Identification of Homogeneous Zones in Kerala Through K-Means Clustering

Zeta Paul, Keerthana A and Archana Nair

Department of Mathematics, Amrita Vishwa Vidyapeetham, Kochi, India, Pincode 682024

zeta.aquinas@gmail.com

*archananair@asas.kh.amrita.edu

Article Info
Page Number: 413 – 433
Publication Issue:
Vol. 71 No. 3s (2022)

Abstract

Rainfall is a major source of water to flora and fauna. The variations in rainfall greatly affect agriculture and the daily life of human beings. Surface temperature is another factor that influences and brings changes in the climatic conditions, which in turn affects the environment as a whole. Over the past few years, the global mean surface temperature follows a steady increasing trend, while the trend of Kerala monsoon and post monsoon rainfall remains unpredictable. These variables has been analyzed and studied to determine the changes and future threats. The temperature and monsoon rainfall data has been extracted for a total of 58 grid points $(0.25° \times 0.25°)$ all over Kerala for the years 1951 to 2019 (69 years). k- means clustering method is applied to form clusters or groups with similar traits. Clustering assisted to regionalize Kerala State into places of homogeneous characteristics with respect to rainfall, temperature and its combinations. The optimal cluster number is obtained by Elbow's method, which is found to be 4. The cluster divisions is 10-22-8-18 for rainfall and 9-16-18-15 for temperature. The cluster results for the combination of rainfall and temperature concluded to be same as that of rainfall alone, implying the greater importance of rainfall in delineating the zones.

Article History
Article Received: 22 April 2022
Revised: 10 May 2022
Accepted: 15 June 2022
Publication: 19 July 2022

Keywords: Clustering, Elbow Method, k-means Clustering, Rainfall, Regionalization of Kerala, Temperature

1. INTRODUCTION:

Kerala is located in the South-Western part of India. The coastal state is also known as God's own country due to its natural beauty, greeneries, rivers, beaches, backwaters and most importantly balanced climate throughout the year, which is neither too hot nor too cold. Kerala contributes largely to the Indian economy, agriculture, tourism and education. The state is basically having an agrarian economy with most of export income coming from spices, seafood and cash crops, which heavily depends on the regional temperature and rainfall received. Since 1970, many international researchers have held up Kerala as the model of development (Ratcliffe, 1978). The mean annual rainfall in Kerala over the years 1901 to 2019 ranges from 870 mm to 3061 mm, with an average of 1792 mm. From the literatures, it is observed that

Kerala has experienced a significant climate change after 1980s. Studies have explored monthly, seasonal and annual rainfall trends in Kerala over different time periods. The time series analysis of monthly and seasonal rainfall in Kerala from 1871 to 1994 found that the northeast monsoon over Kerala has an increasing trend in October and November, while a decreasing trend was observed in December (Krishnakumar et al., 2008). Soman et al. (1988) statistically examined the rainfall time series of Kerala for the period 1901-1980 and observed a decreasing trend in the mean annual rainfall, south west monsoon and north east monsoon. The decreasing trend was found to be significant for the highland areas of Kerala, whereas the coastal regions did not show any significant trend. Nair et al. (2014) had studied the rainfall trends in Kerala on a district level, for each month on each season over the period of 100 years from 1901 to 2000. The maximum monsoon rainfall is obtained in the northern districts of Kerala whereas the winter, pre monsoon and post monsoon average rainfall is maximum over the southern districts of Kerala. The study also supported the decreasing trend in both annual and seasonal rainfall of Kerala. The trend analysis showed that the rainfall distribution in districts like Kottayam, Pathanamthitta, Idukki, Thrissur and Wayanad is asymmetric with change in its distribution and hence more detailed study is required to characterize the rainfall pattern in Kerala. During recent years, Kerala has recorded a significant increase in climatological disasters. Increase in regional atmospheric entropy is being seen as the cause. In August 2018, the state received an extended period of very heavy rainfall as a result of a low-pressure system near the beginning of the month, followed several days later by a monsoon depression (Hunt & Menon, 2020).

