Time Series Analysis of Decadal Precipitation Pattern at Selected Cities of Southern India

T. S. Subbiah*, P. Parthiban*, R. Mahesh* and A. Das*†

*Centre for Environmental Engineering, PRIST Deemed University, Thanjavur, Tamil Nadu, India
†Corresponding author: A. Das; scientists.crd@gmail.com

ABSTRACT

To characterize and explore the short-term climatic patterns over the last decade (Jan. 2009 to Dec. 2018), the present research has been carried out, involving time series analysis of precipitation pattern in three cities of Tamil Nadu, namely, Thanjavur, Nagapattinam, and Chennai, referring to deltaic, coastal and highly urbanized cities of Tamil Nadu, respectively. The study involves time series empirical analysis, decomposition, exponential smoothing, and various stochastic modeling. Herein, the location-specific suitable models are obtained and specific predictions are being carried out, as well.

INTRODUCTION

Rainfall trends in response to climate change and probability analysis have been studied by various researchers (Jayawardene et al. 2005, Parta & Kahya 2006, Kumar & Singh 2011, Obot et al. 2010, Kumar & Jain 2010, Manikandan & Tamilmani 2012, Manivannan et al. 2016, Mohanty et al. 2000) and they have emphasized that the knowledge of location-specific rainfall variations is essential for proper water harvesting and water management practices.

In India, nearly 65% of the cultivable area is under rainfed agriculture. Rainfall is one of the important factors deciding the success of rainfed agriculture of a particular agro-ecological region where the major part of the precipitation is rainfall, which shows definite changes over time (Banze et al. 2018, Sharma et al. 2018). Kumar and Jain (2010) reported that higher or lower rainfall or changes in rainfall distribution would influence the spatial and temporal distribution of runoff, soil moisture, groundwater storage and would alter the frequency of droughts and floods. While the observed monsoon rainfall at the national level does not show any significant trend, regional monsoon variations have been recorded. An increasing trend of monsoon season rainfall has been found along the west coast, northern Andhra Pradesh, and northwestern India (+10% to +12% of the normal over the last 100 years) while a decreasing trend of monsoon seasonal rainfall has been observed over eastern Madhya Pradesh, North-eastern India and some parts of Gujarat and Kerala (6 to 8% of the normal over the last 100 years) (Kumar & Singh 2011, Singh & Kumar 2016).

In fact, comprehensive knowledge of the trend and persistence in rainfall of the area is of great importance because of the economic implications of rain-sensitive operations (Sharma et al. 2015, Sharma & Dubey 2013, Jakhar et al. 2011). Keeping the above in view, the present study has been conducted to analyze the precipitation pattern of selected study areas and evaluation of various modeling options and predictability.

STUDY AREA

The locations selected for the study are Thanjavur, Nagapattinam and Chennai to have a synoptic coverage of major hydro-lithographic domains namely deltaic and agricultural region, coastal and metropolitical region respectively. The study area lies in the following geocoordinates: $10^{\circ}45'00''$ N to $10^{\circ}49'30''$ N, $79^{\circ}6'00''$ E to $79^{\circ}10'30''$E; $10^{\circ}41'32.787''$ N to $10^{\circ}49'50.541''$ N, $79^{\circ}49'33.907''$E to 79° $51'13.19''$E and $12^{\circ}.58'38.203''$ N to $13^{\circ}8'31.381$ N, $80^{\circ}14'7.673''$E to $80^{\circ}18'42.037''$E for Thanjavur, Nagapattinam and Chennai respectively. The base map of the study area is shown in Fig. 1.

