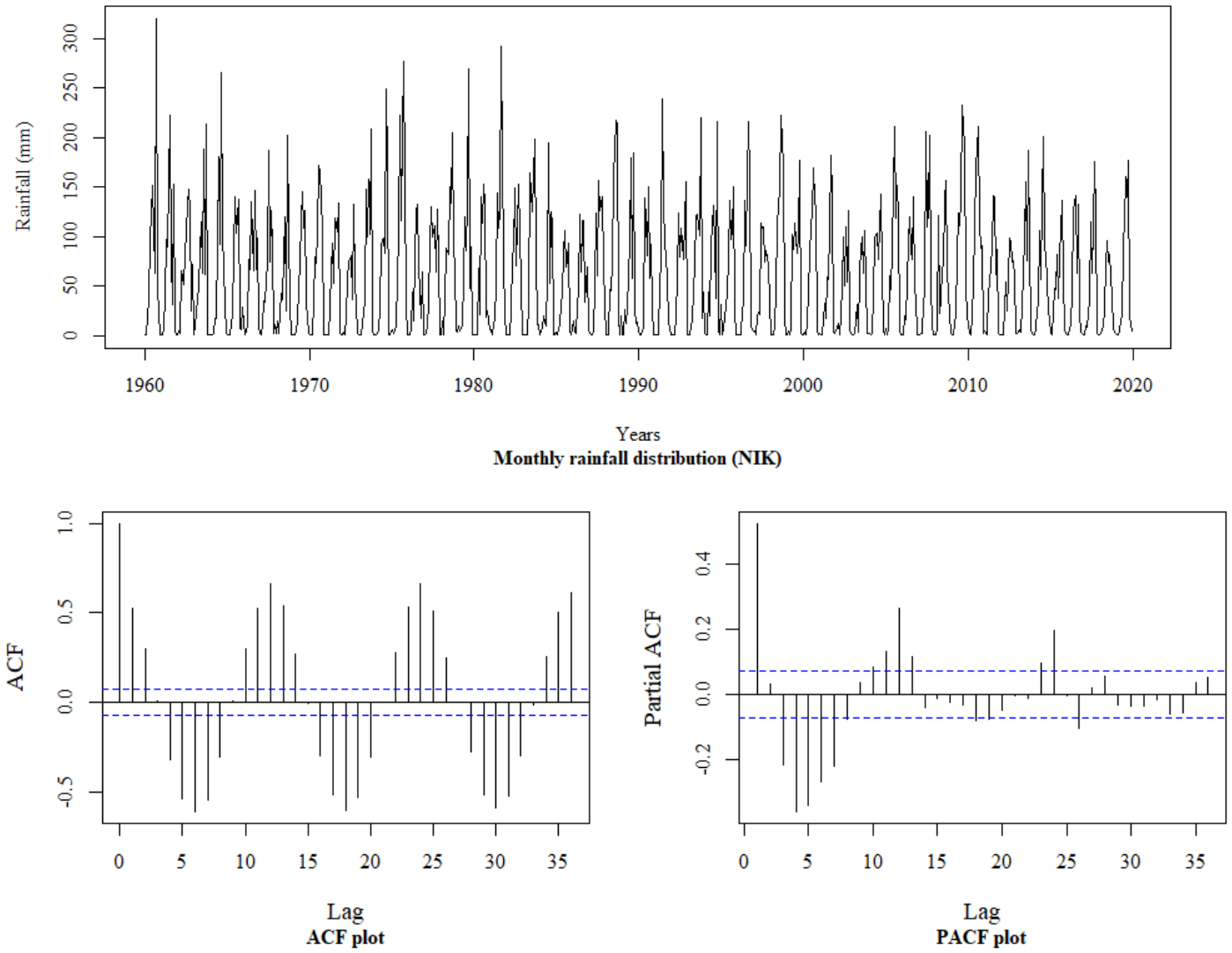

Fig. 4.10: Line graph for monthly rainfall data along with its ACF and PACF plots for NIK subdivision

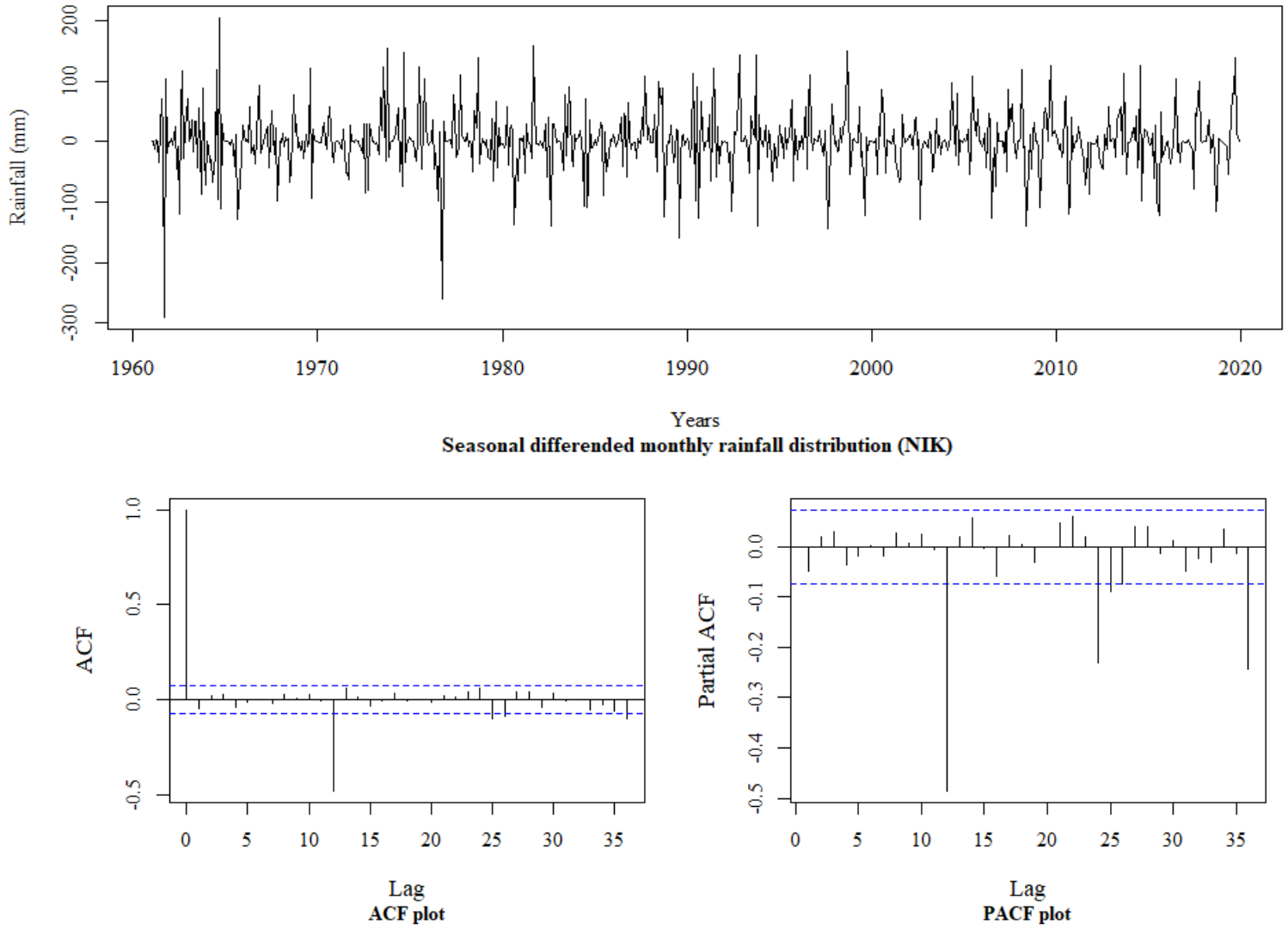

Fig. 4.11: Line graph for seasonally differenced monthly rainfall data along with its ACF and PACF plots for NIK subdivision

Table 4.19: Parameter estimates of H-WES model for monthly rainfall data of NIK subdivision

Model Parameter	Estimates	S.E	t-statistic	p-value
α	0.01 ^{NS}	0.005	1.18	0.24
β	0.00000004 ^{NS}	0.0002	0.0002	1.00
γ	0.001^{NS}	0.01	0.09	0.93

NS: Non-Significant

Table 4.20: Tentatively identified SARIMA models for monthly rainfall data of NIK subdivision

Tentative models	AIC value	Ljung-Box statistic
$(0,0,0)\ (3,1,3)_{12}$	7001.67	1.83^{NS}
$(0,0,0)\ (2,1,2)_{12}$	**6998.62**	**1.81^{NS}**
$(0,0,0)\ (1,1,2)_{12}$	7003.80	0.76^{NS}
$(0,0,0)\ (2,1,1)_{12}$	7002.65	2.71^{NS}
$(1,0,1)\ (2,1,2)_{12}$	7000.48	0.00^{NS}
$(0,0,0)\ (2,1,0)_{12}$	7164.92	2.39^{NS}
$(1,0,0)\ (2,1,0)_{12}$	7164.61	0.00^{NS}
$(0,0,1)\ (2,1,0)_{12}$	7164.82	0.00^{NS}
$(0,0,0)\ (0,1,2)_{12}$	7002.19	2.89^{NS}
$(1,0,0)\ (0,1,2)_{12}$	7001.73	0.00^{NS}
$(0,0,1)\ (0,1,2)_{12}$	7001.99	0.04^{NS}
$(0,0,0)\ (0,1,1)_{12}$	7001.06	2.85^{NS}
$(0,0,0)\ (1,1,0)_{12}$	7204.59	1.62^{NS}

Table 4.21: Parameter estimates of SARIMA $(0,0,0)\ (2,1,2)_{12}$ model for monthly rainfall data of NIK subdivision

Model Parameter	Estimates	S.E	t-statistic	p-value
SAR(1)	-0.78**	0.11	6.83	<0.01
SAR(2)	0.10*	0.04	2.33	0.02
SMA(1)	-0.18^{NS}	0.12	1.51	0.13
SMA(2)	-0.82**	0.11	7.20	<0.01

* Significant at 5% level of significance, ** Significant at 1% level of significance, NS: Non-Significant

Parameters are estimated using maximum likelihood method. Estimated model parameters along with their respective standard errors, t-statistic and p-values are shown in the Table 4.23. Estimated parameters are found to be significant at 1 per cent level of significance, which indicates selected model is adequate model. RMSE value calculated are respectively 53.33 and 60.09 for trained and testing data sets, which are depicted in Table 4.27.

GARCH model: Tentative identified models along with their AIC values are tabulated in the Table 4.24. Initially, four tentative GARCH models were fitted and the model with AIC values (10.21), GARCH (1,1) is selected as best fitted model among the tentative models. Parameters are estimated using maximum likelihood method. Estimated model parameters along with their respective standard errors, t-statistic and p-values are presented in the Table 4.25. Estimated parameter α_0 was found to be significant at 1 per cent level of significance and parameters α_1 and β_1 were non-significant indicates that selected model was not adequate to explain heteroscedasticity in the residual of SARIMA model. RMSE value computed respectively for training and testing data sets are 52.96 and 41.88 as tabulated in Table 4.27.

Artificial Neural Networks (ANN): The most commonly used neural network for time-series forecasting, MLP feed forward neural network, was adopted in this study. Training data set was used to train the MLP-NN model, validation data set was used for validate model parameters (*i.e.*, find optimal number of hidden nodes and activation function), and finally testing data set was used to measure the accuracy of forecasting. ANN model with 12 input nodes, one hidden layer and one output node were fixed prior to fitting ANN model. Three different activation function namely Logistic, Hyperbolic tangent (Tanh) and Linear, and various number of hidden nodes (h = 1, 2, …) were trained one at a time to find the optimal combination of these parameters by computing RMSE value on validation set which are depicted in Table 4.26. Model with lowest RMSE for validation set was considered as best fitted model.

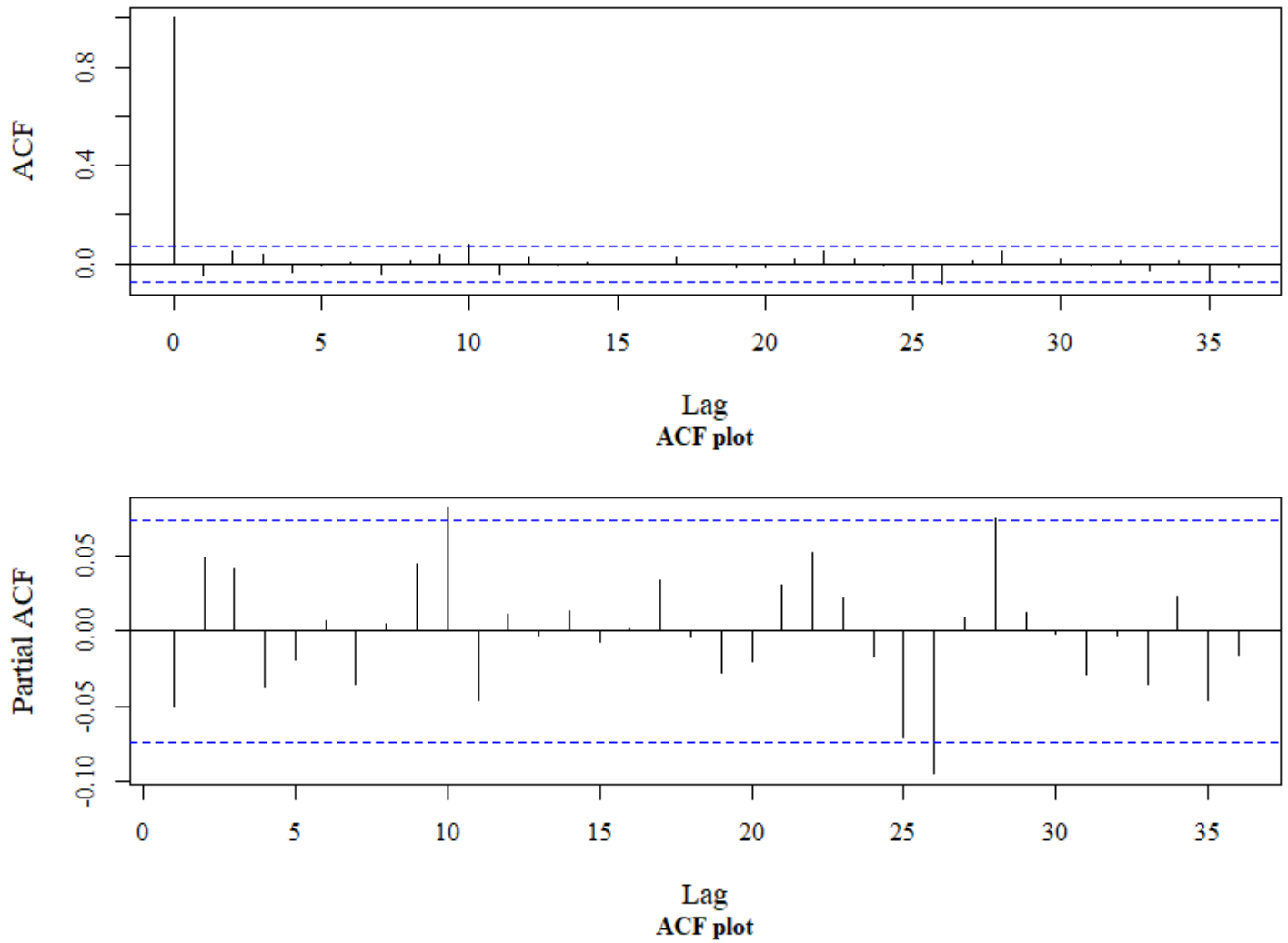

Fig. 4.12: ACF and PACF plots of residuals of SARIMA (0,0,0) $(2,1,2)_{12}$ model for monthly rainfall data for NIK subdivision

Table 4.22: Tentatively identified ARCH models for monthly rainfall data of NIK subdivision

Tentative models	AIC value
ARCH(1)	**10.19**
ARCH(2)	10.74
ARCH(3)	10.80
ARCH(4)	10.75

Table 4.23: Parameter estimates of ARCH (1) model for monthly rainfall data of NIK subdivision

Model Parameter	Estimates	S.E	t-statistic	p-value
α_0	0.13**	0.50	2.62	<0.01
α_1	1352.94**	1.41	957.81	0.01

** Significant at 1% level of significance

Table 4.24: Tentatively identified GARCH models for monthly rainfall data of NIK subdivision

Tentative models	AIC value
GARCH(1,1)	**10.21**
GARCH(2,1)	10.80
GARCH(1,2)	11.10
GARCH(2,2)	11.11

Table 4.25: Estimates of GARCH (1, 1) model parameter for monthly rainfall data of NIK subdivision

Model Parameter	Estimates	S.E	t-statistic	p-value
α_0	141.22**	1.42	99.21	<0.01
α_1	0.000001^{NS}	0.000024	0.05	0.99
β_1	0.91^{NS}	0.95	0.95	0.34

** Significant at 1% level of significance, NS: Non-Significant

Table 4.26: RMSE values for ANN model over different activation functions on validation data set of monthly rainfall data of NIK subdivision

Input nodes	Hidden Nodes	RMSE values		
		Logistic	Tanh	Linear
12	1	16.08	11.15	20.55
12	2	9.10	7.27	15.76
12	3	8.69	8.41	16.20
12	**4**	11.96	**7.24**	13.60
12	5	12.08	9.45	18.09
12	6	13.87	9.53	15.53