The temperatures of Kerala were also studied for varied time scales. Jain & Kumar (2012) had observed that the annual mean temperature, maximum temperature in all seasons, global mean annual and seasonal temperatures all had an increasing trend during the years 1901 to 2007, with an accelerated warming during 1971 to 2007. The sea surface temperature (SST) gradient across the equator over the global oceans (Indian and Pacific) is found to have a high and statistically significant linear correlation coefficient with the date of monsoon onset over Kerala (Preenu et al., 2017). Dash et al., (2007) had observed that not only the sea surface temperatures of Arabian sea and Bay of Bengal shows an increasing trend, the atmospheric surface temperature in India has enhanced by about 1°C and 1.1°C during winter and post-monsoon season, respectively. Moreover in south India, the minimum temperature has a steady increase. Kothawale et al., (2005) studied the temperature data of the years 1901–2003 and reported that all India mean annual temperature shows a significant warming trend of 0.05°C per 10 year during the period 1901–2003. The recent period 1971–2003 in the study has seen a relatively accelerated warming of 0.22 °C per 10 year, which is largely due to global warming.

There was a major shift in rainfall pattern during recent years as the seasonal rainfall during southwest monsoon was declining, where as post-monsoon season had an increasing trend. The drastic change in biophysical resources of Kerala State due to man-made interventions is one of the reasons for the decrease in monsoon rainfall (Krishnakumar et al., 2009). The increase in the frequency of tropical cyclones during the post-monsoon season contributes for the increase in post-monsoon rainfall. A detailed study on the rainfall and temperature at maximum number of location points at Kerala will be helpful for predicting the rainfall and extreme weather conditions. These in turn, benefits to predict the natural calamities and disasters.

Clustering method helps to categorize the data into groups based on similar traits. Several researchers have employed various techniques for regionalising and clustering Kerala State based on rainfall and temperature. Simon & Mohankumar (2004) have employed a technique of factor analysis for grouping the rainfall stations of Kerala into 3 clusters. The method of Principal Component analysis has been applied by Gadgil & Narayana Iyengar (1980) for rainfall clustering of Indian Peninsula. Shirin et al. (2016) used Hierarchical Clustering - Ward's Method and grouped the Kerala region into 7 clusters according to the monthly rainfall data. Parmar & Saket (2017) compared different clustering algorithms used for weather data and proposes a K- medoids clustering algorithm to enhance the accuracy and efficiency in weather forecasting. Varsha and Pai (2018) performed the prognosticate Kerala monsoon rainfall using FRBCS Fuzzy Rule Based Classification. Kulkarni (2017) applied k-means clustering algorithm to probability density function of daily rainfall and divided the Indian landmass in to 5 clusters of homogeneous rainfall characteristics. Yovan Felix et al. (2019) compares fuzzy logic algorithm, Neuro Fuzzy genetic algorithm and k-means clustering algorithm and found that the algorithm has more accuracy (98.1%) compared to the others. They have stated that the proposed k-means clustering algorithm is very useful for the users as it is less time consuming with accurate and effective output.

In this study, the k means clustering method is applied on the Kerala grid point data sets of rainfall and temperature, taken separately as well as together for 68 years.The objective is to group the location points into clusters of similar characteristics. This gives an illustration of how the rainfall and temperature pattern has changed over these years. The manuscript is divided into five sections. Section 1 deals with introduction, which gives an insight into the relevance of the study and reviews of the researches conducted so far. Section 2 describes the data and methodology. This section gives an overview of the geographical features of Kerala, source of the data and clustering methodology. Section 3 explains the result obtained. Section 4 concludes and summarise the study, give suggestions and scope of the present study.

2. DATA AND METHODOLOGY:

2.1 Study Area:

The selected study area is Kerala, a state in the south-western part of India. The land surface area is38,863km^2, which is 1.18 % of the Indian landmass. Kerala is situated between the Arabian Sea to the west and the Western Ghats to the east. Kerala can be divided into three climatically distinct geographical regions (i) Highlands (ii) Midlands and (iii) Lowlands. The High land slopes down from the Western Ghats, which rise to an average height of 900 m with a number of peaks over 1,800 m in height. The midlands lying between the mountains and lowlands are made up of undulating hills and valleys. The lowlands or the coastal area is made of the river deltas, backwaters and the shore of Arabian Sea. Kerala is a land of rivers and backwaters, having forty-four rivers (41 to the west and 3 to the east) that cut across with innumerable number of tributaries and branches. These rivers are comparatively small and are entirely monsoon fed. Most of them turn into rivulet in summer, especially in the upper areas.