The average temperature in Thanjavur is 28.7°C and the precipitation averages 938 mm. The average maximum temperature of Nagapattinam during the summers remains around 35°C. The relative humidity hovers around 60 to 65%.
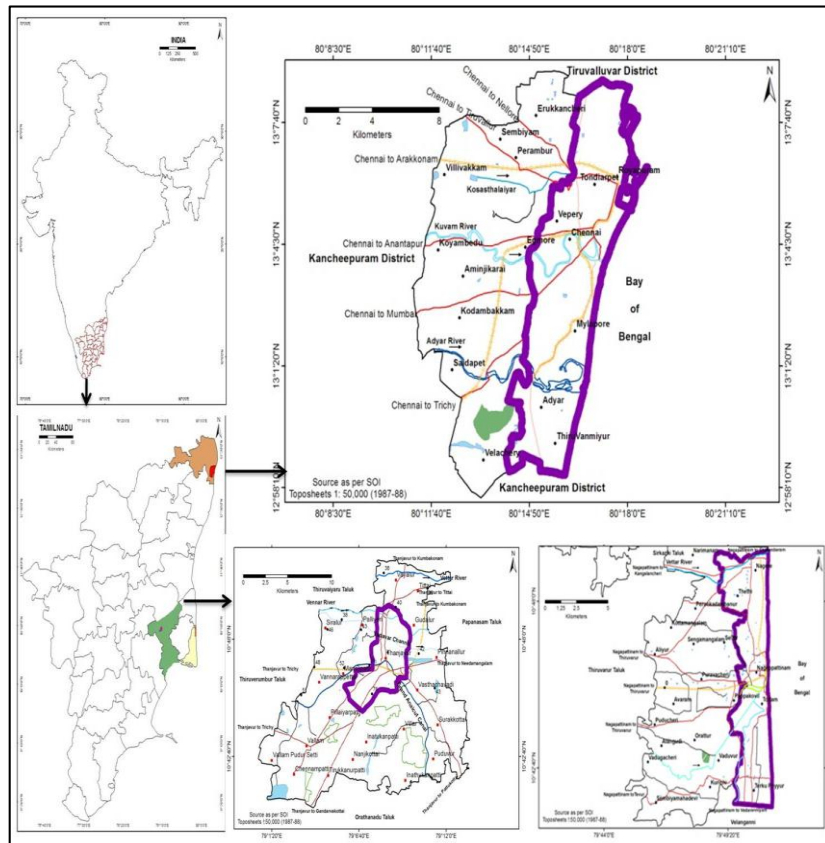
Fig. 1: Location map of the study area.

In Chennai, the average annual temperature is 28.6°C. In a year, precipitation is around 5.43 mm and humidity at 70%.

MATERIALS AND METHODS

General Analyses of Decadal Rainfall Pattern

Monthly rainfall data of 10 years for the period 2009 to 2018 was collected from the meteorological observatory center (IMD) and annual rainfall data was computed from monthly rainfall data. The statistical analysis of rainfall dissimilarity of the study regions (Thanjavur, Nagapattinam, and Chennai) displays a quite good range from 0 to 449.6, 0 to 522.45, and 0 to 780.2 mm respectively with an annual average rainfall value of 60.38, 83.47, and 66.8 mm respectively.

The Descriptive Statistics and Box Whisker Plots of Rainfall Data

Statistical parameters such as mean, median, standard deviation, range, coefficient of variation, skewness, and kurtosis of these rainfalls were also calculated. The trend of annual, seasonal, and monthly rainfall data was calculated using box-whisker plots.

Decomposition of Time Series Data of Study Area

R (Version 1.1.463) was used to carry out the multiplicative decomposition of the time series data in terms of trend, cyclicity, and seasonality. Decompose and forecast functions of the R tool are used to split the time series into seasonality, trend, and error components.

Time Series Modelling

Four standard time series modeling approaches were adopted to explore the frequency domain of the precipitation data, namely:

1. Holt-Winters: The Holt-Winters forecasting algorithm allows users to smooth a time series and use that data to forecast areas of interest.

2. Linear Regression: It is used to predict the value of an outcome variable Y based on one or more input predictor variable (X-variables). The aim is to establish a linear relationship (a mathematical formula) between the predictor variable(s) and the response variable, so that, we can use this formula to estimate the value of the response Y when only the predictors (Xs) values are known.

3. ETS (Error, Trend, and Seasonal): ETS method is an approach method for forecasting time series univariate.

4. ARIMA (Autoregressive Integrated Moving Average): ARIMA is a statistical analysis model that uses time series data to either better understand the data set or to predict future trends.

The Auto-ARIMA algorithm was used to obtain the optimal parameters (p, d & q), rather than using ACF and PACF (plots), manually.

In addition to the standard metrics (namely, mean, standard deviation, and R^2), three other important metrics (viz. AIC (Akaike's Information Criteria), AICc (corrected AIC), and BIC (Bayesian information criteria)) are being employed for model evaluation and selection for ETS and ARIMA, being an unbiased estimate of the model prediction error.