Table 4.27: Forecast accuracy measures for training and testing data sets for best-fitted models of NIK subdivision

	H-WES	SARIMA	ARCH	GARCH	ANN
Training	35.34	35.12	53.53	52.96	**28.04**
Testing	28.70	27.74	60.09	41.88	**25.29**

Table 4.28: Ex-post forecast of monthly rainfall of NIK subdivision by H-WES, SARIMA, ARCH, GARCH and ANN models

Month-year	Actual rainfall (mm)	Forecasted rainfall (mm) for NIK region				
		H-WES	SARIMA	ARCH	GARCH	ANN
Jan-19	0.50	1.20	0.28	24.30	5.06	2.14
Feb-19	0.20	2.30	1.80	38.92	6.63	0.73
Mar-19	3.00	3.40	7.19	47.71	22.24	0.92
Apr-19	17.00	22.60	27.74	52.91	47.70	19.63
May-19	19.00	49.40	52.92	55.97	76.16	39.19
Jun-19	91.00	97.22	103.56	57.77	99.99	111.48
Jul-19	116.00	113.21	109.90	58.83	112.81	115.26
Aug-19	161.00	108.98	108.74	59.46	111.21	138.47
Sep-19	140.00	136.25	144.63	59.82	95.64	144.22
Oct-19	177.00	98.61	106.34	60.04	70.28	97.76
Nov-19	20.00	22.41	22.14	60.17	41.96	15.86
Dec-19	4.00	1.46	5.42	18.27	4.33	1.80
RMSE		**28.70**	**27.74**	**60.09**	**41.88**	**25.29**

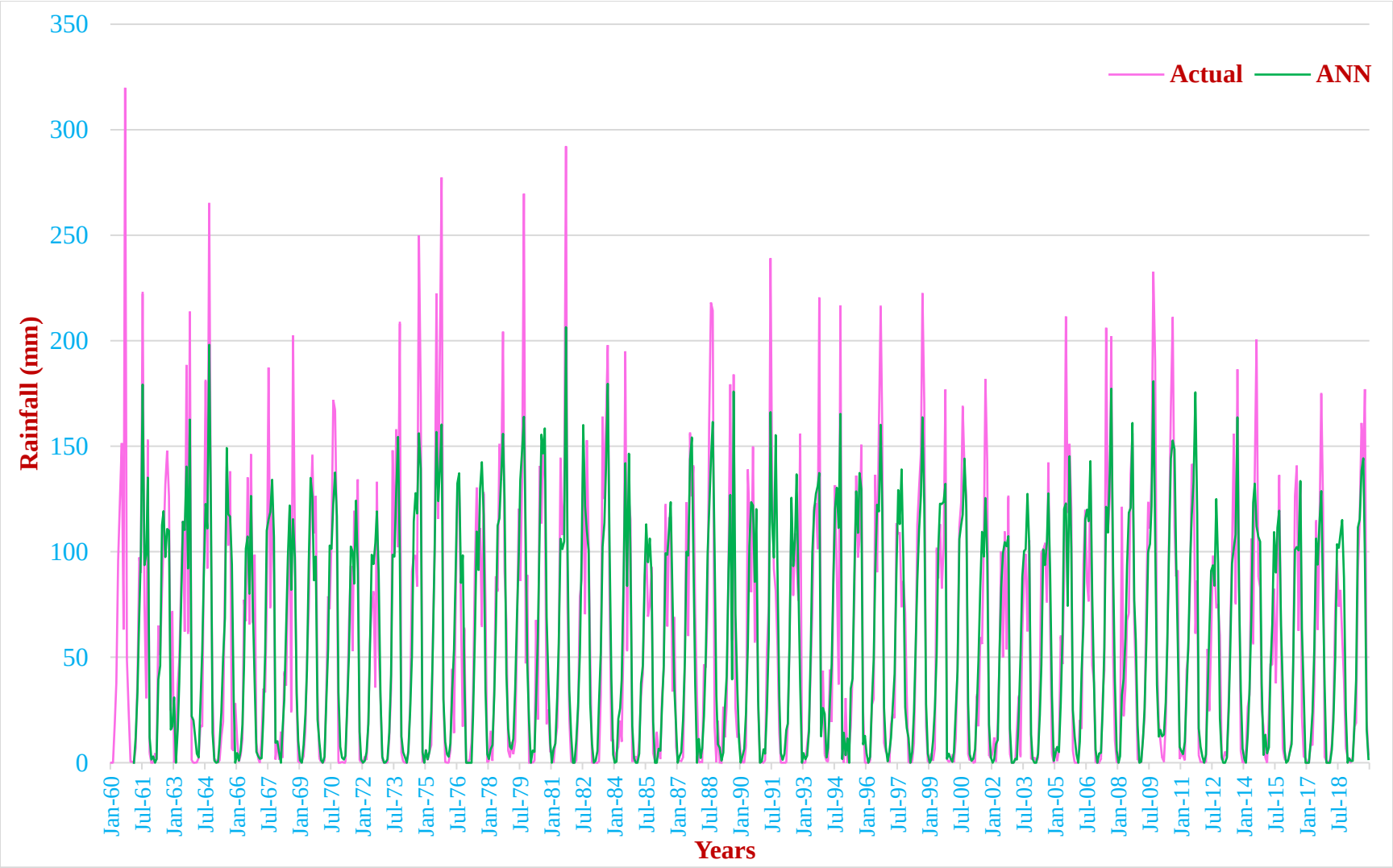

Fig.4.13: Actual and forecasted monthly rainfall (mm) data using ANN model for NIK subdivision

For NIK subdivision, MLP-NN consist of one input layer with 12 input nodes, one hidden layer having four hidden nodes and one output layer with one output node and 'Tanh' activation function was found to be best fitted network for forecasting rainfall with lowest RMSE (7.24) over validation data set. Then, training and validation data sets were clubbed together to form training data set, which was used to train the MLP-NN and to forecast for comparing forecasting accuracy with testing data set. RMSE value computed respectively for training and testing data sets are 28.04 and 25.29 as presented in Table 4.27.

RMSE values were computed for all the fitted models on both training and testing data sets and are reported in Table 4.27. The forecasting performance of all the fitted models are compared based on RMSE value. For training data set, ANN model with lowest RMSE (28.04) has performed better than the other models, followed by SARIMA, H-WES, GARCH and ARCH models with RMSE values 35.12, 35.34, 52.96 and 53.53 respectively. Further, it can also be observed from the Table 4.27 that for testing data also, the ANN model has performed better with lowest RMSE (25.29) as compared to SARIMA (27.74), H-WES (28.70), GARCH (41.88) and ARCH(60.09) models. It may be noted that for both training and testing data, the ANN model performed better than all the other models, which indicates the superiority of ANN model over other models. Therefore, ANN model was selected as a best model for forecasting monthly rainfall of NIK subdivision. The actual rainfall along with predicted by ANN model were plotted in Fig. 4.13. A perusal of Fig. 4.13 indicates that ANN model nicely captures the variations in monthly rainfall of NIK subdivision.

4.3.2 H-WES, SARIMA, ARCH, GARCH and ANN models for forecasting monthly rainfall for SIK subdivision

H-WES model: Different combination of smoothing parameters α, β and γ were tried by grid search method to get suitable H-WES parameters. The parameter estimates of H-WES model along with standard error for SIK meteorological subdivision are given in Table 4.29. The parameter estimates of H-WES model $\alpha = 0.094, \beta = 0.0000003$ and $\gamma = 0.00002$ were found to be suitable for SIK subdivision with lowest RMSE.

The p-values for α, β and γ are respectively <0.01, 1.00 and 0.99 indicates that smoothing parameters of H-WES model are non-significant except α, which was highly significant. Therefore, this model is mostly not suitable for forecasting monthly rainfall for future time period. RMSE value calculated for both training and testing data sets are respectively 37.44 and 35.11 and are tabulated in Table 4.37.

SARIMA model: Distribution pattern of monthly rainfall data along with its ACF and PACF plot for SIK subdivision are depicted in Fig. 4.14. The plots of ACF and PACF at seasonal lags revealed that decay rate was too slow, which indicates that the data was non-stationary at seasonal lags. Hence, seasonal differencing at 12^{th} lag (D = 1) was performed to achieve stationarity. The plots of seasonal differenced monthly rainfall data along with its ACF and PACF plots presented in the Fig. 4.15. ACF and PACF plots of seasonally differenced series indicates that data was stationary. So, by examining the plots of ACF and PACF, tentative models were identified. The list of tentatively identified models along with their AIC values and Ljung-Box statistic for residuals are arranged in the Table 4.30. The parameters of model were estimated using maximum likelihood estimator. To check the adequacy of the model, residual analysis was carried out using Ljung-Box test statistic. Based on the lowest AIC value (7058.89), SARIMA $(0,0,0)(2,1,2)_{12}$was found to be the best-fitted model among the tentative models. Estimated model parameters with their respective standard errors, t-statistic and p-values are exhibited in the Table 4.31.

The ACF and PACF plots of residuals obtained from best-fitted model are displayed in the Fig. 4.16. The Ljung-Box test statistic (0.76) presented in Table 4.30 for best-fitted model was found to be non-significant at 5 per cent level of significance (p-value 0.38) indicates that errors are independent and the model is adequate. RMSE value computed for training and testing data sets were respectively 36.78 and 47.25 and are given in Table 4.37.

ARCH model: The ARCH-LM test was carried out to check the volatility or ARCH effect in the residuals of SARIMA model. ARCH-LM test statistic value of 55.33 with p-value < 0.01 reveals that, there is highly significant ARCH effect in the best-fitted SARIMA model for SIK subdivision. Therefore, ARCH and GARCH models were built for SIK subdivision.

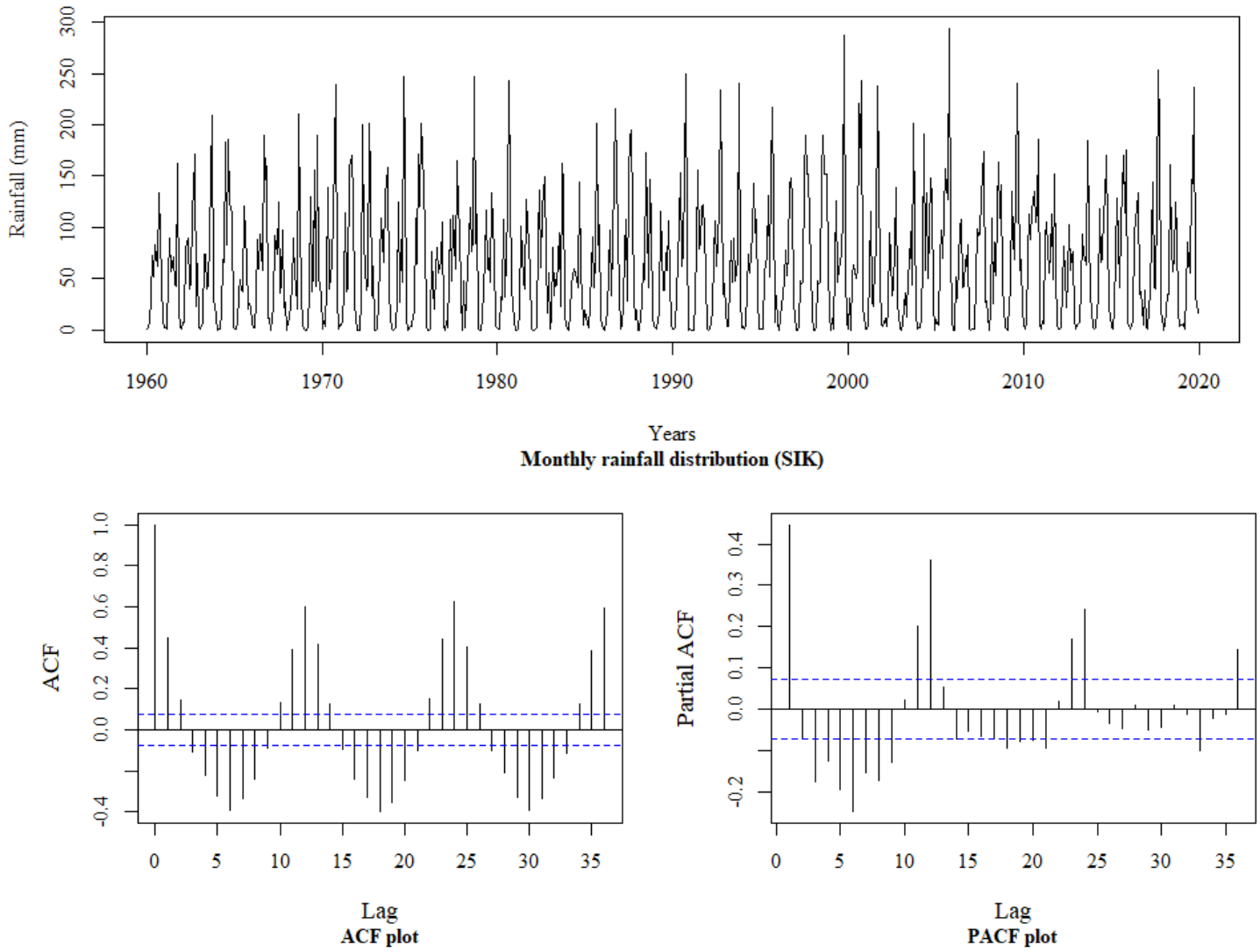

Fig. 4.14: Line graph for monthly rainfall data along with its ACF and PACF plots for SIK subdivision

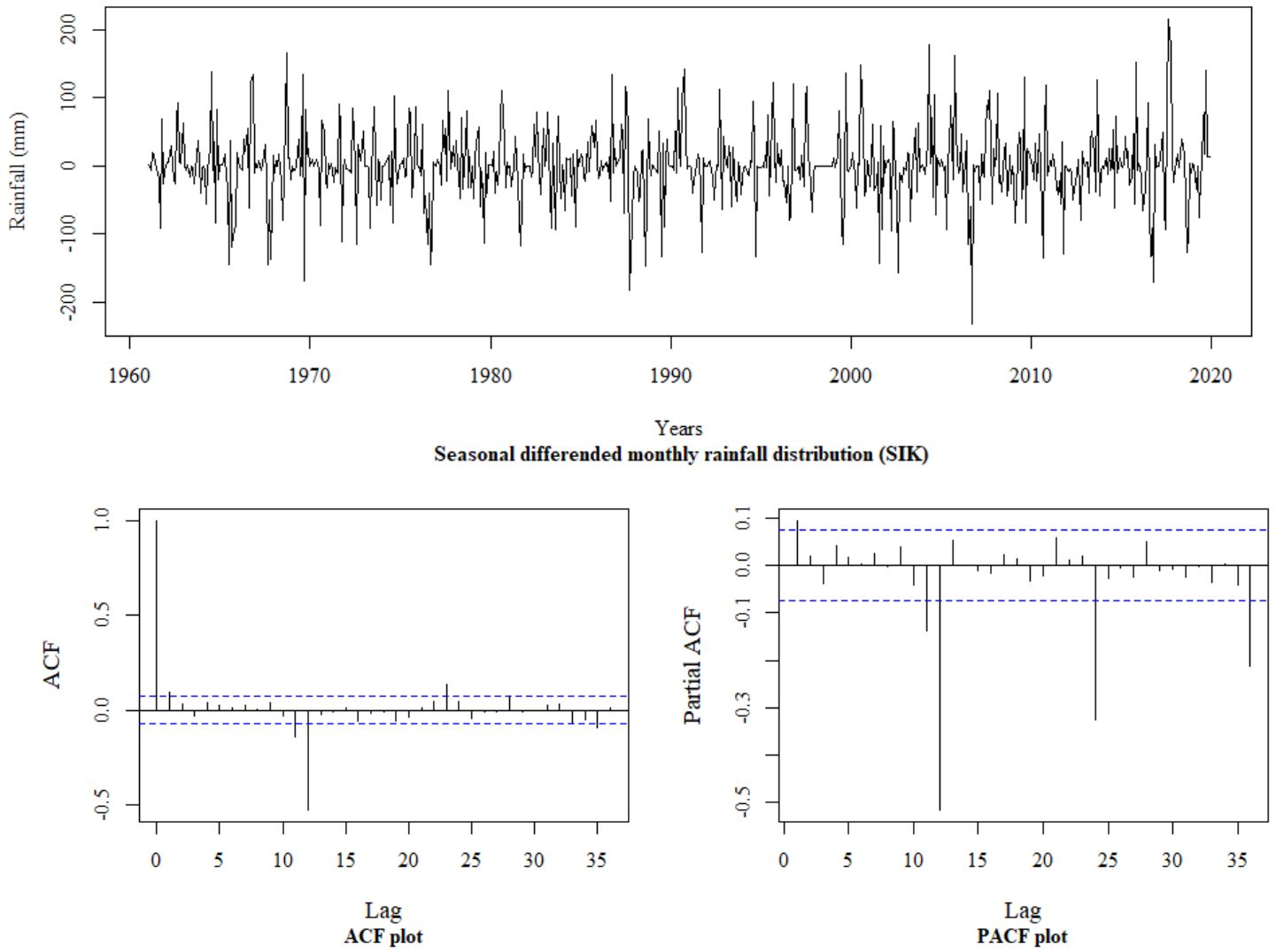

Fig. 4.15: Line graph for seasonally differenced monthly rainfall data along with its ACF and PACF plots for SIK subdivision