Located at the extreme southern tip of the Indian subcontinent, Kerala lies near the centre of the Indian tectonic plate(the Indian Plate) and is subject to comparatively little seismic or volcanic activity. The topography consists of a hot and wet coastal plain, gradually rising in elevation to the high hills and mountains of the Western Ghats. Western Ghats were identified as a biological hotspot by Myers et al., (2000).Kerala lies between north latitudes 8° 17' 30" N to 12° 47' 40" N and east longitudes 74° 27' 47" E to 77° 37' 12" E. The climate is mainly wet and maritime tropical, which is influenced by the heavy rains of monsoon. A total of 58 grid points inside Kerala state were selected as the location points, shown in Figure 1. The considered grid longitude varies from 75 to 77.25 and latitude varies from 8.5 to 12.75.

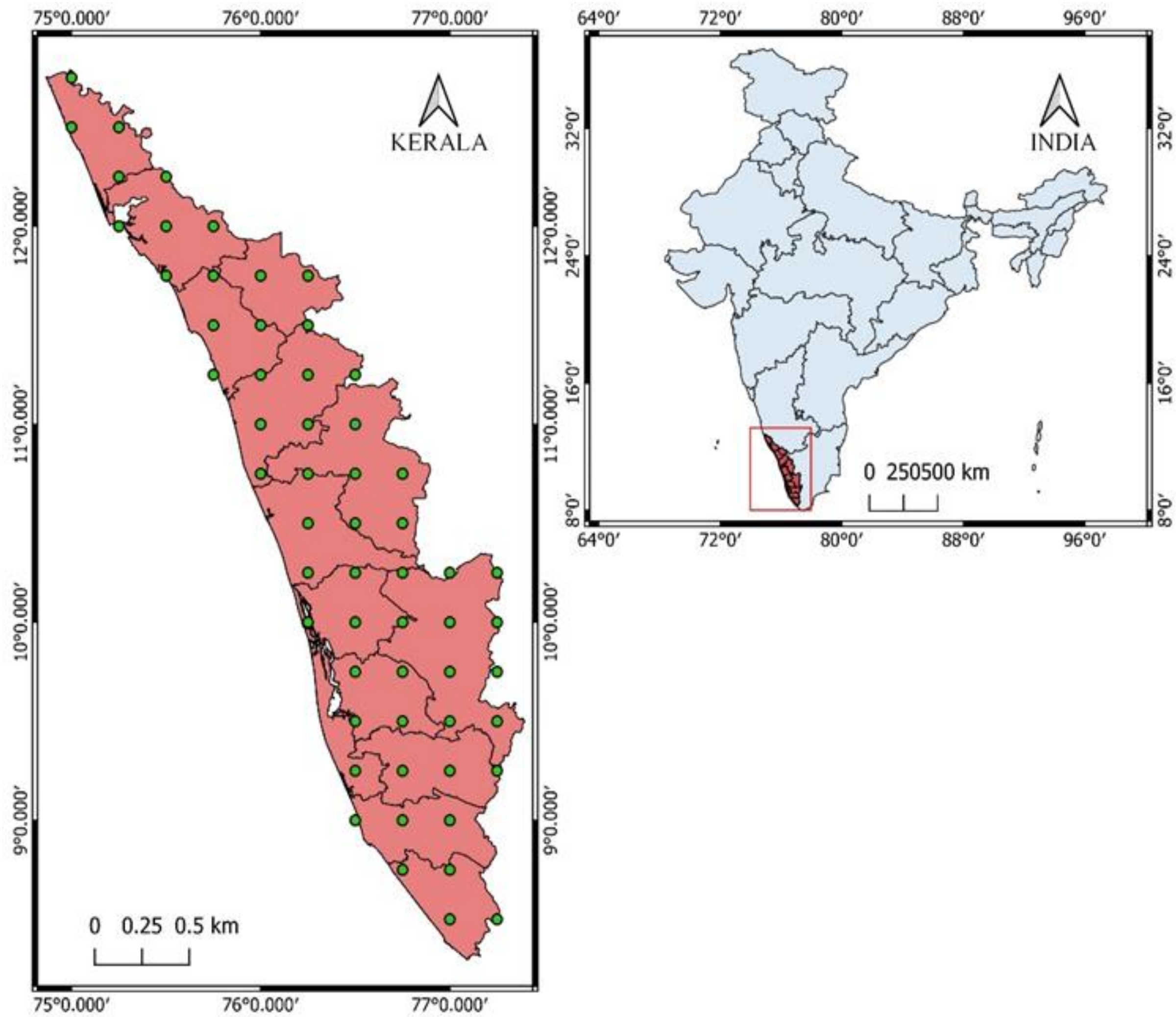


Fig 1. Study Area Map with considered grid point locations

2.2 Data:

The parameters rainfall and temperature of Kerala for the years 1951 to 2019 were downloaded from the website of Indian Meteorological Department (https://www.imdpune.gov.in). To solely analyze the precipitation, rainfall data is taken for monsoon months June to October. The yearly average rainfall in the monsoon months is calculated at each of the considered grid points. The annual average temperature data for each year is also calculated at 58 grid points.