The basic idea of AIC is to penalize the inclusion of additional variables in a model. It adds a penalty that increases the error when including additional terms, thereby discouraging overfitting. The lower the AIC, the better the model (Akaike 1974).

Let k be the number of estimated parameters in the model. Let L be the maximum value of the likelihood function for the model. Then the AIC value of the model is the following.

$$AIC = 2k - \ln(L)$$

AICc is a version of AIC corrected for small sample sizes (Burnham & Anderson 2002), which is defined in terms of the correction term below, where K is the number of parameters and n is the sample size.

$$AICc = AIC + \frac{2k^2 + 2k}{n - k - 1}$$

BIC is a variant of AIC with a stronger penalty for including additional variables in the model.

RESULTS AND DISCUSSION

Time Series Variation of Precipitation Pattern

All three-study area shows a more or less similar pattern of temporal periodicity. The rainfall pattern from 2009-2012 and 2015 shows higher precipitation compared to that during 2013, 2014, 2016, and 2017. The time series of annual rainfall has been presented in Fig. 2.

However, the precipitation in 2018 has been the highest compared to the previous 9 years. Interestingly maximal contribution has been recorded during November 2018 (possibly due to the striking of the Gaja cyclone in November 2018). Among the three study areas, the maximum monthly rainfall had been obtained in Thanjavur during 2009-2014.

As indicated in the summary statistics, Thanjavur experiences the highest mean and median rainfall, compared to the other two study areas (namely, Chennai and Nagapattinam), whereas standard deviation, standard error, and range are the least for Chennai. All the locations show negatively skewed and leptokurtic and positively skewed, the highest being at Nagapattinam. (Table 1).

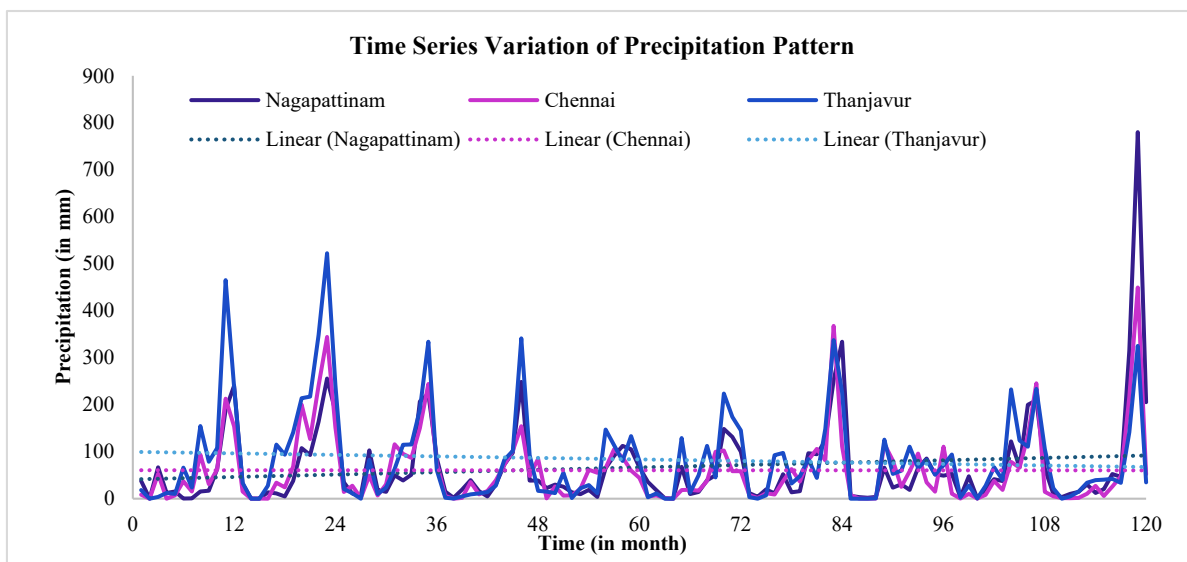### Monthly Variability of Rainfall Pattern Since Last Decade

Fig. 2: Time Series Variation of Precipitation Pattern in the Study Area.