Table 4.29: Parameter estimates of H-WES model for monthly rainfall data of SIK subdivision

Model Parameter	Estimates	S.E	t-statistic	p-value
α	0.094**	0.021	4.56	<0.01
β	0.0000003 ^{NS}	0.002	0.0001	1.00
γ	0.00002 ^{NS}	0.012	0.002	0.99

** Significant at 1% level of significance, NS: Non-Significant

Table 4.30: Tentatively identified SARIMA models for monthly rainfall data of SIK subdivision

Tentative models	AIC value	Ljung-Box statistic
$(0,0,0)\ (3,1,3)_{12}$	7062.78	0.75^{NS}
$(0,0,0)\ (2,1,2)_{12}$	**7058.89**	**0.76^{NS}**
$(0,0,0)\ (1,1,2)_{12}$	7061.64	0.47^{NS}
$(0,0,0)\ (2,1,0)_{12}$	7188.82	2.14^{NS}
$(1,0,0)\ (2,1,0)_{12}$	7188.71	0.00^{NS}
$(0,0,1)\ (2,1,0)_{12}$	7188.75	0.00^{NS}
$(0,0,0)\ (0,1,2)_{12}$	7190.61	0.00^{NS}
$(1,0,0)\ (0,1,2)_{12}$	7061.13	0.02^{NS}
$(0,0,1)\ (0,1,2)_{12}$	7061.16	0.01^{NS}
$(0,0,0)\ (0,1,1)_{12}$	7062.85	0.00^{NS}
$(0,0,0)\ (1,1,0)_{12}$	7204.59	1.62^{NS}

Table 4.31: Parameter estimates of SARIMA $(0,0,0)\ (2,1,2)_{12}$ model of monthly rainfall data for SIK subdivision

Model Parameter	Estimates	S.E	t-statistic	p-value
SAR(1)	0.72**	0.17	4.31	<0.01
SAR(2)	0.06^{NS}	0.04	1.47	0.14
SMA(1)	-1.72**	0.16	10.57	<0.01
SMA(2)	0.72**	0.16	4.45	<0.01

** Significant at 1% level of significance, NS: Non-Significant

Tentative identified models along with their AIC values are presented in the Table 4.32. Initially, four tentative ARCH models were fitted and the model with AIC values (10.46), ARCH(1) is selected as best fitted model among the tentative models. Parameters are estimated using maximum likelihood method. Estimated model parameters along with their respective standard errors, t-statistic and p-values are tabulated in the Table 4.33. The p-values for estimated parameters α_0 (<0.01) and α_1 (0.40) respectively indicates highly significance and non-significance of parameters, which indicates that the fitted model was inadequate model. RMSE value calculated are respectively 50.72 and 67.00 for trained and testing data sets, which are tabulated in Table 4.37.

GARCH model: Tentative identified models along with their AIC values are tabulated in the Table 4.34. Initially, four tentative GARCH models were fitted and the model with AIC values (10.48), GARCH (1,1) is selected as best fitted model among the tentative models. Parameters are estimated using maximum likelihood method. Estimated model parameters along with their respective standard errors, t-statistic and p-values are formulated in the Table 4.35. Estimated parameters α_0 and β_1 were found to be significant at 1 per cent level of significance and parameter α_1 was non-significant indicating that selected model is not adequate to explain heteroscedasticity in the residuals of SARIMA model. RMSE value computed respectively for training and testing data sets are 45.05 and 52.24 as tabulated in Table 4.37.

Artificial Neural Networks (ANN): The most commonly used neural network for time-series forecasting, MLP feed forward neural network, was adopted in this study. Training data set was used to train the MLP-NN model, validation data set was used for validate model parameters (*i.e.,* find optimal number of hidden nodes and activation function), and finally testing data set was used to measure the accuracy of forecasting. ANN model with 12 input nodes, one hidden layer and one output node were fixed prior to fitting ANN model. Three different activation function namely Logistic, Hyperbolic tangent (Tanh) and Linear, and various number of hidden nodes (h = 1, 2, …) were trained one at a time to find the optimal combination of these parameters by computing RMSE value on validation set, which are arranged in Table 4.36. Model with lowest RMSE for validation set was considered as best fitted model.

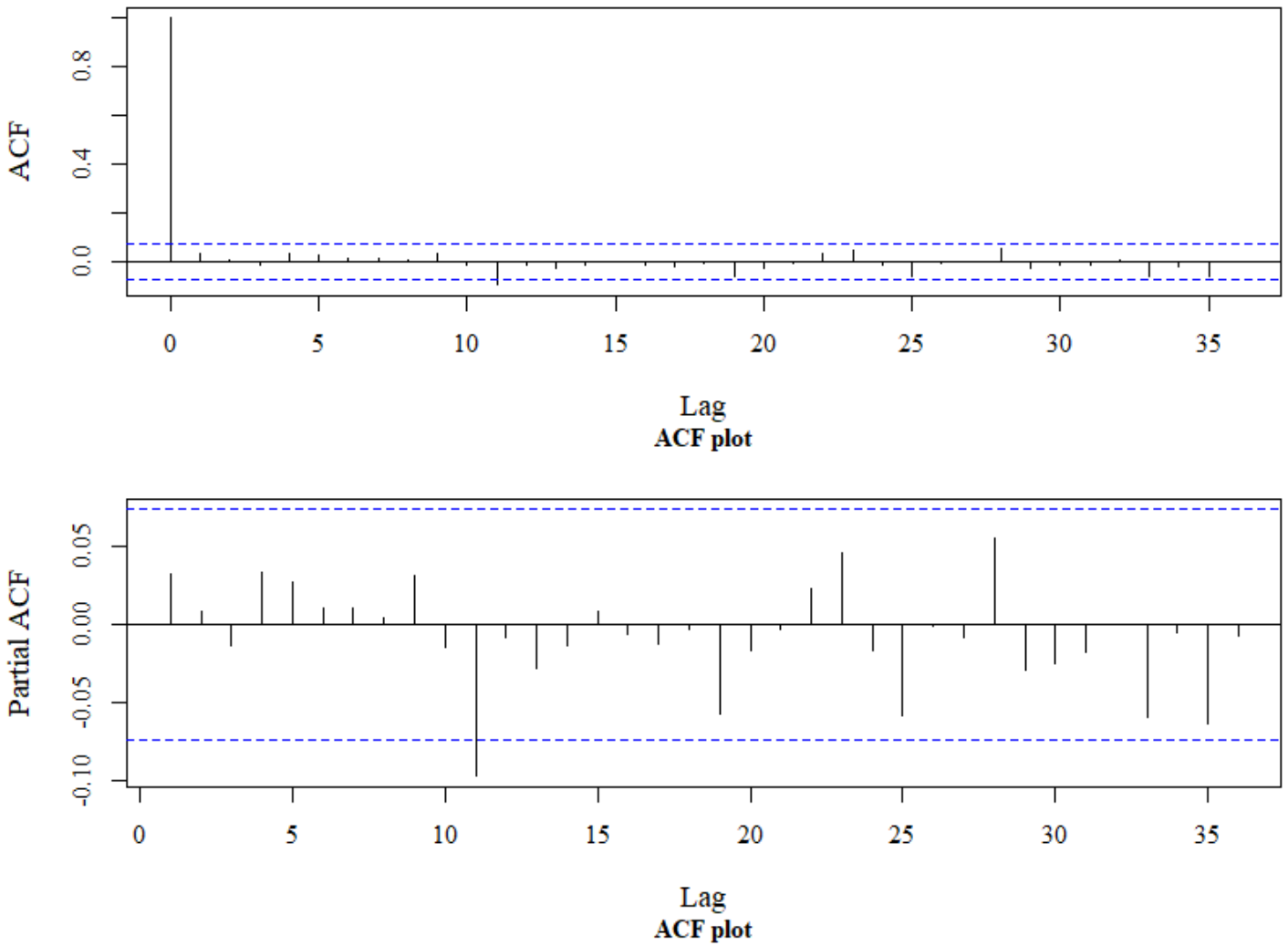

Fig. 4.16: ACF and PACF plots of residuals of SARIMA (0,0,0) $(2,1,2)_{12}$ model for monthly rainfall data for SIK subdivision

Table 4.32: Tentatively identified ARCH models for monthly rainfall data of SIK subdivision

Tentative models	AIC value
ARCH(1)	**10.46**
ARCH(2)	10.79
ARCH(3)	10.80
ARCH(4)	10.83

Table 4.33: Estimates of ARCH (1) model parameter for monthly rainfall data of SIK subdivision

Model Parameter	Estimates	S.E	t-statistic	p-value
α_0	1699.98**	1.60	1060.99	<0.01
α_1	0.17^{NS}	0.19	0.85	0.40

** Significant at 1% level of significance, NS: Non-Significant

Table 4.34: Tentatively identified GARCH models for monthly rainfall data of SIK subdivision

Tentative models	AIC value
GARCH(1,1)	**10.48**
GARCH(2,1)	10.81
GARCH(1,2)	11.03
GARCH(2,2)	11.03

Table 4.35: Estimates of GARCH (1,1) model parameter for monthly rainfall data of SIK subdivision

Model Parameter	Estimates	S.E	t-statistic	p-value
α_0	158.11**	1.62	97.50	<0.01
α_1	0.00000001^{NS}	0.000003	0.0047	0.94
β_1	0.92**	0.03	3.67	<0.01

** Significant at 1% level of significance, NS: Non-Significant

Table 4.36: RMSE values for ANN model for different activation functions on validation data of monthly rainfall data of SIK subdivision

Input nodes	Hidden Nodes	RMSE values		
		Logistic	Tanh	Linear
12	1	27.47	27.56	2.99
12	2	24.32	20.30	1.45
12	3	15.18	17.88	1.63
12	**4**	18.81	16.16	**0.57**
12	5	10.50	18.04	1.26
12	6	15.56	19.07	3.86

Table 4.37: Forecast accuracy measures for training and testing data sets for best-fitted models of SIK subdivision

	H-WES	SARIMA	ARCH	GARCH	ANN
Training	37.44	36.78	50.72	45.05	**33.22**
Testing	35.11	47.25	67.00	52.24	**33.31**

Table 4.38: Ex-post forecast of monthly rainfall of SIK subdivision by H-WES, SARIMA, ARCH, GARCH and ANN models

Month-year	Actual rainfall (mm)	Forecasted rainfall (mm) for SIK region				
		H-W	SARIMA	ARCH	GARCH	ANN
Jan-19	4.20	0.00	1.13	23.42	13.71	3.42
Feb-19	6.30	0.00	0.49	44.15	10.82	0.67
Mar-19	1.00	6.88	9.9	59.98	21.52	13.33
Apr-19	34.00	38.55	31.81	69.17	42.93	37.12
May-19	85.00	87.71	51.57	72.35	69.28	97.96
Jun-19	67.00	61.37	107.6	71.37	93.49	74.81
Jul-19	55.00	74.13	108.68	68.27	109.06	74.53
Aug-19	146.00	88.43	125.43	64.73	111.81	91.00
Sep-19	143.00	135.45	165.82	61.81	101.02	148.63
Oct-19	237.00	134.24	99.15	59.96	79.60	143.32
Nov-19	33.00	49.64	1.39	59.16	53.31	58.89
Dec-19	17.00	9.09	8.22	59.15	29.23	14.49
RMSE		**35.10**	**47.25**	**67.00**	**52.24**	**33.31**

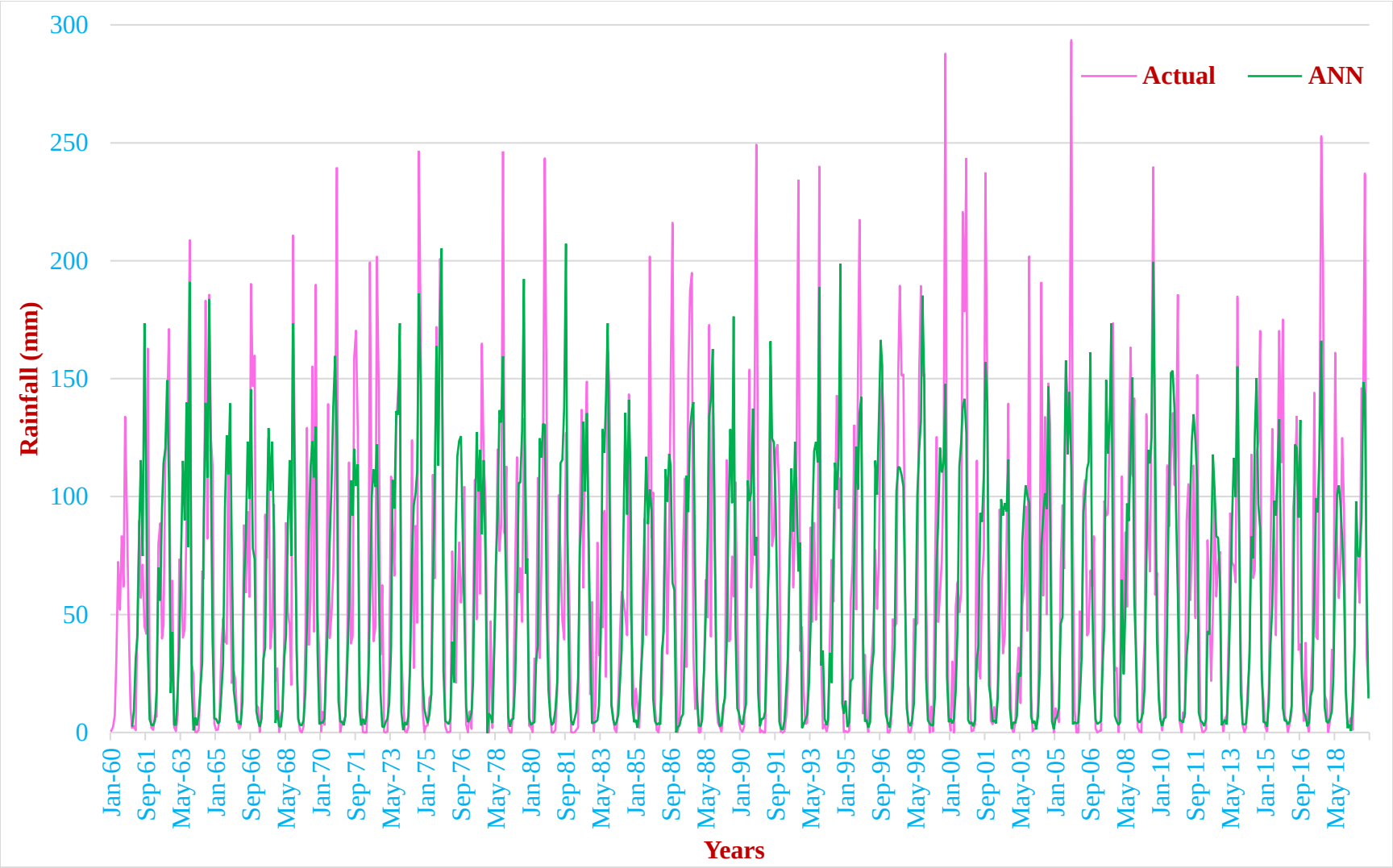

Fig. 4.17: Actual and forecasted monthly rainfall (mm) data using ANN model for SIK subdivision

For SIK subdivision, MLP-NN consist of one input layer with 12 input nodes, one hidden layer having four hidden nodes and one output layer with one output node and 'Linear' activation function was found to be best fitted network for forecasting rainfall with lowest RMSE (0.57) over validation data set. Machado *et al.* (2011) and Roman *et al.* (2012) obtained similar result. Then, training and validation data sets were clubbed together to form training data set, which was used to train the ANN and to forecast for comparing forecasting accuracy with testing data set. RMSE value computed respectively for training and testing data sets are 33.22 and 33.31 as given in Table 4.37.

RMSE values were computed for all the fitted models on both training and testing data sets and are reported in Table 4.37. The forecasting performance of all the fitted models are compared based on RMSE value. For training data set, ANN model with lowest RMSE (33.22) has performed better than the other models, followed by SARIMA, H-WES, GARCH and ARCH models with RMSE values 36.78, 37.44, 45.05 and 50.72 respectively. Further, it can also be observed from the Table 4.37 that for testing data also, the ANN model has performed better with lowest RMSE (33.31) as compared to H-WES (35.11), SARIMA (47.25), GARCH (52.24) and ARCH (67.00) models. It may be note that for both training and testing data, the ANN model performed better than all the other models, which indicates the superiority of ANN model over other models. Therefore, ANN model was selected as a best model for forecasting monthly rainfall of SIK subdivision. The actual rainfall along with predicted by ANN model were plotted in Fig. 4.17. A perusal of Fig. 4.17 indicates that ANN model nicely captures the variations in monthly rainfall of SIK subdivision.

4.3.3 H-WES, SARIMA, ARCH, GARCH and ANN models for forecasting monthly rainfall for Malnad subdivision

H-WES model: Different combination of smoothing parameters α, β and γ were tried by grid search method to get suitable H-WES parameters. The parameter estimates of H-WES model along with standard error for Malnad meteorological subdivision are presented in Table 4.39. The parameter estimates of H-WES model $\alpha = 0.04$, $\beta = 0.00003$ and $\gamma = 0.19$ were found to be suitable for Malnad subdivision with lowest RMSE.