Monthly rainfall and temperature data were also examined to explore the seasonal trends. The annual average of both parameters at each grid point is examined for a pattern in that place. While the parameters individually form two different data sets, a third data set as a combination of rainfall and temperature is also taken for the study. This allows addressing the combined effect of rainfall and temperature. That is, clustering is performed on the three data sets rainfall, temperature and rainfall-temperature.

2.3 Methodology:

Clustering is a common technique in data analysis, which gives an idea about the structure and distribution of the data. Clustering can be defined as a process of grouping a set of objects or data in such a way that the objects or datapoints in the same group are more similar to each other, than to those in the other groups. This is an exploratory data mining process and a common statistical data analysis technique used in many fields. The various clustering methods include Hierarchical Clustering, k-means clustering and Fuzzy C-means clustering method. In this study, k-means clustering is used to form clusters with respect to the variables rainfall and temperature.

k-means clustering is used to partition the dataset into k pre-defined distinct non-overlapping subgroups (clusters), where each data point belongs to only one group. The method tries to make the intra-cluster data points as similar (near) as possible while keeping the clusters as different (far) as possible. The goal of k-means is to group data points into k distinct non-overlapping subgroups. Dabbura (2018) proposed a k means clustering method which assigns data points to a cluster such that sum of the squared distance between the data points and the cluster's centroid (arithmetic mean of all the data points that belong to that cluster) is at the minimum. Lesser the variation within clusters, the more homogeneous (similar) are the data points within the same cluster. That is, the objective function is $J = \sum_{i=1}^{m} \sum_{k=1}^{K} w_{ik} \left\| x^i - \mu_k \right\|^2$, where $w_{ik} = 1$ for data point x^i, if it belongs to cluster k; otherwise $w_{ik} = 0$ and μ_k is the centroid of x^i's cluster. This is similar to a minimization problem with two stages. Firstly, minimize J w.r.t. w_{ik}keeping $\mu_k$fixed and update cluster assignments (E step).

$$\frac{\partial J}{\partial w_{ik}} = \sum_{i=1}^{m} \sum_{k=1}^{K} \left\| x^i - \mu_k \right\|^2 \text{ implies } w_{ik} = \begin{cases} 1 & \text{if } k = \operatorname{argmin}_j \left\| x^i - \mu_j \right\|^2 \\ 0 & \text{otherwise} \end{cases}$$

That is, assign the data point x^i to the closest cluster judged by its sum of squared distance from cluster's centroid. In the second part, differentiate J w.r.t. μ_k and re-compute the centroids after the cluster assignments from previous step (M step).

$\frac{\partial J}{\partial \mu_k} = 2 \sum_{i=1}^{m} w_{ik} (x^i - \mu_k) = 0$ implies $\mu_k = \frac{\sum_{i=1}^{m} w_{ik} x^i}{\sum_{i=1}^{m} w_{ik}}$, which translates in to re-computing the centroid of each cluster to reflect the new assignments.

k-means clustering algorithm requires k as an input and it may not be obvious from the data. Various methods like Elbow method, Silhouette analysis and gap statistic are used to find the optimal number of clusters. In this study, Elbow method is used to fix the number of clusters k. Elbow method is based on the sum of squared distance (SSE) between data points and their assigned cluster's centroids. The number k is picked at the spot where SSE starts to flatten out forming an elbow shape. The steps involved in k-means algorithm are as follows:

1. Specify number of clusters k.
2. Initialize centroids by shuffling and randomly selecting k data points without replacement.
3. Keep iterating until there is no change to the centroids. That is, the assignment of data points to clusters are not changing.
4. Compute the sum of the squared distance between data points and all centroids.
5. Assign each data point to the closest cluster (centroid).
6. Compute the centroids for the clusters by taking the average of all data points that belong to each cluster.

The k-means approach to solve a problem is called Expectation-Maximization. The E-step is assigning the data points to the closest cluster and the M-step is computing the centroid of each cluster.