This was estimated using box plots (Figs. 3, 4 and 5) and during November-December (NE monsoon), all the places experience the highest rainfall, with another secondary peak during May-June (South-West Monsoon) indicating two distinct monsoon showers contributing the majority of the rainfall.

As evident from box plots, in all three locations, there are two distinct periods of heavy shower primarily during October to December corresponding to Northeast monsoon and secondarily during May to August corresponding to Southwest monsoon. The onset of the southwest monsoon is early in Thanjavur and Nagapattinam (i.e. May) compared to Chennai (i.e. June) and correspondingly its duration is longer in Chennai followed by Nagapattinam and Thanjavur .

The sporadic showers as indicated by the outliers in the box plot is limited from January to June in Thanjavur and Nagapattinam whereas it extended up to August in the case of Chennai. Interestingly in all three locations, the minimum average rainfall as indicated by the first quartile is higher than 20 mm from August to December, the highest minimum being during October to November. Similarly, the highest monthly mean rainfall as indicated by the third quartile is observed during May to December in all the locations with more than 100 mm rainfall.

Table 1: Descriptive statistics.

| S. No	Parameters	Chennai	Thanjavur	Nagapattinam
1	Mean	60.387	83.471	66.864
2	Standard Error	7.131	9.102338941	8.905
3	Median	35.055	48.585	36.765
4	Mode	0	0	0
5	Standard Deviation	78.123	99.711	97.559
6	Sample Variance	6103.351	9942.308	9517.834
7	Kurtosis	7.394	4.499	23.807
8	Skewness	2.439	1.979	3.974
9	Range	449.6	522.45	780.16
10	Minimum	0	0	0.04
11	Maximum	449.6	522.45	780.2
12	Sum	7246.45	10016.55	8023.77
13	Confidence Level (95.0%)	14.121	18.023	17.634
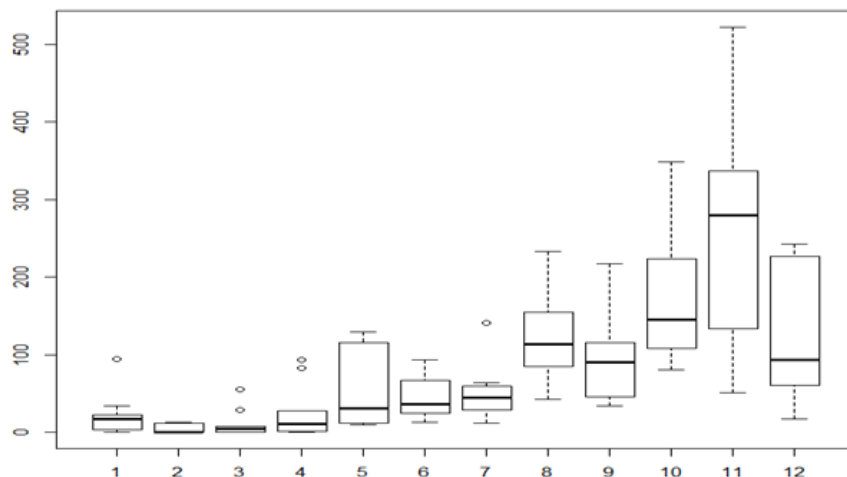
Fig. 3: Box Plots for Monthly Rainfall at Thanjavur from 2009 to 2018.

Decomposition of Multiplicative Series

On the multiplicative decomposition of the time series data of precipitation of study area, although the trend appears to be similar (i.e. primary peak during March) there is a distinct decrease in rainfall over the years compared to 2003 except 2008. In the case of Chennai, however, the magnitude of rainfall during 2008 is much lesser than the other two study areas (Fig. 5)

The seasonality of the rainfall pattern over the years is similar in all three locations in relation to the highest rainfall during November whereas, in the case of Chennai, there are two more peaks which are the secondary and tertiary peaks corresponding to May and August through the decade and only less significant peak in Thanjavur during May. Nagapattinam, shows no distinct secondary and tertiary peaks, indicating uni-model periodicity throughout the years studied.

Time Series Modeling and Forecasting of the Rainfall Data

Various models (namely, Holt-Winters, ETS, Linear Regression, and ARIMA) were attempted to model and forecast rainfall pattern, and ARIMA (0,0,1) was found to be the best model for all three locations for non-seasonal components, as reflected by the least value of error as represented by AIC, although seasonality components do vary.