The p-values for α, β and γ are respectively <0.01, 1.00 and <0.01 indicates that smoothing parameters of H-WES model are highly significant except β, which was non-significant. Therefore, this model can be employed for forecasting monthly rainfall for future time period. RMSE value calculated for both training and testing data sets are respectively 87.08 and 162.94 and are tabulated in Table 4.47.

SARIMA model: Distribution pattern of monthly rainfall data along with its ACF and PACF plot for Malnad subdivision are presented in Fig. 4.18. The plots of ACF and PACF at seasonal lags revealed that decay rate was too slow, which indicates that the data was non-stationary at seasonal lags. Hence, seasonal differencing at 12^{th} lag (D = 1) was performed to achieve stationarity. The plots of seasonal differenced monthly rainfall data along with its ACF and PACF plots are presented in the Fig. 4.19. ACF and PACF plots of seasonally differenced series indicates that data was stationary. So, by examining the plots of ACF and PACF, tentative models were identified. The list of tentatively identified models along with their AIC values and Ljung-Box statistic for residuals are tabulated in the Table 4.40. The parameters of model were estimated using maximum likelihood estimator. To check the adequacy of the model, residual analysis was carried out using Ljung-Box test statistic. Based on the lowest AIC value (8214.38), SARIMA $(2,0,0)(0,1,1)_{12}$was found to be the best-fitted model among the tentative models. Estimated model parameters with their respective standard errors, t-statistic and p-values are exhibited in the Table 4.41.

The ACF and PACF plots of residuals obtained from best-fitted model are displayed in the Fig. 4.20. The Ljung-Box test statistic (0.01) depicted in Table 4.40 for best-fitted model was found to be non-significant at 5 per cent level of significance (p-value 0.91) indicates that errors are independent and the model is adequate. RMSE value computed for training and testing data sets were respectively 250.49 and 161.15 and are presented in Table 4.47.

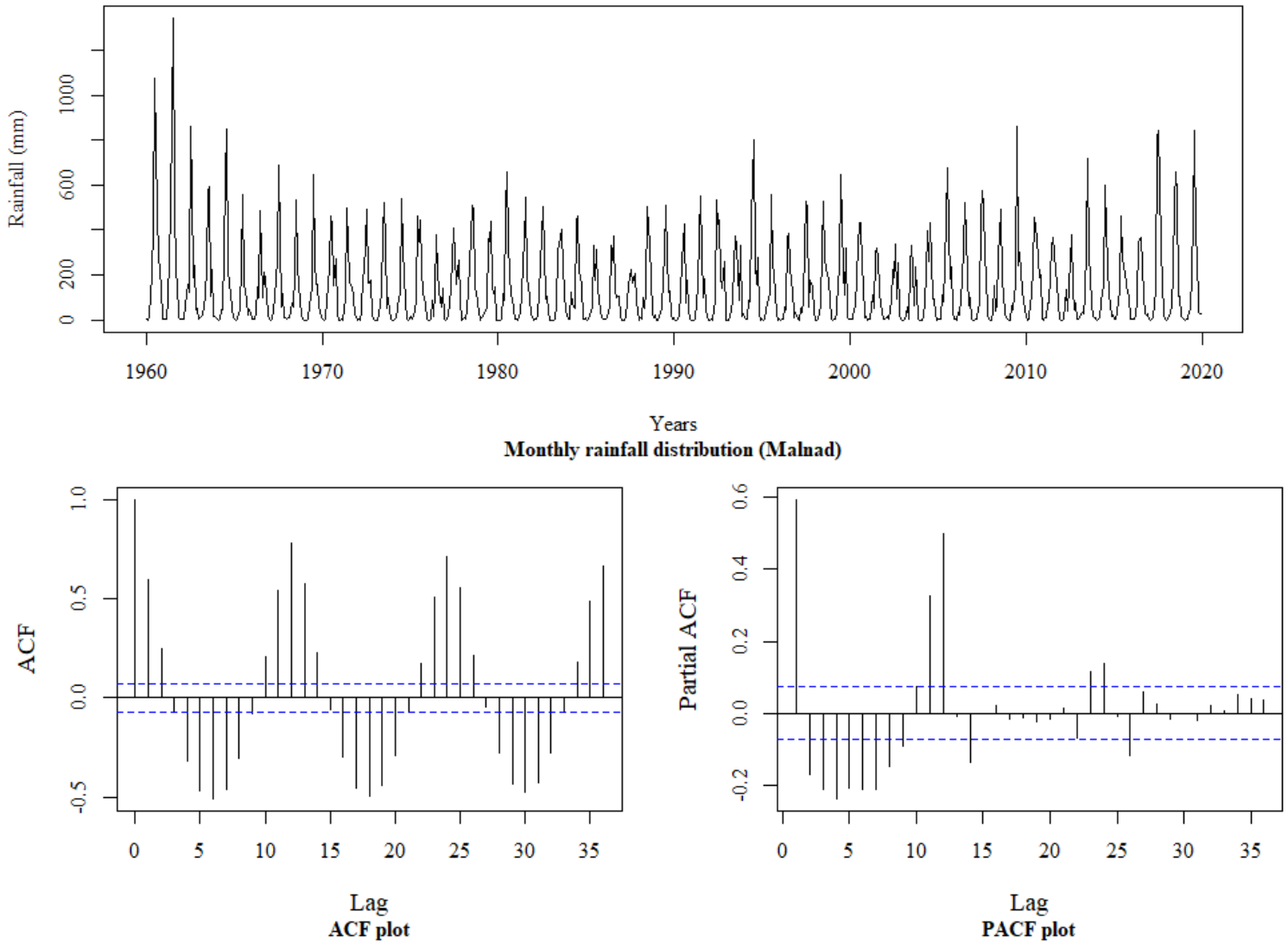

Fig. 4.18: Line graph for monthly rainfall data along with its ACF and PACF plots for Malnad subdivision

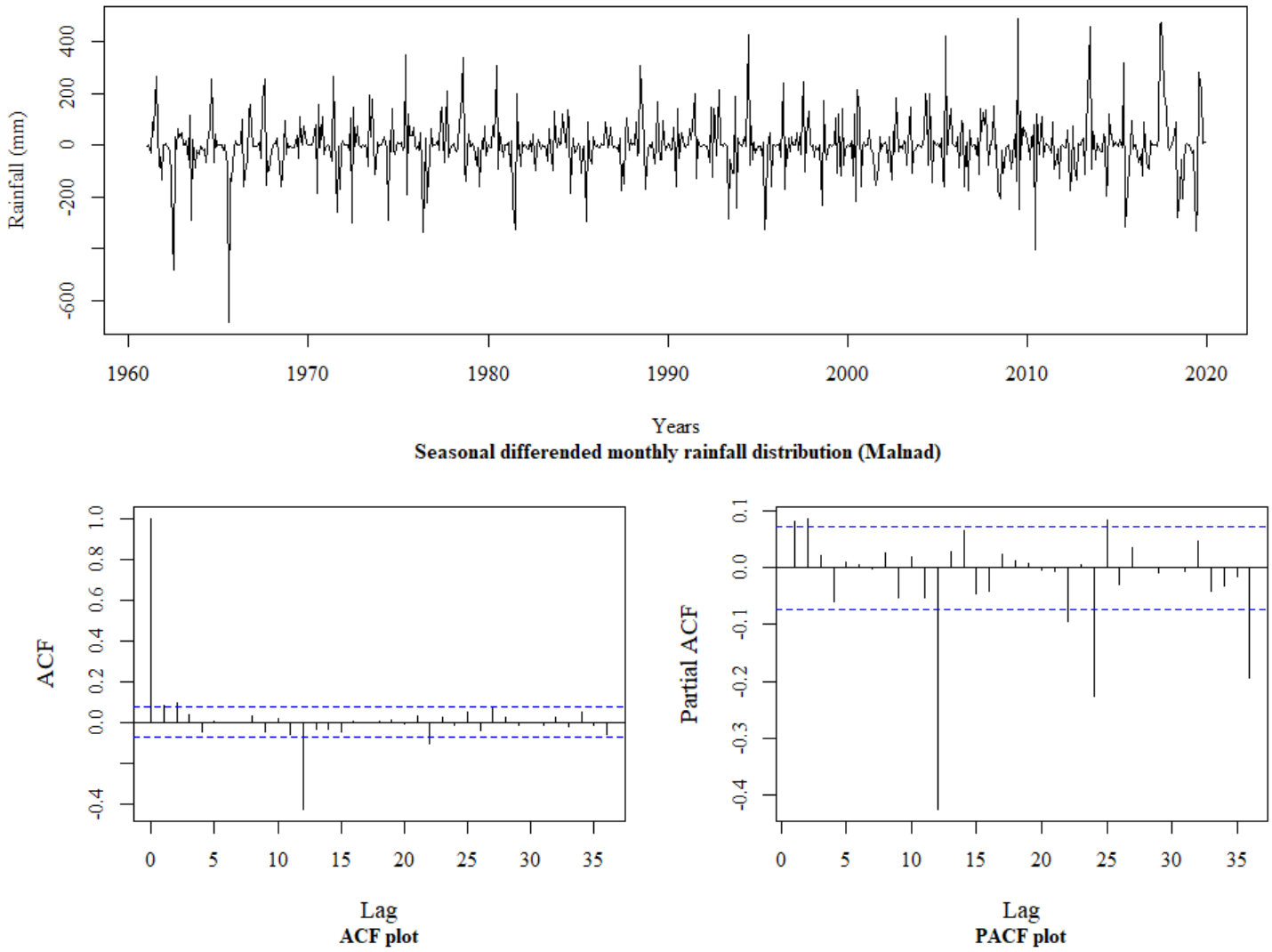

Fig. 4.19: Line graph for seasonally differenced monthly rainfall data along with its ACF and PACF plots for Malnad subdivision

Table 4.39: Parameter estimates of H-WES model for monthly rainfall data of Malnad subdivision

Model Parameter	Estimates	S.E	t-statistic	p-value
α	0.04**	0.01	3.34	<0.01
β	0.00003^{NS}	0.001	0.03	0.98
γ	0.19**	0.02	9.13	<0.01

** Significant at 1% level of significance, NS: Non-Significant

Table 4.40: Tentatively identified SARIMA models for monthly rainfall data of Malnad subdivision

Tentative models	AIC value	Ljung-Box statistic
$(0,0,0)\ (2,1,3)_{12}$	8225.95	11.29**
$(0,0,0)\ (2,1,2)_{12}$	8232.74	11.13**
$(0,0,0)\ (1,1,2)_{12}$	8214.43	0.65^{NS}
$(0,0,0)\ (2,1,0)_{12}$	8294.49	2.50^{NS}
$(1,0,0)\ (2,1,0)_{12}$	8294.01	0.04^{NS}
$(0,0,1)\ (2,1,0)_{12}$	8294.50	0.02^{NS}
$(2,0,0)\ (0,1,1)_{12}$	**8214.38**	**0.01^{NS}**
$(0,0,0)\ (0,1,2)_{12}$	8235.87	11.64**
$(1,0,0)\ (0,1,2)_{12}$	8225.54	0.33^{NS}
$(1,1,1)\ (2,1,0)_{12}$	8290.33	0.23^{NS}
$(2,1,0)\ (0,1,1)_{12}$	8349.12	1.10^{NS}
$(0,1,2)\ (0,1,1)_{12}$	8218.73	0.03^{NS}

Table 4.41: Parameter estimates SARIMA $(2,0,0)\ (0,1,,1)_{12}$ model for monthly rainfall data of Malnad subdivision

Model Parameter	Estimates	S.E	t-statistic	p-value
AR(1)	0.12**	0.04	3.03	<0.01
AR(2)	0.15**	0.04	3.98	<0.01
SMA(1)	-0.77**	0.03	23.63	<0.01

** Significant at 1% level of significance

ARCH model: The ARCH-LM test was carried out to check the volatility or ARCH effect in the residuals of SARIMA model. ARCH-LM test statistic value of 117.08 with p-value < 0.01 reveals that, there is highly significant ARCH effect in the best-fitted SARIMA model for Malnad subdivision. Therefore, ARCH and GARCH models were built for Malnad subdivision.

Tentative identified models along with their AIC values are tabulated in the Table 4.42. Initially, four tentative ARCH models were fitted and the model with AIC values (12.54), ARCH (1) is selected as best fitted model among the tentative models. Parameters are estimated using maximum likelihood method. Estimated model parameters along with their respective standard errors, t-statistic and p-values are depicted in the Table 4.43. The p-values for estimated parameters α_0(<0.01) and α_1(<0.01) are indicates high significance of parameters, which indicates selected model was adequate model. RMSE value calculated are respectively 145.15 and 220.19 for trained and testing data sets, which are tabulated in Table 4.47.

GARCH model: Tentative identified models along with their AIC values are tabulated in the Table 4.44. Initially, four tentative GARCH models were fitted and the model with AIC values (12.60), GARCH (1,1) is selected as best fitted model among the tentative models. Parameters are estimated using maximum likelihood method. Estimated model parameters along with their respective standard errors, t-statistic and p-values are organized in the Table 4.45. Estimated parameters α_0, α_1 and β_1 were found to be significant at 1 per cent level of significance indicates that selected model was not adequate to explain heteroscedasticity in the residual of SARIMA model. RMSE value computed respectively for training and testing data sets are 132.59 and 254.96 as depicted in Table 4.47.

Artificial Neural Networks (ANN): The most commonly used neural network for time-series forecasting, MLP feed forward neural network, was adopted in this study. Training data set was used to train the MLP-NN model, validation data set was used for validate model parameters (*i.e.,* find optimal number of hidden nodes and activation function), and finally testing data set was used to measure the accuracy of forecasting.

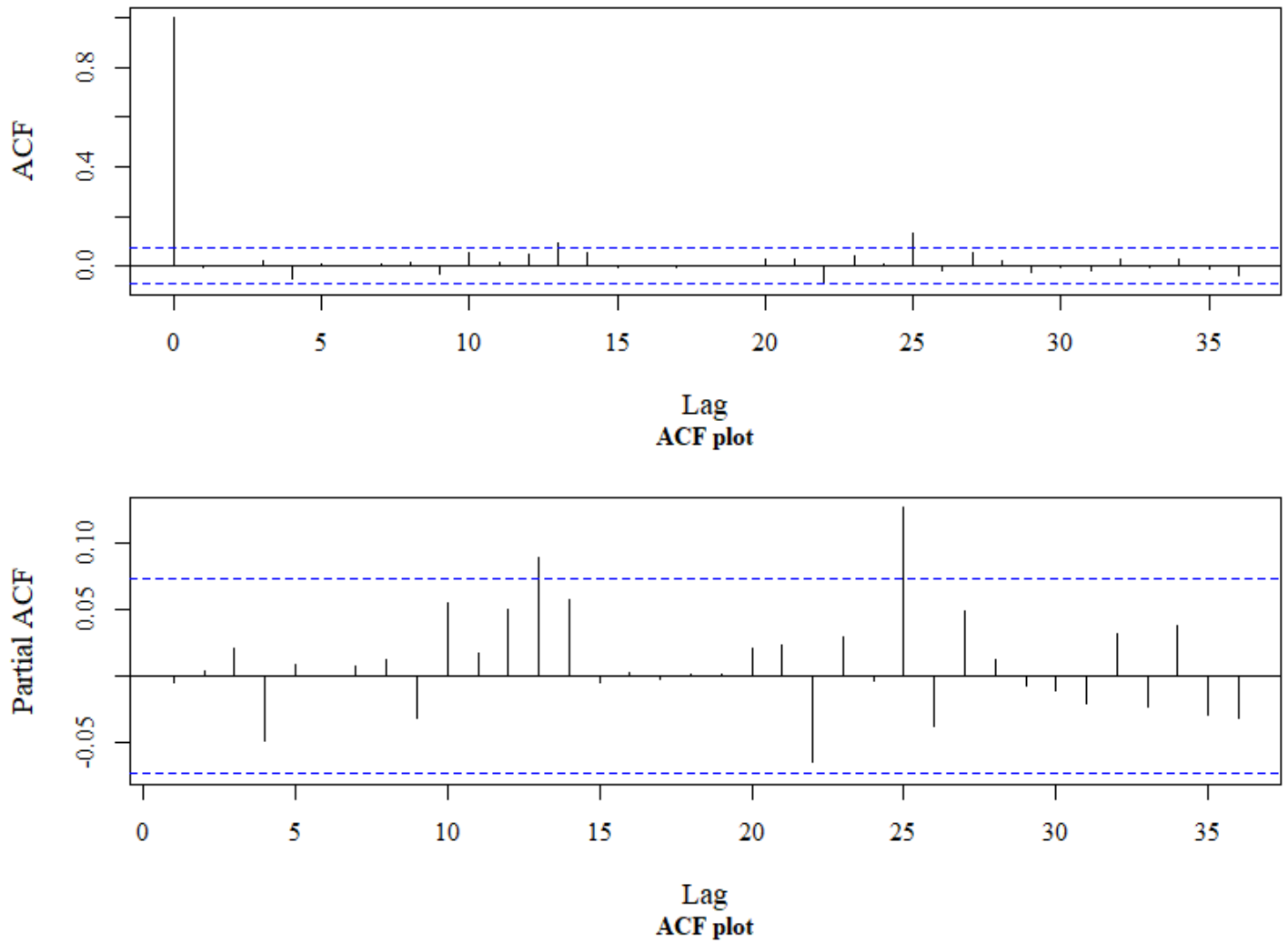

Fig. 4.20: ACF and PACF plots of residuals of SARIMA (2,0,0) $(0,1,1)_{12}$ model for monthly rainfall data for Malnad subdivision

Table 4.42: Tentatively identified ARCH models for monthly rainfall data of Malnad subdivision

Tentative models	AIC value
ARCH(1)	**12.54**
ARCH(2)	12.74
ARCH(3)	12.74
ARCH(4)	12.71

Table 4.43: Parameter estimates ARCH (1) model for monthly rainfall data of Malnad subdivision

Model Parameter	Estimates	S.E	t-statistic	p-value
α_0	11504.14**	3.89	2961.10	<0.01
α_1	0.36**	0.10	3.50	<0.01

** Significant at 1% level of significance

Table 4.44: Tentatively identified GARCH models for monthly rainfall data of Malnad subdivision

Tentative models	AIC value
GARCH(1,1)	**12.60**
GARCH(2,1)	12.77
GARCH(1,2)	12.76
GARCH(2,2)	12.77

Table 4.45: Estimates of GARCH (1,1) model parameter for monthly rainfall data of Malnad subdivision

Model Parameter	Estimates	S.E	t-statistic	p-value
α_0	300.80**	3.57	84.20	<0.01
α_1	0.02**	0.004	4.04	<0.01
β_1	0.97**	0.01	122.73	<0.01

** Significant at 1% level of significance

ANN model with 12 input nodes, one hidden layer and one output node were fixed prior to fitting ANN model. Three different activation function namely Logistic, Hyperbolic tangent (Tanh) and Linear, and various number of hidden nodes (h = 1, 2, …) were trained one at a time to find the optimal combination of these parameters by computing RMSE value on validation set which are tabulated in Table 4.46. Model with lowest RMSE for validation set was considered as best fitted model.