3. RESULT AND DISCUSSION:

3.1 General Scenario:

The general trend of rainfall and temperature at each of the 58 gridded points of Kerala is being analyzed. Kerala State has an average annual precipitation of about 3000 mm. The rainfall is controlled by the South-west and North-east monsoons. About 90% of the rainfall occurs during monsoon months. The high intensity storms prevailing during the monsoon months lead to heavy discharges in all the rivers. The rainfall has a huge variation in the data from 2.1 mm to 28.2 mm, whereas the temperature ranges from 23.53°C to 27.77°C. The average rainfall in the monsoon months (June to October) of the years 1951 to 2019 is calculated and plotted. An equal quantile classification of average rainfall and temperature into 4 groups is shown in Figure 2.

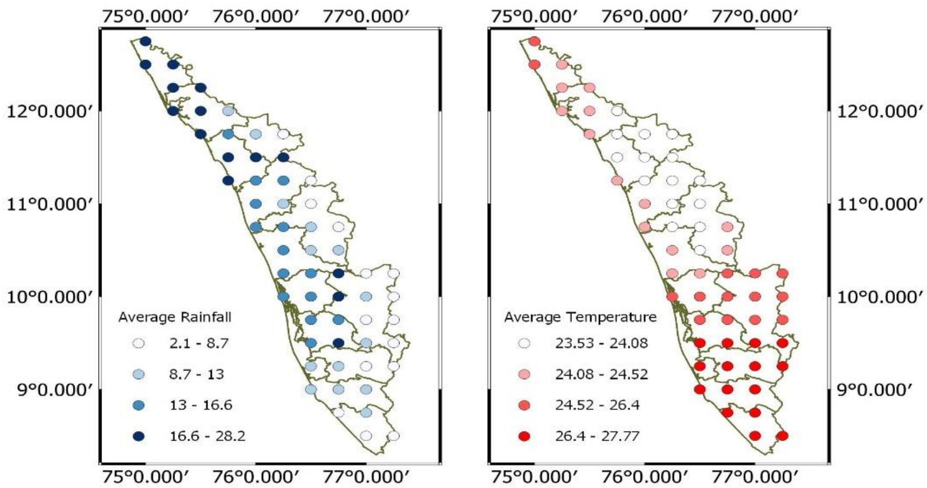


Fig 2. Annual average rainfall and temperature for the years 1951 to 2019

The gridded points which receive daily average rainfall between 2.1 mm to 8.7 mm are grouped as low rainfall areas. A total of 15 gridded points classified under this group are distributed over the eastern border of middle and southern Kerala, covering some parts of the districts Trivandrum, Pathanamthitta, Idukki, Palakkad and Wayanad. The second group named as normal rainfall areas receives 8.7mm to 13mm daily average rainfall. There were 14 gridded points in normal rainfall class, mostly distributed over the districts Kollam, Pathanamthitta, Alappuzha, Idukki, Palakkad, Wayanad and some parts of Kottayam, Malappuram and Kannur. The third group named as heavy rainfall areas have an average daily rainfall range 13 mm to 16.6 mm. These areas include 14 gridded points distributed over the middle Kerala covering the districts Kozhikode, Malappuram, Thrissur, Ernakulam, Kottayam and northern part of Alappuzha. The rest 15 gridded points of Kerala receive rainfall greater than 16.6mm up to a daily average 28.2mm. As a result, these wells may be named as extreme rainfall areas, mainly concentrated on the northern end of Kerala and some regions in the middle Kerala. These grid points cover the entire Kasaragod district, majority of Kannur districts and few parts of the districts Kozhikode, Idukki and Kottayam.

The average temperature of Kerala stay within a narrow range from 23.53°C to 27.77°C of width 4.24°C. The places with average temperature varying from 23.53°C to 24.08°C are classified as low temperature areas, which include 15 gridded points of Kerala mostly covering the middle-east districts Wayanad, Kozhikode, Malappuram, Palakkad and the south-east part of Kannur. The temperature ranging from 24.08°Cto 24.52°C were grouped together to form the second group of 14 gridded points distributed over the western border of the Northern half of Kerala and cover the districts Kasaragod, Kannur, Kozhikode, Malappuram, Palakkad and Thrissur. The 14 gridded points with average temperature between 24.52°Cand 26.4°Cforms the third group. These grid points are at the north end of Kasaragod district and concentrated within latitudes 9.75°C to 10.25°C covering the districts Idukki, Ernakulam, Kottayam and Alappuzha. The places with temperature ranging from 26.4°C to 27.77°C are categorized as highest temperature areas, containing 17 gridded points and are strictly lying on the southern end covering the districts Trivandrum, Kollam, Pathanamthitta and the southern border of the districts Alappuzha, Kottayam and Idukki.