This means all the time series data of the study area is stationary, as indicated by the integrated parameter zero. Besides, the autoregression of the data series is also zero, thereby establishing the requirement of only first-order moving average-term to model the rainfall data for the whole decade. Thus, the required model (non-seasonal) for forecasting the rainfall data is given by,
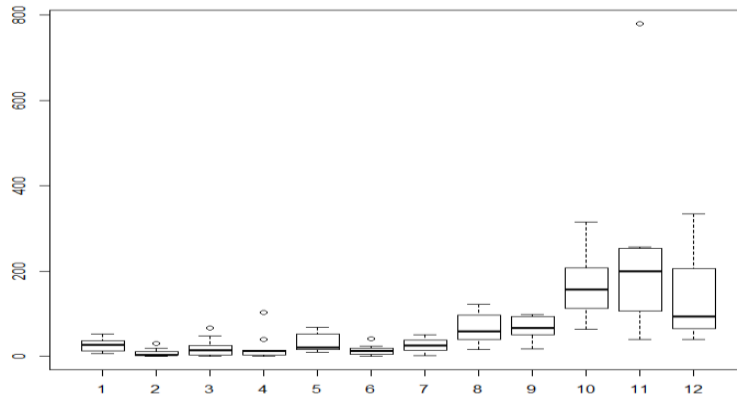

Fig. 4: Box Plots for Monthly Rainfall at Nagapattinam from 2009 to 2018.
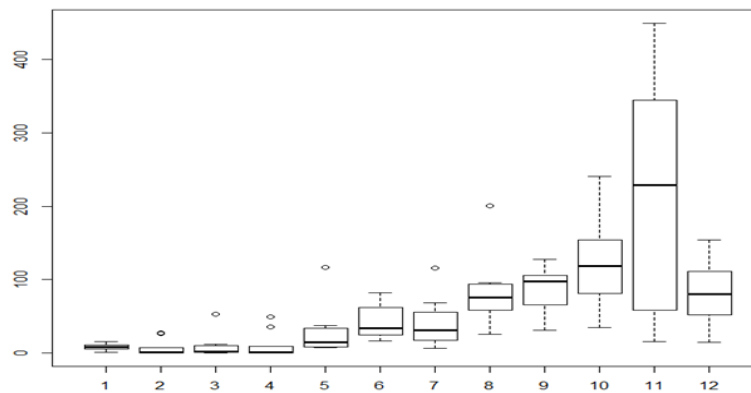

Fig. 5: Box Plots for Monthly Rainfall at Chennai from 2009 to 2018.

$$X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

Where μ is the mean of the series, the θ_1 is the coefficient of the model and the εt is white noise error terms. However to be specific, even among the first-order moving average model that satisfies all the three time series data of the precipitation, the complete model, involving the seasonal component as well, for the locations are given by ARIMA (0,0,1) (2,0,0) [12], ARIMA (0,0,1) (0,1,1) [12], and ARIMA (0,0,1) (1,0,0) [12] for Thanjavur, Nagapattinam, and Chennai, respectively. This means, the seasonal component of the locations does show distinct variation in the model, as represented by autoregression of 2nd order for Thanjavur, first-order differentiated and first-order moving average for Nagapattinam, and first-order autoregression for Chennai.

The diagnostics of various models attempted are presented in Table 2, which shows that considering standard error, the most suitable models for Thanjavur is ARIMA followed by ETS, whereas for Chennai, the most suitable models are ETS and Holt-Winters, and for Nagapattinam, Linear Regression and ETS are the most suitable models. Hence ETS and ARIMA are by far the best models for all the three places considered.

CONCLUSIONS

Based on the studies carried out, it is clear that the same seasonality and cyclicity pattern of rainfall of the three distinct study areas has been maintained over time, i.e., heavy shower primarily during October to December corresponding to Northeast monsoon and secondarily during May to August corresponding to Southwest monsoon. However, the onset and duration of the monsoon as well as the magnitude and variability do show distinct variability in these locations, with increasing sporadic variation between January to June, as reflected by the outliers. The time series analysis of precipitation pattern is stationary in all three locations, with the non-seasonal component of the precipitation pattern obeying first-order moving average, whereas the non-seasonal component does show very distinct variation, as observable in the additional secondarily peak in Chennai and tertiary peak in Thanjavur and varying seasonal terms in ARIMA Thanjavur (2,0,0), Nagapattinam (0,0,1) and Chennai (1,0,0). As far as the suitability of the modeling method is concerned, ETS and ARIMA are found to be the best forecasting models for all three locations, as inferred from the diagnostic matrices.