For Malnad subdivision, MLP-NN consist of one input layer with 12 input nodes, one hidden layer having five hidden nodes and one output layer with one output node and 'Logistic' activation function was found to be best fitted network for forecasting rainfall with lowest RMSE (85.03)over validation data set. Machado *et al.* (2011) and Roman *et al.* (2012) obtained similar result. Then, training and validation data sets were clubbed together to form training data set, which was used to train the MLP-NN and to forecast for comparing forecasting accuracy with testing data set. RMSE value computed respectively for training and testing data sets are 62.52 and 153.89 as depicted in Table 4.47**.**

RMSE values were computed for all the fitted models on both training and testing data sets and are reported in Table 4.47. The forecasting performance of all the fitted models are compared based on RMSE value. For training data set, ANN model with lowest RMSE (62.52) has performed better than the other models, followed by H-WES, GARCH, ARCH and SARIMA models with RMSE values 87.08, 132.59, 145.15 and 250.49 respectively. Further, it can also be observed from the Table 4.47 that for testing data also, the ANN model has performed better with lowest RMSE (153.89) as compared to H-WES (162.94), SARIMA (165.15), ARCH (220.19) and GARCH (254.96) models. It may be note that for both training and testing data, the ANN model performed better than all the other models, which indicates the superiority of ANN model over other models. Therefore, ANN model was selected as a best model for forecasting monthly rainfall of Malnad subdivision. The actual rainfall along with predicted by ANN model were plotted in Fig. 4.21. A perusal of Fig. 4.21 indicates that ANN model nicely captures the variations in monthly rainfall of Malnad subdivision.

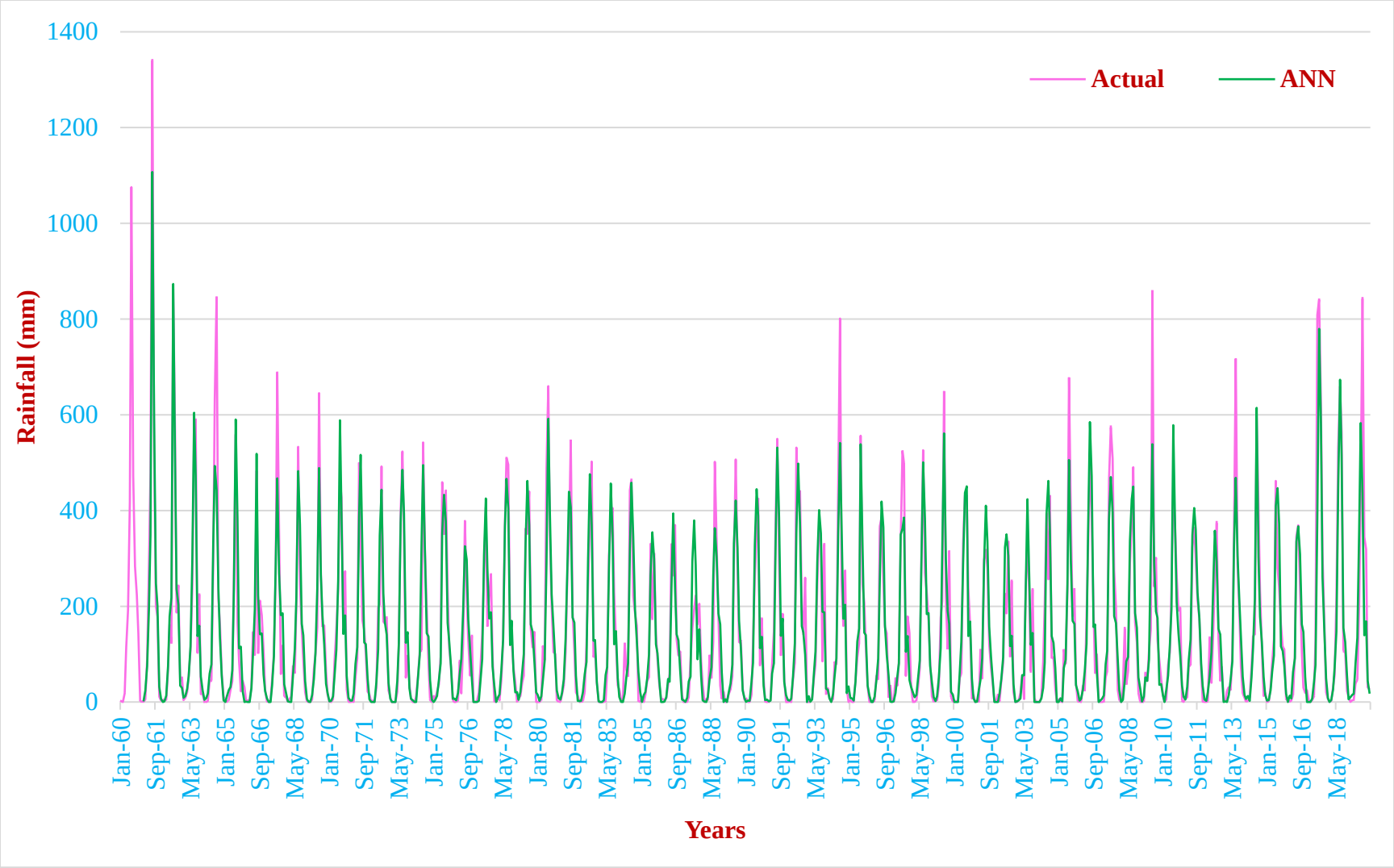

Fig. 4.21: Actual and forecasted monthly rainfall (mm) data using ANN model for Malnad subdivision

Table 4.46: RMSE values for ANN model for different activation functions on validation data of monthly rainfall data of Malnad subdivision

Input nodes	Hidden Nodes	RMSE values		
		Logistic	Tanh	Linear
12	1	118.44	120.08	123.44
12	2	110.68	110.59	112.31
12	3	101.01	97.84	108.41
12	4	104.24	101.32	98.06
12	**5**	**85.03**	94.53	87.39
12	6	114.30	99.48	104.48
12	7	124.76	104.95	118.14

Table 4.47: Forecast accuracy measures for training and testing data sets for best-fitted models of Malnad subdivision

	H-WES	SARIMA	ARCH	GARCH	ANN
Training	87.08	250.49	145.15	132.59	**62.52**
Testing	162.94	165.15	220.19	254.96	**153.89**

Table 4.48: Ex-post forecast of monthly rainfall of Malnad subdivision by H-WES, SARIMA, ARCH, GARCH and ANN models

Month-year	Actual rainfall (mm)	Forecasted rainfall (mm) for Malnad region				
		H-WES	SARIMA	ARCH	GARCH	ANN
Jan-19	0.30	30.59	2.21	0.00	13.71	10.08
Feb-19	3.60	33.12	3.11	7.14	10.82	14.68
Mar-19	4.00	49.22	18.81	53.51	21.52	17.03
Apr-19	39.00	80.82	49.19	108.32	42.93	80.65
May-19	46.00	157.31	133.74	156.07	69.28	133.84
Jun-19	203.00	448.46	444.06	187.08	93.49	167.36
Jul-19	446.00	573.67	562.83	198.31	109.06	582.71
Aug-19	844.00	431.54	421.68	192.31	111.81	435.89
Sep-19	345.00	204.87	182.78	175.13	101.02	334.40
Oct-19	319.00	137.53	112.88	153.90	79.60	270.35
Nov-19	31.00	67.02	41.72	134.83	53.31	44.32
Dec-19	26.00	33.94	11.06	121.97	29.23	19.81
RMSE		**162.94**	**165.15**	**220.19**	**254.96**	**128.79**

4.3.4 H-WES, SARIMA, ARCH, GARCH and ANN models for forecasting monthly rainfall for Coastal subdivision

H-WES model: Different combination of smoothing parameters α, β and γ were tried by grid search method to get suitable H-WES parameters. The parameter estimates of H-WES model along with standard error for Coastal meteorological subdivision are given in Table 4.49. The parameter estimates of H-WES model $\alpha = 0.01$, $\beta = 0.00002$ and $\gamma = 0.002$ were found to be suitable for Coastal subdivision with lowest RMSE. The p-values for α, β and γ are respectively 0.07, 0.92 and 0.85 indicates that smoothing parameters of H-WES model are non-significant. Therefore, this model cannot be employed for forecasting monthly rainfall for future time period. RMSE value calculated for both training and testing data sets are respectively 129.35 and 263.95 and are tabulated in Table 4.57.

SARIMA model: Distribution pattern of monthly rainfall data along with its ACF and PACF plot for Coastal subdivision are depicted in Fig. 4.22. The plots of ACF and PACF at seasonal lags revealed that decay rate was too slow, which indicates that the data was non-stationary at seasonal lags. Hence, seasonal differencing at 12^{th} lag (D = 1) was performed to achieve stationarity. The plots of seasonal differenced monthly rainfall data along with its ACF and PACF plots are presented in the Fig. 4.23. ACF and PACF plots of seasonally differenced series indicates that data was stationary. So, by examining the plots of ACF and PACF, tentative models were identified. The list of tentatively identified models along with their AIC values and Ljung-Box statistic for residuals are presented in the Table 4.50. The parameters of model were estimated using maximum likelihood estimator. To check the adequacy of the model, residual analysis was carried out using Ljung-Box test statistic. Based on the lowest AIC value (8783.09), SARIMA (0,0,0) $(0,1,2)_{12}$was found to be the best-fitted model among the tentative models. Estimated model parameters with their respective standard errors, t-statistic and p-values are exhibited in the Table 4.51.

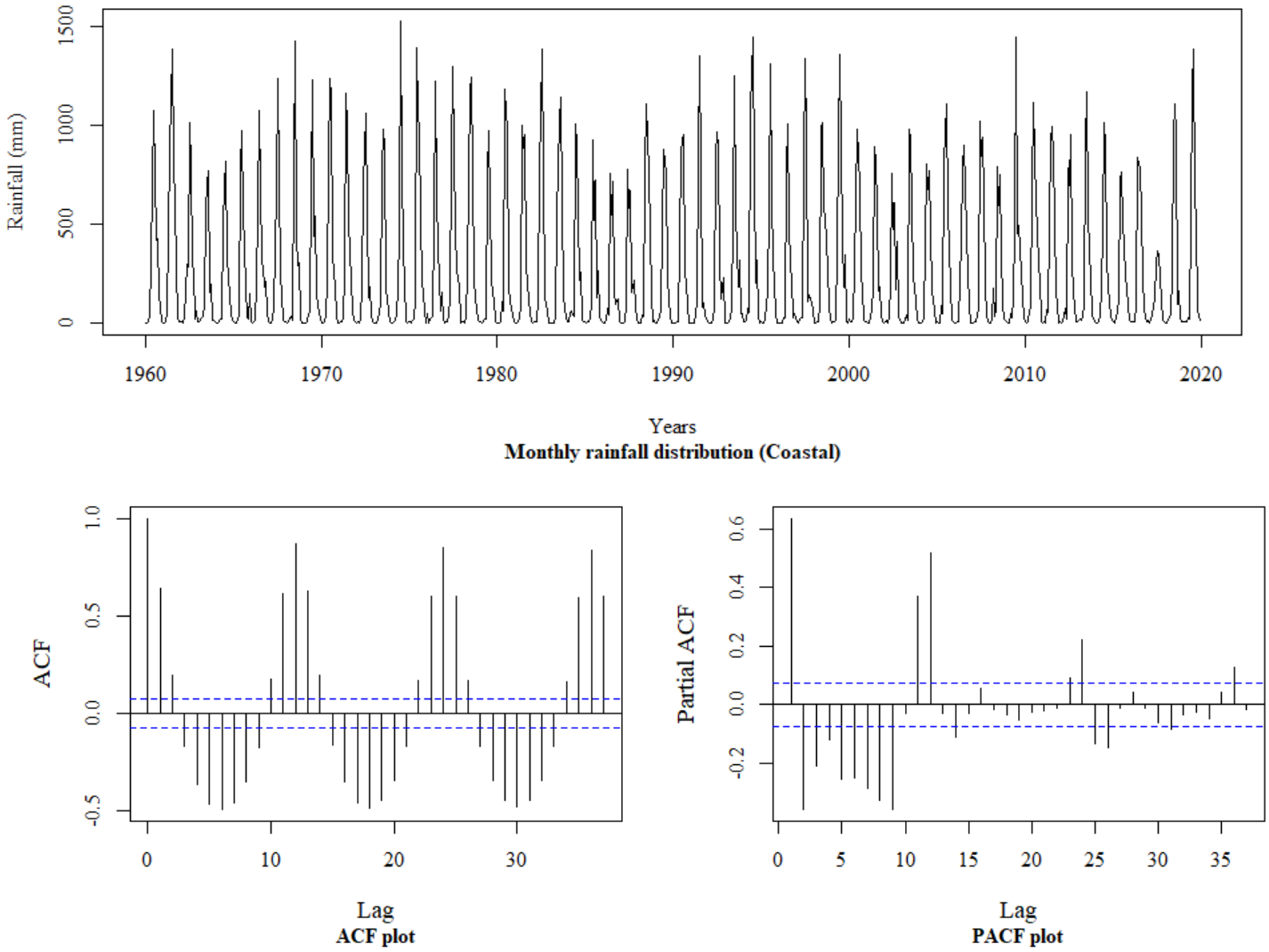

Fig. 4.22: Line graph for monthly rainfall data along with its ACF and PACF plots for Coastal subdivision

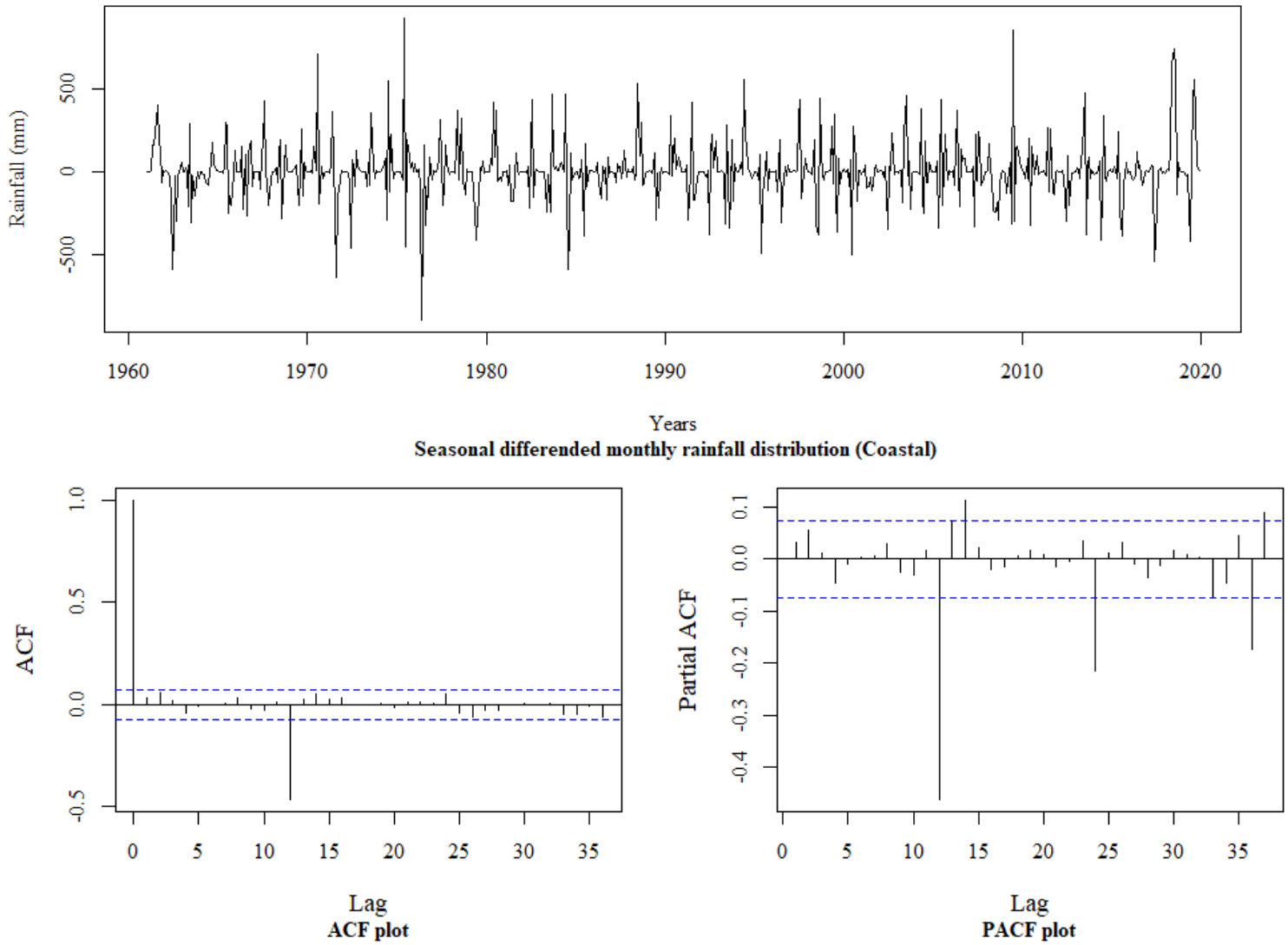

Fig. 4.23: Line graph for seasonally differenced monthly rainfall data along with its ACF and PACF plots for Coastal subdivision