A very few places in Kerala, in particularly three grid points have received significantly very low or no rainfall throughout all the 69 years under study. These three grid point locations are 9° 25' N, 77°25' E;9°75' N, 77° 25' E and 10° 00' N, 77° 25' E, which belong to the eastern part of Idukki and Pathanamthitta districts. The average daily rainfall at these places is less than 5 mm except for 1 or 2 years. The deviation in rainfall from the annual average at each gridded points (major latitudes and longitudes) is also calculated (Figure not shown). This lead to the conclusion that except for a few years, particularly for 1987, 2005, 2018 and 2019, each grid point location of Kerala has followed similar rainfall pattern throughout all the years. In the year 1984, exactly one gridded point 10°N, 77° E showed excess rainfall with 269% greater than the average rainfall which is the fourth highest recorded rainfall during the study period. In 2005, almost all places in Kerala received very low rainfall compared to all the other years in the study period and hence 2005 can be considered as a drought period in Kerala. But, a very few places particularly three gridded points received excess rainfall in 2005 which are the highest recorded daily average rainfall during the period 1951-2019. The gridded point 12° N, 75° 5' E received 361% of the average rainfall in 2005, the gridded point 12° 25' N, 75° 5' E received 345% of the average rainfall in 2005 and the gridded point 12° 5' N, 75° 25' E received 278% of the average rainfall in 2005.

3.2 Month-wise and Year-wise Changes in Rainfall and Temperature:

The month-wise average rainfall over the years 1951-2019 is shown in Table 1. The rainfall for the seven months from May to November contributes 92% of annual rainfall. As per the rainfall data of places in Kerala during the years 1951-2019, the average rainfall in Kerala is highest in July month with 587.84 mm, which is 23.25% contribution to the annual rainfall. In addition, July month rainfall has the highest standard deviation. The Southeast monsoon starts in June and the rainfall increases and reaches to its highest in July. The lowest rainfall is observed in January month with 10.54 mm and has the lowest standard deviation of 12.46. The standard deviation is observed to be directly proportional to the average rainfall except for the months May and September, whereas significantly high variance in the amount of rainfall received is seen. It is observed that there is an increasing trend in rainfall from January to July, as the season moves from winter to summer to south west monsoon. Once the south west monsoon rainfall starts, a decreasing trend follows from July to September. With the onset of north east monsoon, October receives higher rainfall than September and the following months. As a result, a decreasing trend follows from October to December. It seems to be interesting that in Kerala, rainfall during the winter season January-February is lower than the rainfall during the summer March-April-May months.

Table 1 Month-wise average rainfall over the years 1951-2019

Month	AverageRainfall (mm)	Standard Deviation	Percentage contribution to annual rainfall (%)
January	10.54	12.46	0.42
February	14.90	15.86	0.59
March	34.72	33.16	1.37
April	98.88	40.85	3.91
May	187.43	108.20	7.41
June	507.74	161.89	20.08
July	587.84	190.73	23.25
August	395.46	127.05	15.64
September	230.61	109.43	9.12
October	265.54	83.99	10.50
November	152.68	83.78	6.04
December	42.23	34.22	1.67

The average temperature in Kerala state ranges from 23.83°C to 27.42°C, with an average of 25.14°C of annual average temperatures during the years 1951 to 2019.The highest temperature is recorded in the month of April due to summer and minimum temperature is in winter (January).The average temperature follows a breaking increase-decrease pattern. Opening with January month, the temperature keeps on increasing till June. Due to the south-west monsoon that starts from June, the temperature is lower than May. As the monsoon continues, a slight increase in temperature is seen until September. The arrival of north-east monsoon brings along a slight decrease in temperatures up to the year end. The standard deviations resulted in higher value 1.57 (write units) in July and lower value 1.09 in April. The deviations of monthly average temperature from the annual average temperature were also calculated. A total of 8 negative deviations and 4 positive deviations were observed. The summer months March, April and May along with the initial monsoon month June have the positive deviations.