The rainfall pattern over the years is similar in all three locations in relation to the highest rainfall during November whereas in the case of Chennai during May and August through the decade and only less significant in Thanjavur during May. In the case of Nagapattinam, there are no distinct variations indicating uni-model periodicity throughout the years studied. The time series study of the three locations also reveals that the most suitable models for Thanjavur are ARIMA followed by ETS, whereas for Chennai, the most suitable models are ETS and Holt-Winters and for Nagapattinam, Linear Regression and ETS are the most suitable models. Hence ETS and ARIMA are by far the best models for all the three places considered.

Table 2: Diagnostics for Models Used

| Model	Locations	Mean	SE	Additional Diagnostics
Holt-Winter	Thanjavur	42.95	3.2145	-
	Nagapattinam	63.45	18.4235	-
	Chennai	58.51	14.2541	-
Linear	Thanjavur	44.62	29.5967	-
	Nagapattinam	72.84	15.6951	-
	Chennai	59.5	19.2205	-
ETS	Thanjavur	61.30	4.3652	$\sigma^2 = 6721$, AIC=1599.51, AICC= 1604.13, BIC=1641.32
	Nagapattinam	60.1178	16.3915	$\sigma^2 = 3621$, AIC=1431.087, AICC=1435.70, BIC=1472.90
	Chennai	43.2	11.2395	$\sigma^2 = 3256$, AIC=1236.02, AICC=1568.65, BIC=1352.91
ARIMA	Thanjavur	46.83	0.2507	$\sigma^2 = 4594$, AIC=1235.08 AICC=1235.4, BIC=1245.8
	Nagapattinam	70.54	20.2281	$\sigma^2 = 5671$, AIC=1387.63 AICC=1387.28, BIC=1398.43
	Chennai	61.7962	16.9503	$\sigma^2 = 3810$, AIC=1341.2, AICC=1341.73, BIC=1355.14
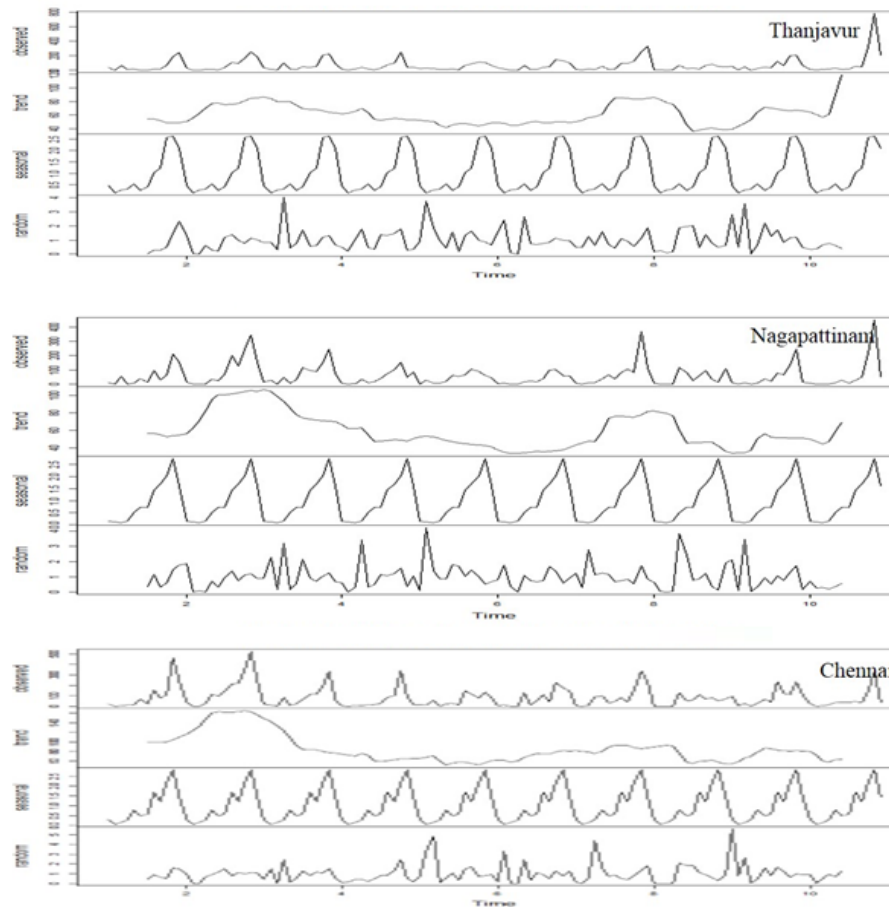
Fig. 6: Decomposition of Multiplicative Series for the Study Area during 2009 to 2018.