Table 4.49: Parameter estimates of H-WES model for monthly rainfall data of Coastal subdivision

Model Parameter	Estimates	S.E	t-statistic	p-value
α	0.01^{NS}	0.003	1.85	0.07
β	0.00002^{NS}	0.0002	0.10	0.92
γ	0.002^{NS}	0.01	0.20	0.85

NS: Non-Significant

Table 4.50: Tentatively identified SARIMA models for monthly rainfall data of Coastal subdivision

Tentative models	AIC value	Ljung-Box statistic
$(0,0,0)\ (2,1,3)_{12}$	8783.87	0.31^{NS}
$(0,0,0)\ (2,1,2)_{12}$	8783.20	0.68^{NS}
$(0,0,0)\ (2,1,0)_{12}$	8887.85	0.62^{NS}
$(0,0,0)\ (0,1,2)_{12}$	**8783.09**	**0.21^{NS}**
$(0,0,1)\ (2,1,0)_{12}$	8889.65	0.00^{NS}
$(2,0,0)\ (0,1,1)_{12}$	8793.09	0.21^{NS}
$(1,0,0)\ (0,1,2)_{12}$	8794.87	0.00^{NS}
$(1,0,0\ (1,1,0)_{12}$	8931.48	0.00^{NS}
$(0,0,1)\ (0,1,2)_{12}$	8794.90	0.00^{NS}

Table 4.51: Parameter estimates SARIMA $(0,0,0)\ (0,1,2)_{12}$ model of monthly rainfall data for Coastal subdivision

Model Parameter	Estimates	S.E	t-statistic	p-value
SMA(1)	-0.84**	0.04	22.96	<0.01
SMA(2)	-0.09*	0.04	2.49	0.01

** Significant at 1% level of significance

Table 4.52: Tentatively identified ARCH models for monthly rainfall data of Coastal subdivision

Tentative models	AIC value
ARCH(1)	**14.04**
ARCH(2)	14.13
ARCH(3)	14.54
ARCH(4)	14.14

The ACF and PACF plots of residuals obtained from best-fitted model are displayed in the Fig. 4.24. The Ljung-Box test statistic (0.21) presented in Table 4.50 for best-fitted model was found to be non-significant at 5 per cent level of significance (p-value 0.65) indicates that errors are independent and the model is adequate. RMSE value computed for training and testing data sets were respectively 129.99 and 261.90 and are depicted in Table 4.57.

ARCH model: The ARCH-LM test was carried out to check the volatility or ARCH effect in the residuals of SARIMA model. ARCH-LM test statistic value of 81.97 with p-value < 0.01 reveals that, there is highly significant ARCH effect in the best-fitted SARIMA model for Coastal subdivision. Therefore, ARCH and GARCH models were built for Coastal subdivision.

Tentative identified models along with their AIC values are illustrated in the Table 4.52. Initially, four tentative ARCH models were fitted and the model with AIC values (14.04), ARCH (1) is selected as best fitted model among the tentative models. Parameters are estimated using maximum likelihood method. Estimated model parameters along with their respective standard errors, t-statistic and p-values are tabulated in the Table 4.53. The p-values for estimated parameters α_0(<0.01) and α_1(<0.01) are indicates high significance of parameters, which indicates selected model was adequate model. RMSE value calculated are respectively 273.28 and 440.35 for trained and testing data sets, which are tabulated in Table 4.57.

GARCH model: Tentative identified models along with their AIC values are tabulated in the Table 4.54. Initially, four tentative GARCH models were fitted and the model with AIC values (14.07), GARCH (1,1) is selected as best fitted model among the tentative models. Parameters are estimated using maximum likelihood method. Estimated model parameters along with their respective standard errors, t-statistic and p-values are tabulated in the Table 4.55. Estimated parameters α_0 and β_1 were found to be significant at 1 per cent level of significance and α_1 was found to be non-significant at 5 per cent level of significance indicates that selected model was not adequate to explain heteroscedasticity in the residual of SARIMA model. RMSE value computed respectively for training and testing data sets are 272.75 and 433.17 as reported in Table 4.57.

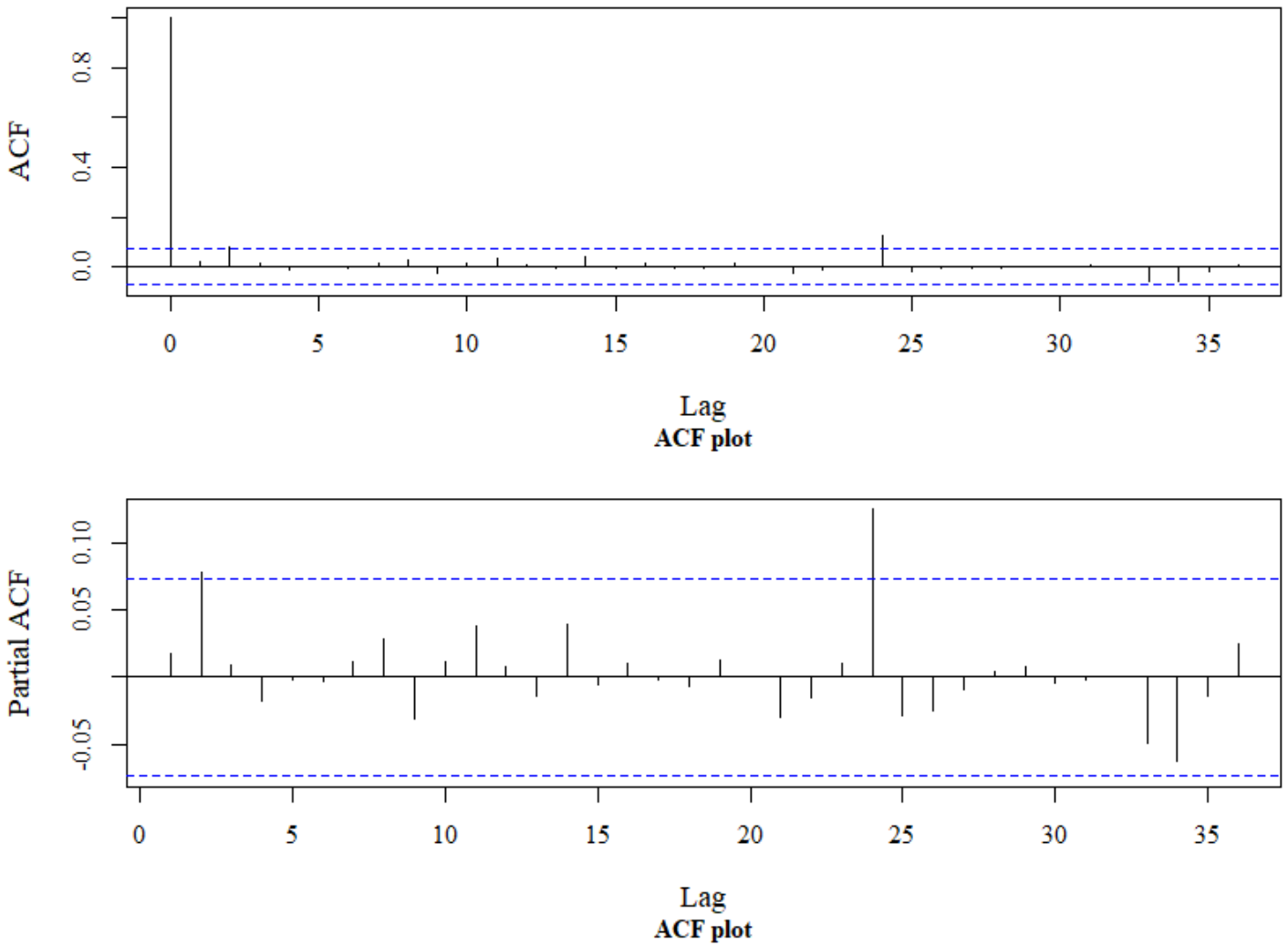

Fig. 4.24: ACF and PACF plots of residuals of SARIMA (0,0,0) $(0,1,2)_{12}$ model for monthly rainfall data for Coastal subdivision

Table 4.53: Parameter estimates ARCH (1) model for monthly rainfall data of Coastal subdivision

Model Parameter	Estimates	S.E	t-statistic	p-value
α_0	62945.76**	26.99	2331.95	<0.01
α_1	0.16**	0.01	3.33	<0.01

** Significant at 1% level of significance

Table 4.54: Tentatively identified GARCH models for monthly rainfall data of Coastal subdivision

Tentative models	AIC value
GARCH(1,1)	**14.07**
GARCH(2,1)	14.12
GARCH(1,2)	14.12
GARCH(2,2)	14.18

Table 4.55: Estimates of GARCH (1,1) model parameter for monthly rainfall data of Coastal subdivision

Model Parameter	Estimates	S.E	t-statistic	p-value
α_0	26887.74**	25.56	1051.84	<0.01
α_1	0.00000008^{NS}	0.00002	0.004	0.99
β_1	0.64**	0.07	9.30	<0.01

** Significant at 1% level of significance, NS: Not Significant

Table 4.56: RMSE values for ANN model for different activation functions on validation data of monthly rainfall data of Coastal subdivision

Input nodes	Hidden Nodes	RMSE values		
		Logistic	Tanh	Linear
12	1	6.72	6.00	36.32
12	2	21.45	25.88	38.84
12	3	10.22	13.44	36.47
12	**4**	3.87	**1.39**	58.67
12	5	4.72	14.43	66.03
12	6	23.05	26.17	55.33

Artificial Neural Networks (ANN): The most commonly used neural network for time-series forecasting, MLP feed forward neural network, was adopted in this study. Training data set was used to train the MLP-NN model, validation data set was used for validate model parameters (*i.e.,* find optimal number of hidden nodes and activation function), and finally testing data set was used to measure the accuracy of forecasting. ANN model with 12 input nodes, one hidden layer and one output node were fixed prior to fitting ANN model. Three different activation function namely Logistic, Hyperbolic tangent (Tanh) and Linear, and various number of hidden nodes (h = 1, 2, …) were trained one at a time to find the optimal combination of these parameters by computing RMSE value on validation set which are given in Table 4.56. Model with lowest RMSE for validation set was considered as best fitted model.

For Coastal subdivision, MLP-NN consist of one input layer with 12 input nodes, one hidden layer having four hidden nodes and one output layer with one output node and ‘Tanh activation function was found to be best fitted network for forecasting rainfall with lowest RMSE (1.39) over validation data set. Then, training and validation data sets were clubbed together to form training data set, which was used to train the MLP-NN and to forecast for comparing forecasting accuracy with testing data set. RMSE value computed respectively for training and testing data sets are 98.80 and 238.67 as tabulated in Table 4.57**.**

RMSE values were computed for all the fitted models on both training and testing data sets and are reported in Table 4.57. The forecasting performance of all the fitted models are compared based on RMSE value. For training data set, ANN model with lowest RMSE (98.80) has performed better than the other models, followed by H-WES, SARIMA, GARCH and ARCH models with RMSE values 129.35, 129.99, 272.75 and 273.28 respectively. Further, it can also be observed from the Table 4.57 that for testing data also, the ANN model has performed better with lowest RMSE (238.67) as compared to SARIMA (261.90), H-WES (263.95), GARCH (433.17) and ARCH (440.35) models. It may be note that for both training and testing data, the ANN model performed better than all the other models, which indicates the superiority of ANN model over other models.

Table 4.57: Forecast accuracy measures for training and testing data sets for best-fitted models of Coastal subdivision

	H-WES	SARIMA	ARCH	GARCH	ANN
Training	129.35	129.99	273.28	272.75	**98.80**
Testing	263.95	261.90	440.35	433.17	**238.67**

Table 4.58: Ex-post forecast of monthly rainfall of Coastal subdivision by H-WES, SARIMA, ARCH, GARCH and ANN models

Month-Year	Actual rainfall (mm)	Forecasted rainfall (mm) for Coastal region				
		H-WES	SARIMA	ARCH	GARCH	ANN
Jan-19	0.70	0.00	2.30	91.50	108.99	4.27
Feb-19	0.60	0.00	1.52	95.77	113.22	3.28
Mar-19	4.00	12.62	14.70	87.99	105.31	19.88
Apr-19	23.00	18.37	35.37	99.15	116.75	45.11
May-19	17.00	117.81	143.19	116.48	133.97	104.25
Jun-19	525.00	777.72	801.43	175.67	193.50	745.14
Jul-19	1185.00	994.66	967.93	297.98	315.94	960.56
Aug-19	1384.00	696.61	723.32	311.55	327.56	869.76
Sep-19	646.00	278.09	276.00	290.35	306.54	229.43
Oct-19	512.00	168.98	180.35	213.08	228.80	140.58
Nov-19	56.00	47.69	55.26	134.05	150.42	36.13
Dec-19	12.00	17.89	9.08	89.32	106.36	15.02
RMSE		**263.94**	**261.90**	**440.35**	**433.17**	**238.67**

Therefore, ANN model was selected as a best model for forecasting monthly rainfall of Coastal subdivision. Somvanshi *et. al.* (2006) also reported that ANN model outperformed the ARIMA model for forecasting the rainfall data. The actual rainfall along with predicted by ANN model were plotted in Fig. 4.25. A perusal of Fig. 4.25 indicates that ANN model nicely captures the variations in monthly rainfall of Coastal subdivision.

Since ANN model performed better than all the other models for all the four meteorological subdivision of Karnataka, it was further used to forecast monthly rainfall of all the four subdivision. Monthly rainfall data of each subdivisions for the period of 60 year from 1960 to 2019 were used to build the ANN model, then this trained model intern was used for forecasting next year rainfall data *i.e.*, for 2020. The forecasted monthly rainfall data along with normal rainfall as per IMD for each subdivision for the year 2020 are reported in Table 4.59.

Results in Table 4.59 revealed that for NIK subdivision, estimated rainfall (mm) for the month of January (-1.25), February (-0.38), March (-1.43), May (-0.94), June (-23.98), July (107.02) and October (-6.66) will be less than normal rainfall, whereas for April (5.67), August (6.22), September (11.31), November (0.20) and December (1.16) months will be more than normal rainfall. For SIK subdivision, expected rainfall for the month of January (-1.69) and December (-11.04) will be less than normal rainfall, and for the month of February (2.10), March (14.52), April (6.53), May (22.03), June (1.44), July (2.35), August (19.63), September (1.32), October (6.37) and November (2.31) will be more than normal rainfall. In Malnad subdivision, the forecasted rainfall for the month of May (-74.58), June (-81.42), July (-34.80), November (-19.53) and December (-6.69) will be less than normal rainfall and for the month of January (8.23), February (10.09), March (47.83), April (1.95), August (63.74), September (29.44) and October (12.33) will be more than normal rainfall. In addition, for Coastal subdivision, in the month of April (-7.37), May (-16.28), July (-102.32), November (-0.43) and December (-6.40) less rainfall, whereas in the month of January (2.85), February (6.58), March (5.32), June (26.59), August (333.46), September (109.17) and October (106.03) more rainfall can be expected than the normal rainfall.