Table 2 Month-wise average temperature over the years 1951-2019

Month	Average Temperature (°C)	Standard Deviation	Deviation from the Annual Average Temperature
January	23.83	1.44	-1.32
February	24.97	1.29	-0.17
March	25.93	1.20	0.78
April	27.42	1.09	2.28
May	27.07	1.24	1.93
June	25.28	1.49	0.13
July	24.52	1.57	-0.63
August	24.58	1.52	-0.57
September	24.96	1.42	-0.18
October	24.91	1.24	-0.23
November	24.39	1.24	-0.76
December	23.88	1.42	-1.27

During the period of 69 years from 1951 to 2019, the average daily rainfall in Kerala is 2.4 mm. The anomalies (Xi-X)/SD were plotted for these years which showed 33 years have positive anomaly whereas 36 years have negative anomaly, shown in Figure 3.The highest positive deviation is 3.3, which is observed in the year 1961. The highest negative deviation is -2.6, which is observed in 2005. Over these years, the highest number of continuous negative deviations is 6 for the years 1985-1990 and 2000-2005. The highest number of continuous positive deviations is 9 for the years 1953-1961. There were 13 positive to negative deviations and 14 negative to positive deviations. That is, a greater number of fluctuations are observed in the rainfall pattern over Kerala grid points for the considered years.

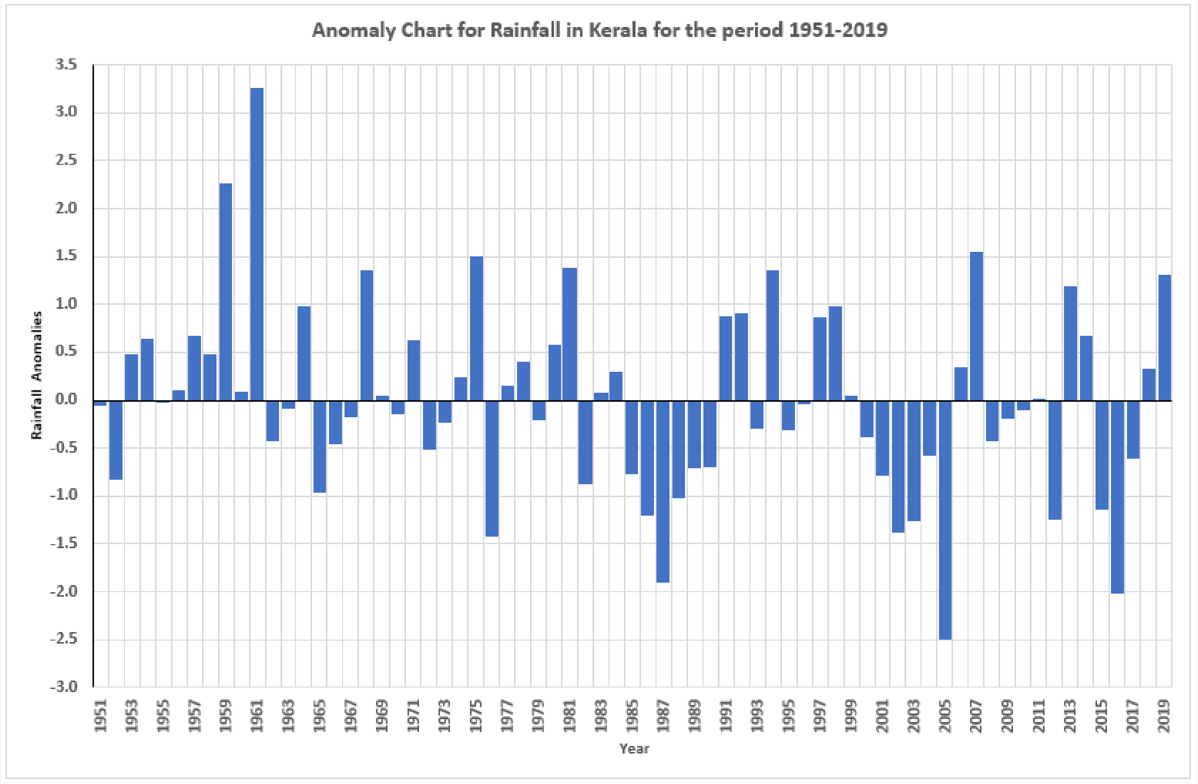


Fig 3. Anomaly Chart for Rainfall in Kerala for the period 1951-2019

During the period of 69 years from 1951 to 2019, the average daily temperature in Kerala is 25.14°C.The average temperature of Kerala has an increasing trend which is evident from the anomaly chart of temperature of Kerala for the years 1951 to 2019, shown in Figure 4.The anomalies (Xi-X)/SD showed that 31 years have positive values whereas 34 years have negative values. The year 1985 can be regarded as the starting point for this change since from 1985 onwards 80% of the years there is increase in temperature from the average and the rest 20% has only very less decrease with anomaly value between 0 and -0.3. The highest positive deviation is 2.1, which is observed in the year 1987. The highest negative deviation is -2.1, which is observed in 1956 and 1971. Over these years, the highest number of continuous negative deviations is 15 obtained for the years 1951-1965, while the highest number of continuous positive deviations are 17, obtained for the years 2001-2017.