ACKNOWLEDGEMENTS

REFERENCES

Akaike, H. 1974. A new look at the statistical model identification. IEEE Trans. Automat. Contr., 19(6): 716-723. doi:10.1109/TAC.1974.1100705, MR 0423716.

Banze, F., Guo, J. and Xiaotao, S. 2018. Impact of climate change on precipitation in Zambeze river basin in Southern Africa. Nat. Env. and Pol. Tech., 17(4): 1093-1103.

Burnham, K.P. and Anderson, D.R. 2002. Model Selection and Multi-model Inference: A Practical Information-Theoretic Approach (2nd ed.). Springer-Verlag, Berlin/Heidelberg, Germany.

Jakhar, P., Gowda, H.C.H., Naik, B.S. and Barman, D. 2011. Probability analysis of rainfall characteristics of Semiliguda in Koraput, Orissa. Indian. J. Soil Cons., 39(1): 9-13.

Jayawardene, H.K.W.I., Sonnadar, D.U.J. and Jayawardene, D.R., 2005. Trends of rainfall in Sri Lanka over the last century. Sri Lanka J. Phy., 6: 7-17.

Kumar, C.P. and Singh, S. 2011. Impact of climatic on water resources and national water mission of India. In: National seminar on global warming and its impact on water resources held on 14th January, 2011 at Kolkata (West Bengal).

Kumar, V. and Jain, S.K. 2010. Rainfall trends in Ganga-Brahmputra-Meghna river basins of India. J. Hydrol., 33: 59-66.

Manikandan, M. and Tamilmani, D. 2012. Statistical analysis of special pattern of rainfall trends in Parambikualam Aliyar subbasin, Tamilnadu. Indian. Water Resour. Soc., 32(1-2): 40-49.

Manivannan, S., Khola, O.P.S. and Dinesh, D. 2016. Probability analysis of weekly rainfall for crop planning in Nilgiris hills of Tamil Nadu. J. Agromet., 18(1): 163-164.

Mohanty, S., Marathe, R.A. and Singh, S. 2000. Probability models for prediction of annual maximum daily rainfall for Nagpur. J. Soil Water Cons., 44(1-2): 38-40.

Obot, N.I., Chendo, M.A.C., Udo, S.O. and Ewona, I.O. 2010. Evaluation of rainfall trends in Nigeria for 30 years. Int. J. Phys. Sci., 5(14): 2217-2222.

Parta, T. and Kahya, E. 2006. Trend analysis in Turkish precipitation data. Hydrol. Process, 20: 2011-2026.

Sharma, K.K. and Dubey, S.K. 2013. Probability analysis of rainfall for planning water harvesting and irrigation in semi-arid region of Uttar Pradesh. Indian. J. of Soil Cons., 41(1): 14-19.

Sharma, K.K., Singh, A.K. and Dubey, S.K. 2015. Analysis of one-day probable maximum precipitation for designing soil and water conservation structures in Agra, U. P. J. Agromet., 17(2): 268-270.

Sharma, K.K., Singh, A.K., and Dubey, S.K. 2018. Rainfall trend analysis for crop planning under rainfed conditions in district Agra of Uttar Pradesh. Mausam, 69(4): 599-606.

Singh, S. and Kumar, S. 2016. Trend analysis of rainfall of Sagar district, Madhya Pradesh. Indian. J. of Soil Cons., 44(1): 44-49.