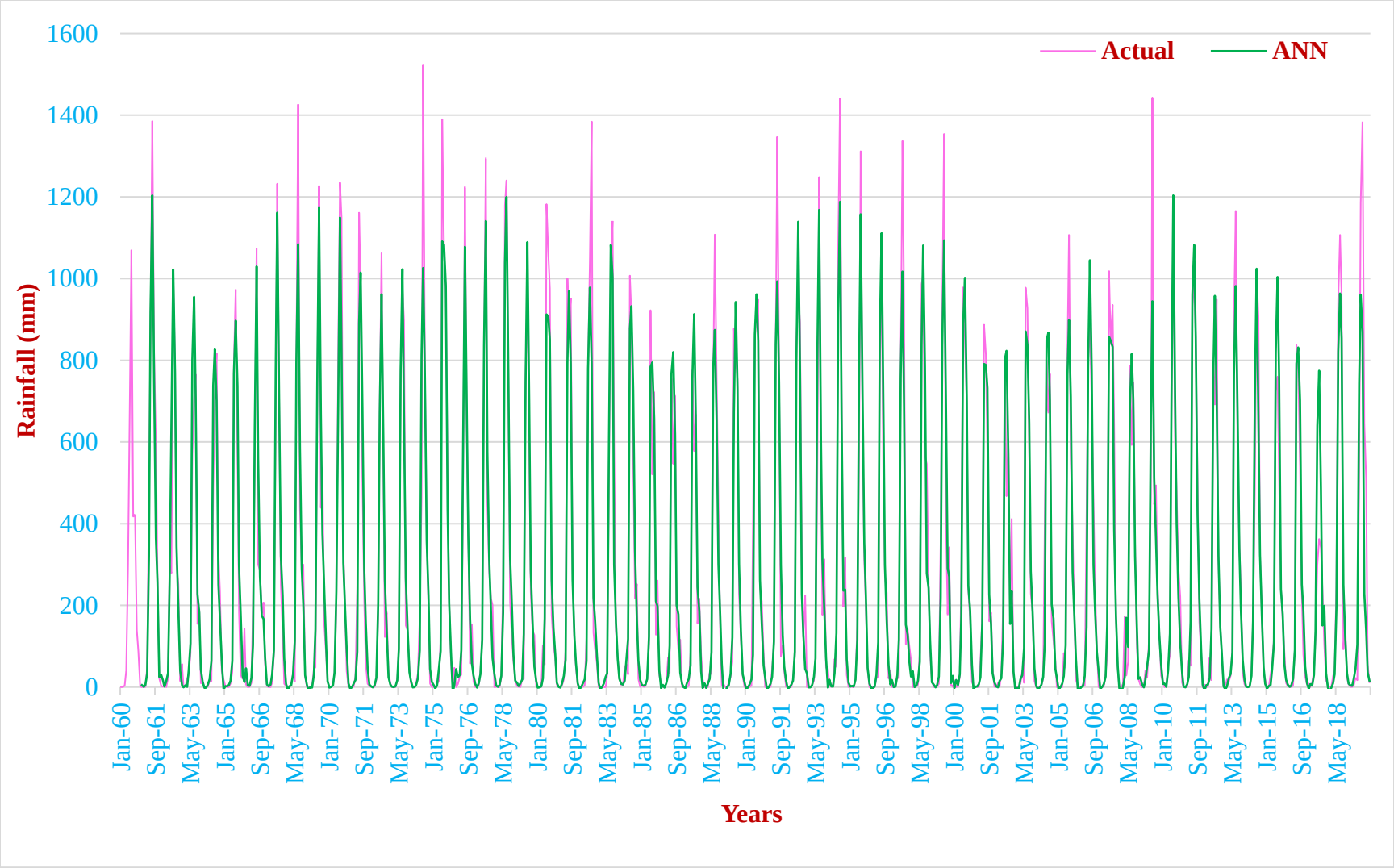

Fig.4.25: Actual and forecasted monthly rainfall (mm) data using ANN model for Coastal subdivision

Table 4.59: One-step-ahead forecasts of monthly rainfall (mm) by ANN model along with normal rainfall (mm) for all the meteorological subdivision of Karnataka

Month-Year	NIK		SIK		Malnad		Coastal	
	Normal	Forecasted	Normal	Forecasted	Normal	Forecasted	Normal	Forecasted
Jan-20	2.10	0.85	1.80	0.01	2.00	10.23	1.00	3.85
Feb-20	1.80	1.42	3.50	5.60	2.40	12.49	0.40	6.98
Mar-20	5.00	3.57	8.00	22.52	11.00	58.83	5.00	10.32
Apr-20	25.00	30.67	41.00	47.53	56.00	57.95	31.00	23.63
May-20	55.00	54.06	96.00	118.03	112.00	37.42	135.00	118.72
Jun-20	101.00	91.28	64.00	65.44	358.00	276.58	813.00	839.59
Jul-20	131.00	107.02	79.00	81.35	598.00	563.20	1156.00	1053.68
Aug-20	118.00	124.22	81.00	100.63	382.00	445.74	760.00	1093.46
Sep-20	145.00	156.31	135.00	136.32	165.00	194.44	290.00	399.17
Oct-20	111.00	104.34	146.00	152.37	161.00	173.33	187.00	293.03
Nov-20	28.00	28.20	50.00	52.31	55.00	35.47	61.00	60.57
Dec-20	7.00	8.16	14.00	2.96	13.00	6.31	13.00	6.60

V SUMMARY AND CONCLUSION

Agriculture is backbone of India's economy and is largely dependent on the monsoon. Rainfall plays a dominant role in agricultural production and productivity, despite advancement in many technological fronts. Therefore, an adequate understanding of the spatial and temporal dynamics of rainfall is important for the planning and management of agriculture related activities. However, such is also often difficult, because rainfall occurs as a result of complex and nonlinear interactions among numerous atmospheric, land, and ocean processes.

Karnataka state is facing frequent drought conditions due to uncertainty of rainfall distribution and spatio-temporal variation in different regions. Due to these variations in rainfall distribution, the state is experiencing extreme weather conditions like drought or flood in different regions. Frequent drought and floods have caused heavy loss to crops and lives, while society loses its livelihood and rhythm of life and people are forced to migrate for their sustenance. Rainfall also influences the quantity of water available in the rivers, which in turn influences the amount of drinking water available to the population and the amount of electricity that can be generated in the hydroelectric power stations in the state.

Shifting pattern analysis is an efficient tool to understand the fundamental information in hydro-meteorological data such as rainfall, discharge, temperature, etc. Especially, the result of a reasonable change point detection method can be effectively used in the prediction of flood and drought because it provides a key to resolve the inhomogeneous problem by climate change.

Forecast of rainfall is essential for planning and management of water resources in India, wherein about 65 per cent of the total cultivated land is under the rainfed agriculture system. Monthly and seasonal rainfall forecasts provide useful information for water resource management, agricultural planning and associated crop insurance application.

As rainfall is an unpredictable natural phenomenon, statistical analysis of distribution patterns and amount of rainfall plays an important role. The study on the statistical analysis of rainfall trend may help the government in agricultural planning and

policymaking, it helps farmers in contingency planning of the crop, adopting the farm production practices, management of the farm production, etc.

Based on the common rainfall distribution pattern, Karnataka is classified into four meteorological subdivisions by KSNDMC namely North Interior Karnataka (NIK), South Interior Karnataka (SIK), Coastal and Malnad. For present study, the secondary data of monthly rainfall (mm) of four meteorological subdivisions of Karnataka were collected from All India Coordinated Research Project (AICRP), Agro-Meteorology, UAS, GKVK, Bengaluru and Karnataka State Natural Disaster Monitoring Center (KSNDMC), Yelahanka, Bengaluru, for the period of 60 years from 1960-2019.

Considering the above points, the following objectives are laid out for the present study:

1. To analyze the trend in the rainfall pattern in meteorological subdivisions of Karnataka
2. To study the shift in rainfall distribution pattern
3. To evaluate a suitable non-linear time series models for forecasting rainfall

To analyze the pattern of the trend, monthly, seasonal and annual rainfall data are analyzed with non-parametric Mann-Kendal (M-K), Modified Mann-Kendall (MM-K) tests and Sen's slope estimator are employed. To know the shift in rainfall distribution pattern Likelihood Ratio test is employed. To find out best time-series models for forecasting monthly rainfall in different meteorological subdivisions of Karnataka, nonlinear models such as ARCH, GARCH and ANN models were employed along with linear SARIMA and Holt-Winters exponential smoothing for comparing forecasting performance of these models.

5.1 To analyze the trend in the rainfall pattern in meteorological subdivisions of Karnataka

To analyze trend in rainfall, collected monthly rainfall (mm) data for the period of 60 years from 1960-2019 were converted into month-wise, seasonal (winter, pre-monsoon, monsoon and post-monsoon) and annual rainfall (mm) data. The statistical analysis was

done separately for each subdivision, and for all data sets. Mann-Kendall (M-K) and Modified Mann-Kendall (MM-K) tests were used to know the presence or absence of monotonic trend and Sen's slope estimator method to quantify the trend in rainfall data.

The M-K test results revealed that for all data sets have no monotonic trend except for February month of Coastal region, which has monotonic increasing trend. This may be due to the influence of serial correlation in the rainfall data. Therefore, to overcome the effect of serial correlation, if any, MM-K test was employed.

The MM-K test for NIK subdivision revealed that January, February, March and August months have monotonic increasing trend and April, May, July and September months have monotonic decreasing trend whereas June, October, November and December have no monotonic trend in monthly rainfall data. Winter season have monotonic increasing trend, monsoon and post-monsoon seasons have monotonic decreasing trend, and pre-monsoon season have no monotonic trend in seasonal rainfall data. Annual rainfall data have monotonic decreasing trend.

The MM-K test for SIK subdivision revealed that February, March, April, June and August months have monotonic increasing trend whereas, January, May, July, September, October, November and December have no monotonic trend in monthly rainfall data. Winter, pre-monsoon and monsoon seasons have monotonic increasing trend and post-monsoon season have no monotonic trend in seasonal rainfall data. Annual rainfall data have monotonic increasing trend.

The MM-K test for Malnad subdivision revealed that June month have monotonic increasing trend and remaining months namely January, February, March, April, May, July, August, September, October, November and December have no monotonic trend in monthly rainfall data. Winter, pre-monsoon, monsoon and post-monsoon seasons have no monotonic trend in seasonal rainfall data. Annual rainfall data have no monotonic trend.

The MM-K test for coastal subdivision revealed that January, February, March and April months have monotonic increasing trend and July month have monotonic decreasing trend whereas namely May, June, August, September, October, November and December

have no monotonic trend in monthly rainfall data. Winter season have monotonic increasing trend and pre-monsoon, monsoon and post-monsoon seasons have no monotonic trend in seasonal rainfall data. Annual rainfall data have monotonic increasing trend.

Sen's slope value towards zero for monthly and seasonal rainfall data for all region indicates that there is a negligible rate of increasing or decreasing change in rainfall data. Relatively higher value of Sen's slope values for annual rainfall data for all subdivision indicates that high rate of decreasing change in rainfall for NIK subdivision, high rate of increasing change in Malnad and Coastal subdivisions.

5.2 To study the shift in rainfall distribution pattern

To analyze rainfall shifting pattern (shifting point), collected monthly rainfall (mm) data for the period of 60 years from 1960-2019 were converted into month-wise, seasonal (winter, pre-monsoon, monsoon and post-monsoon) and annual rainfall (mm) data. The statistical analysis was done separately for each subdivision, and for all data sets. The Likelihood Ratio test was used to know the rainfall pattern-shifting year with respect to average rainfall over the period.

The Likelihood Ratio for NIK subdivision showed that shifting point (year) in monthly rainfall was found in the month of February (2007), March (2013), April (1963), June (1972), August (1968), October (1975), November (1987) and December (1989), and there were no shift in rainfall was observed in the month of January, May, July and September. Seasonal rainfall data revealed shifting point in all seasons *viz.* winter (2015), pre-monsoon (1963), monsoon (1960) and post-monsoon (2002). For annual rainfall, shifting point was observed in the year 2010.

For SIK subdivision, shifting point in monthly rainfall was found in the all the months for SIK subdivision *i.e.*, January (2014), February (1998), March (2013), April (1993), May (2003), June (1976), July (2016), August (1994), September (1989), October (1990), November (2015) and December (1972). Shifting point in seasonal rainfall was

found in the all seasons namely winter (1993), pre-monsoon (2003), monsoon (1970) and post-monsoon (1990). For annual rainfall, shifting point was observed in the year 1968.

For Malnad subdivision, shifting point in monthly rainfall was found in the month of February (1963), March (2012), May (1962), June (2014), July (1962), August (2016), September (2005) and December (1966) and there were no shift in rainfall was observed in the month of January, April, October and November. Shifting point in seasonal rainfall was found for the season winter (2009), pre-monsoon (1962) and monsoon (1964) and no shifting pattern is observed for post-monsoon season. For annual rainfall, shifting point was observed in the year 2016.

For Coastal subdivision, shifting point in monthly rainfall was found in the month of March (2016), April (2000), May (1962), June (1967), July (1999), September (2004), November (1965) and December (2008) and there were no shift in rainfall was observed in the month of January, February, August and October. Shifting point in seasonal rainfall was found for the season winter (1998) and pre-monsoon (1962) and no shifting pattern is observed for monsoon and post-monsoon seasons. For annual rainfall, shifting point was observed in the year 1966.

5.3 To evaluate a suitable non-linear time series models for forecasting rainfall

For rainfall forecasting purpose, the monthly rainfall data was used for the period of 60 years from 1960-2019. Forecasting was done separately for each subdivision. For forecasting purpose, two linear time-series models *viz.* H-WES and SARIMA models, and three nonlinear time-series *viz.* ARCH, GARCH and MLP-NN (ANN) were used.

For NIK subdivision, model with smoothing parameters $\alpha = 0.01$, $\beta = 0.00000004$ and $\gamma = 0.001$ was found to be suitable H-WES model. Monthly rainfall data was seasonally differenced to achieve stationarity. The SARIMA, ARCH and GARCH models with orders respectively SARIMA $(0,0,0)(2,1,2)_{12,}$ ARCH(1) and GARCH(1,1) were found to be best-fitted based on lowest RMSE values. Then MLP-NN (ANN) with one input layer with 12 input nodes, one hidden layer with four hidden nodes and one output layer with one output node and 'Tanh' activation function is found to be best fitted

ANN model. among this five best fitted models ANN was best model with lowest RMSE value for both training (28.04) and testing (25.29) data sets.

For SIK subdivision, model with smoothing parameters $\alpha = 0.094$, $\beta = 0.0000003$ and $\gamma = 0.00002$ was found to be suitable H-WES model. Monthly rainfall data was seasonally differenced to achieve stationarity. The SARIMA, ARCH and GARCH models with orders respectively SARIMA $(0,0,0)(2,1,2)_{12,}$ ARCH(1) and GARCH(1,1) were found to be best-fitted based on lowest RMSE values. Then MLP-NN (ANN) with one input layer with 12 input nodes, one hidden layer with four hidden nodes and one output layer with one output node and linear activation function is found to be best fitted ANN model. among this five best fitted models ANN was best model with lowest RMSE value for both training (33.22) and testing (33.31) data sets.

For Malnad subdivision, model with smoothing parameters $\alpha = 0.04$, $\beta = 0.00003$ and $\gamma = 0.19$ was found to be suitable H-WES model. Monthly rainfall data was seasonally differenced to achieve stationarity. The SARIMA, ARCH and GARCH models with orders respectively SARIMA $(2,0,0)(0,1,1)_{12,}$ ARCH(1) and GARCH(1,1) were found to be best-fitted based on lowest RMSE values. Then MLP-NN (ANN) with one input layer with 12 input nodes, one hidden layer with five hidden nodes and one output layer with one output node and 'Logistic' activation function is found to be best fitted ANN model. among this five best fitted models ANN was best model with lowest RMSE value for both training (62.52) and testing (153.89) data sets.

For Coastal subdivision, model with smoothing parameters $\alpha = 0.01, \beta = 0.00002$ and $\gamma = 0.002$ was found to be suitable H-WES model. Monthly rainfall data was seasonally differenced to achieve stationarity. The SARIMA, ARCH and GARCH models with orders respectively SARIMA $(0,0,0)(0,1,2)_{12,}$ ARCH(1) and GARCH(1,1) were found to be best-fitted based on lowest RMSE values. Then MLP-NN (ANN) with one input layer with 12 input nodes, one hidden layer with four hidden nodes and one output layer with one output node and 'Tanh' activation function is found to be best-fitted ANN model. Among these five best-fitted models ANN was best model with lowest RMSE value for both training (98.80 and testing (238.67) data sets.

5.4 CONCLUSION:

This study has analyzed the month-wise, seasonal and annual rainfall trend employing three non-parametric tests *viz.* Mann-Kendall (M-K), Modified Mann-Kendall (MM-K) tests and Sen's slope estimator for all the four meteorological subdivisions of Karnataka using 60 years (1960-2019) rainfall data. Use of M-K test showed non-significant trend with all data sets except February month for Coastal region, which has monotonic increasing trend. Whereas, MM-K test showed significant monotonic increasing and decreasing trend for majority of data sets. Results of Sen's slope estimator revealed that the rate of increasing or decreasing change for any data set having significant trend is negligible.

The shift year inferred from the Likelihood Ratio test revealed the probable year of shift in month-wise, seasonal and annual rainfall data. It was observed that rainfall is not uniform and varies across months, seasons, years and subdivisions. Rainfall pattern shifting years are scattered between the time period 1962-2016, this evidently explains about dynamic nature of rainfall distribution pattern over time and space.

For monthly rainfall forecasting, five time-series models namely, H-WES, SARIMA, ARCH, GARCH and ANN model were used. Based on the lowest RMSE values for both training and testing data sets it was found that ANN model was performed better than other four models for all the subdivisions. Hence, ANN model can be used for forecasting monthly rainfall data in meteorological subdivisions of Karnataka.

VI REFERENCES

ADARSH, S. AND REDDY, J. M., 2015, Trend analysis of rainfall in four meteorological subdivisions of southern India using nonparametric methods and discrete wavelet transforms. *Int. J. Climatol.*, **35**(6)**:** 1107-1124.

ACHAL, L., 2013, A study on agricultural commodity price volatility using dynamic neural networks. M.Sc. (Agri.) *Thesis* (Unpub.), Ind. Agril. Res. Inst., New Delhi.

ALVAREZ, E. E. AND DEY, D. K., 2009, Bayesian isotonic change point analysis. *Ann. Inst. Stat. Math.*, **61**(2)**:** 355-370.

ANONYMOUS, 2017, Report on drought vulnerability assessment in Karnataka, KSNDMC, Bangalore, 1-103.

ANONYMOUS, 2019, Annual report (2018-19), Department of Agriculture, Cooperation and Farmers Welfare, New Delhi, 1-221.

AZAD, S., DEBNATH, S. AND RAJEEVAN, M., 2015, Analyzing predictability in Indian monsoon rainfall: a data analytic approach. *Environ. Processes*, **2**(4)**:** 717-727.

AZAM, M., MAENG, S. J., KIM, H. S., LEE, S. W. AND LEE, J. E., 2018, Spatial and temporal trend analysis of precipitation and drought in South Korea. *Water*, **10**(6)**:** 765-792.

BARI, S. H., RAHMAN, M. T. U., HOQUE, M. A. AND HUSSAIN, M. M., 2016, Analysis of seasonal and annual rainfall trend in the northern region of Bangladesh. *A J. Forecasting, Pract. Appl., Trng. Techniques Model.*, **100**(176)**:**148-158.

BAYAZIT, M. AND ONOZ, B., 2007, To prewhiten or not to prewhiten in trend analysis?. *Hydrol. Sci. J.*, **52**(4)**:** 611-624.

BHOWMIK, M., DAS, N., AHMED, I. AND DEBNATH, J., 2017, Rainfall frequency analysis to predict flood in West Tripura district, Tripura, North-East India. *Int. J. Geomatics Geosciences*, **7**(3)**:** 310-320.

CHAKRABORTY, S., PANDEY, R. P., CHAUBE, U. C. AND MISHRA, S. K., 2013, Trend and variability analysis of rainfall series at Seonath River Basin, Chhattisgarh (India). *Int. J. Appl. Sci. Engr. Res.*, **2**(4)**:** 425-43.

CHATTOPADHYAY, S. AND CHATTOPADHYAY, G., 2008, A comparative study among different neural net learning algorithms applied to rainfall time series. Meteorological Applications. *J. Forecasting, Pract. Appl., Trng. Techniques Model.*, **15**(2)**:** 273-280.

CHATTOPADHYAY, S. AND CHATTOPADHYAY, G., 2010, Univariate modelling of summer-monsoon rainfall time series: comparison between ARIMA and ARNN. *Comptes. Rendus. Geoscience*, **342**(2)**:** 100-107.

CHEN, J. AND GUPTA, A.K., 2011, *Parametric statistical change point analysis: with applications to genetics, medicine, and finance*. Birkhauser Basel, Springer Science and Business Media, Spring Street, New York.

DACHIAN, S., 2010, On limiting likelihood ratio processes of some change-point type statistical models. *J. Stat. Plan. Inference*, **140**(9)**:** 2682-2692.

DELITALA, A., CESARI, D., CHESA, P. AND WARD. M., 2000, Precipitation over Sardinia (Italy) during the 194-1993 rainy season and associated large scale climate variation, *Int. J. Climatol.*, **20**(5)**:** 519-541.

DETTE, H. AND GOSMANN, J., 2020, A likelihood ratio approach to sequential change point detection for a general class of parameters. *J. Am. Stat. Assoc.*, **115**(531)**:** 1361-1377.

DORE, M. H. I., 2005, Climate change and changes in global precipitation patterns: What do we know?. *Environ. Int.*, **31**(8)**:** 1167–1181.

GAJBHIYE, S., MESHRAM, C., MIRABBASI, R. AND SHARMA, S.K., 2016, Trend analysis of rainfall time series for Sindh river basin in India. *Theor. Appl. Climatol.*, **125**(4)**:** 593-608.

GARBRECHT, J. D., ZHANG, X. C., SCHNEIDER, J. M. AND STEINER, J. L., 2010, Utility of seasonal climate forecasts in management of winter-wheat grazing. *Appl. Eng. Agric.*, **26**(5)**:** 855-866.

GHOSH, H., PAUL, R. K. AND PRAJNESHU., 2010, Nonlinear time series modeling and forecasting for periodic and ARCH effects. *J. Stat. Theory Pract.*, **4**(1)**:** 27-44.

GOMBAY, E. AND HORVATH, L., 1994, An application of the maximum likelihood test to the change-point problem. *Stoch. Process. Their Appl.*, **50**(1)**:** 161-171.

GUHATHAKURTA, P., RAJEEVAN, M., SIKKA, D. R. AND TYAGI, A., 2015, Observed changes in southwest monsoon rainfall over India during 1901-2011. *Int. J. Climatol.*, **35**(8)**:** 1881-1898.

HAMED, K.H., 2008, Trend detection in hydrologic data: the Mann-Kendall trend test under the scaling hypothesis. *J. hydrology*, **349**(3-4)**:** 350-363.

HAMED, K.H. AND RAO, A.R., 1998, A modified Mann-Kendall trend test for autocorrelated data. *J. hydrol.*, **204**(1-4)**:** 182-196.

HAWKINS, D.M., 1977, Testing a sequence of observations for a shift in location. *J. Am. Stat. Assoc.*, **72**(357)**:** 180-186.

HONG, W.C., 2008, Rainfall forecasting by technological machine learning models. *Appl. Math. Comput.*, **200**(1)**:** 41-57.

HUNG, N.Q., BABEL, M.S., WEESAKUL, S. AND TRIPATHI, N.K., 2009, An artificial neural network model for rainfall forecasting in Bangkok, Thailand. *Hydrol. Earth Syst. Sci.*, **13**(8)**:** 1413-1425.

HU, Z., LIU, S., ZHONG, G., LIN, H. AND ZHOU, Z., 2020, Modified Mann-Kendall trend test for hydrological time series under the scaling hypothesis and its application. *Hydrol. Sci. J*., **7:**1-20

HUSAK, G. J., MICHAELSEN, J. AND FUNK, C., 2007, Use of the gamma distribution to represent monthly rainfall in Africa for drought monitoring applications. *Int. J. Climatol.*, **27**(7)**:** 935-944.

ISIOMA, I. N., RUDOLPH, I. I. AND OMENA, A. L., 2018, Non-parametric Mann-Kendall test statistics for rainfall trend analysis in some selected states within the coastal region of Nigeria. *Civil Constr. Environ. Eng.*, **3**(1)**:** 17-28.

JAIN, A. AND KUMAR, A. M., 2007, Hybrid neural network models for hydrologic time series forecasting. *Apld. Soft Computing*, **7**(2)**:** 585-592.

JAISWAL, R. K., LOHANI, A. K. AND TIWARI, H. L., 2015, Statistical analysis for change detection and trend assessment in climatological parameters. *Environ. Process.*, **2**(4)**:** 729-749.

JANDHYALA, V., FOTOPOULOS, S., MACNEILL, I. AND LIU, P., 2013, Inference for single and multiple change-points in time series. *J. Time Ser. Anal.*, **34**(4)**:** 423-446.

JOSEPH, A. AND TAMILMANI, D., 2017, Markov chain model of weekly rainfall probability and dry and wet spells for agricultural planning in Coimbatore in western zone of Tamil Nadu. *Ind. J. Soil Conserv.*, **45**(1)**:** 66-71.

KARIM, T. H., KEYA, D. R. AND AMIN, Z. A., 2018, Temporal and spatial variations in annual rainfall distribution in Erbil province. *Outlook on Agri.*, **47**(1)**:** 59-67.

KILLICK, R. AND ECKLEY, I., 2014, Changepoint: An R package for changepoint analysis. *J. stat. softw.*, **58**(3)**:** 1-19.

KRISHNAKUMAR, K. N., RAO, G. P. AND GOPAKUMAR, C. S., 2009, Rainfall trend in twentieth century over Kerala, India. *Atmos. Environ.*, **43**(11)**:** 1940-1944.

KULKARNI, A. AND VON STORCH, H., 1995, Monte Carlo experiments on the effect of serial correlation on the Mann-Kendall test of trend. *Meteorologische Zeitschrift*, **4**(2)**:** 82-85.

KUMAR, V., JAIN, S. K. AND SINGH, Y., 2010, Analysis of long-term rainfall trend in India. *Hydrol. Sci. J.*, **55**(4)**:** 484-496.

KUMAR, A., TRIPATHI, P., GUPTA, A., SINGH, K. K., SINGH, P. K., SINGH, R. AND TRIPATHI, A., 2018, Rainfall variability analysis of Uttar Pradesh for crop planning and management. *Mausam*, **69**(1)**:** 141-146.

LEE, S. AND KIM, S. U., 2016, Comparison between change-point detection methods with synthetic rainfall data and application in South Korea. *KSCE J. Civ. Eng.*, **20**(4)**:** 1558-1571.

LIU, Y., ZOU, C. AND ZHANG, R., 2008, Empirical likelihood ratio test for a change-point in linear regression model. *Commun. Stat. Theory Methods*, **37**(16)**:** 2551-2563.

LUK, K. C., BALL, J. E. AND SHARMA, A., 2001, an application of artificial neural networks for rainfall forecasting. *Math. Comput. Model.*, **33**(6)**:** 683-693.

MACHADO, F., MINE, M., KAVISKI, E. AND FILL, H., 2011, Monthly rainfall-runoff modelling using artificial neural networks. *Hydrol. Sci. J.*, **56**(3)**:** 349-361.

MAPURISA, B. AND CHIKODZI, D., 2014, An assessment of trend of monthly contributions to seasonal rainfall in South-Eastern Zimbabwe. *Am. J. Climate Change*, **3**(1)**:** 50-59.

MAKRIDAKIS, S., STEVEN, C. W. AND HYNDMAN, R. J., 1998, *Forecasting methods and applications*, 3rd edition, John Wiley & sons, New York.

MEHDIZADEH, S., BEHMANESH, J. AND KHALILI, K., 2018, New approaches for estimation of monthly rainfall based on GEP-ARCH and ANN-ARCH hybrid models. *Water resour. Manag.*, **32**(2)**:** 527-545.

MODARRES, R. AND DA SILVA, V. D. P. R., 2007, Rainfall trend in arid and semi-arid regions of Iran. *J. Arid Envir.*, **70**(2)**:** 344-355.

MOHAN KUMAR, T.L. AND PRAJNESHU, 2014, An application of nonlinear least squares support vector machine using particle swarm optimization technique. *Int. J. Agric. Stat. Sci.*, **10**(2)**:** 401-404.

MOHAN KUMAR, T. L., MALLIKARJUNA, H. B., SINGH, N. M. AND SATHISHGOWDA, C. S., 2011, Comparative study of univariate time series technique for forecasting of onion price. *Int. J. Com. Bus. Manage,* **4**(2)**:** 304-308.

MOHAN KUMAR, T. L., SATHISHGOWDA, C. S., MUNIRAJAPPA, R. AND SURENDRA, H. S., 2012, Nonlinear statistical growth models for describing trend in area under coffee production in India. *Mysore J. Agric. Sci.,* **46** (4)**:** 745-750.

MONDAL, A., KUNDU, S. AND MUKHOPADHYAY, A., 2012, Rainfall trend analysis by Mann-Kendall test: A case study of north-eastern part of Cuttack district, Orissa. *Int. J. Geology, Earth Environ. Sci.*, **2**(1)**:** 70-78.

NAMRATHA, K. S., 2019, Evaluation of statistical models for forecasting of rainfall in meteorological subdivisions of Karnataka. M.Sc. (Agri.) *Thesis (Unpub.),* Univ. Agric. Sci., Bangalore.

NARAYANAN, P., BASISTHA, A. AND SACHDEVA, K., 2016, Understanding trend and shifts in rainfall in parts of northwestern India based on global climatic indices. *Weather*, **71**(8)**:** 198-203.

OSMAN, Y. Z. AND ABDELLATIF, M. E., 2016, Improving the accuracy of downscaling rainfall by combining predictions of different statistical downscale models. *Water Sci.*, **30**(2)**:** 61-75.

PAL, I. AND AL-TABBAA, A., 2011, Assessing seasonal precipitation trend in India using parametric and non-parametric statistical techniques. *Theor. Appl. Climatol.*, **103**(1-2)**:** 1-11.

PANDEY, P.K., TRIPURA, H. AND PANDEY, V., 2019, Improving prediction accuracy of rainfall time series by hybrid SARIMA-GARCH modeling. *Nat. Resour. Res.*, **28**(3)**:** 1125-1138.

RAHMAN, A., ANIK, A., FARHANA, Z., DEVNATH, S. AND AHMED, Z., 2018, Pattern recognition of rainfall using wavelet transform in Bangladesh. *Open J. Stat.*, **8**(1)**:** 134-145.

RAMANA, R.V., KRISHNA, B., KUMAR, S. R. AND PANDEY, N. G., 2013, Monthly rainfall prediction using wavelet neural network analysis. *Water resour. Mngmt.*, **27**(10)**:** 3697-3711.

ROMAN, U. C., PATEL, P. L. AND POREY, P. D., 2012, Prediction of missing rainfall data using conventional and artificial neural network techniques. *J. Hydraulic Engn.*, **18**(3)**:**224-231.

SHARMA, S. AND SAHA, A.K., 2017, Statistical analysis of rainfall trend over Damodar river basin, India. *Arab. J. Geosci.*, **10**(15)**:** 319-331.

SOMVANSHI, V. K., PANDEY, O. P., AGRAWAL, P. K., KALANKER, N. V., PRAKASH, M. R. AND CHAND, R., 2006, Modeling and prediction of rainfall using artificial neural network and ARIMA techniques. *J. Ind. Geophys. Union*, **10**(2)**:** 141-151.

SONALI, P., AND KUMAR, D. N., 2013, Review of trend detection methods and their application to detect temperature changes in India. *J. Hydrol.*, **476:** 212 -227.

SRIDHARA, S., GOPAKKALI, P. AND NANDINI, R., 2020, Trend analysis of precipitation and temperature over different districts of Karnataka: an aid to climate change detection and cropping system option. *Int. J. Envi. Climate Change,* **10**(3)**:** 15-25.

SRINIVASAREDDY, H. S., SHIVAKUMARNAIKLAL, N. G., KEERTHY, PRASAD GARAG, EMILY PRABHA JOTHI AND CHALLA, O., 2019, Drought vulnerability assessment in Karnataka: Through composite climatic index. *Mausam,* **70**(1)**:** 159-170.

TADESSE, K. B. AND DINKA, M. O., 2017, Application of SARIMA model to forecasting monthly flows in Waterval River, South Africa. *J. Water Land Dev.*, **35**(1)**:** 229-236.

TOTH, E., BRATH, A. AND MONTANARI, A., 2000, Comparison of short-term rainfall prediction models for real-time flood forecasting. *J. Hydrol.*, **239**(1)**:** 132-147.

TYAGI, A., 2008, Forecasters guide. *India Meteorological Department (IMD), Pune, India, 16*.

VIGNESH, R., JOTHIPRAKASH, V. AND SIVAKUMAR, B., 2019, Spatial rainfall variability in peninsular India: a nonlinear dynamic approach. *Stoch. Environ. Res. Risk Assess.,* **33**(2)**:** 465-480.

VIJAYA KUMAR P., BAL, S.K. AND SUBBA RAO A.V. M., 2019, All India Coordinated Project on Agrometeorology, Annual Report (2018-19), ICAR-Central Research Institute for Dryland Agriculture, Hyderabad, 1-144.

VERMA, M. K., VERMA, M. K. AND SWAIN, S., 2016, Statistical analysis of precipitation over Seonath river basin, Chhattisgarh, India. *Int. J. App. Engr. Res.*, **11**(4)**:** 2417-2423.

YAO, Y. C. AND DAVIS, R. A., 1986, The asymptotic behavior of the likelihood ratio statistics for testing shift in mean in a sequence of independent normal variates. *Sankhya*, **48**(3)**:** 339-353.

YAU, C.Y. AND ZHAO, Z., 2016, Inference for multiple change points in time series via likelihood ratio scan statistics. *J. Roy. Stat. Soc. B.*, **78**(4)**:** 895-916.

YUE, S. AND WANG, C., 2004, The Mann-Kendall test modified by effective sample size to detect trend in serially correlated hydrological series. *Water Resour. Manag.*, **18**(3)**:** 201-218.

YUE, S., PILON, P., PHINNEY, B. AND CAVADIAS, G., 2002, The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrol. Process.*, **16**(9)**:** 1807-1829.

YUSOF, F., KANE, I.L. AND YUSOP, Z., 2013, Hybrid of ARIMA-GARCH modeling in rainfall time series. *J. Technol.*, **63**(2)**:** 27-34.

ZARENISTANAK, M., DHORDE, A.G. AND KRIPALANI, R.H., 2014, Trend analysis and change point detection of annual and seasonal precipitation and temperature series over southwest Iran. *J. Earth System Sci.*, **123**(2)**:** 281-295.

ZHAO, J. H., DONG, Z. Y. AND ZHAO, M. L., 2009, A statistical model for flood forecasting. *Australas. J. Water Res.*, **13**(1)**:** 43-52.