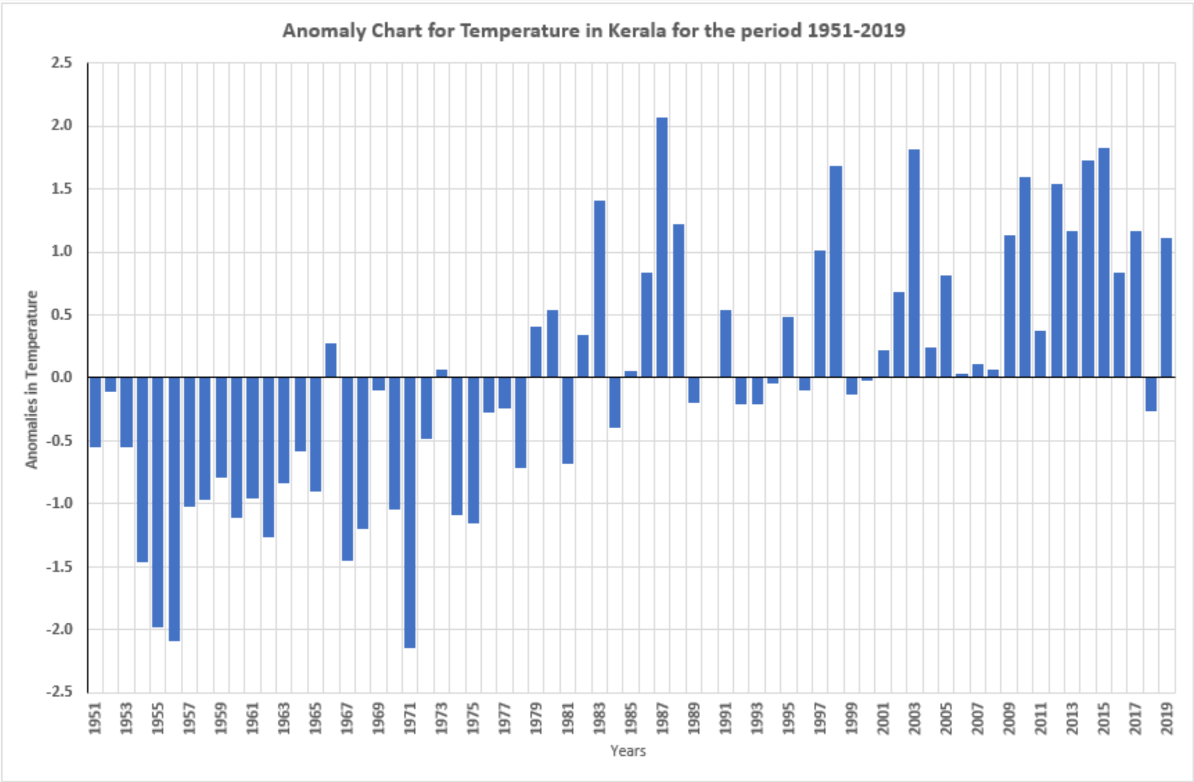


Fig 4. Anomaly Chart for Temperature in Kerala for the period 1951-2019

3.3 Decade-wise Analysis:

At each gridded point, the deviation of average rainfall of a particular decade from that of the previous decade was calculated and are (deviation) plotted, as shown in Figure 5. The gridded points with positive deviation indicate a raise in rainfall in the present decade compared to the past decade and are marked by blue shades. The negative deviation indicates the decrease in rainfall, and is marked by red shades. Significant increase in rainfall was noticed for the deviation between the decades 1981-1990 and 1991-2000. The number of red shaded grid points is 25, 36, 51, 7, 51 and 17, while the numbers of blue shaded grid points are 33, 22, 7, 51, 7 and 41. The initial three decadal deviations showed increase in red shades and decrease in blue shades. That is, comparing each decade from 1951 to 1980, a rise in number of grid points with decreasing trend in rainfall was observed. By the last decade of 20st century, most of the grid points showed increase in rainfall than the previous decade. The inverse scenario followed in the next decade and finally ended with an increase in blue shaded grid points in the last decade.

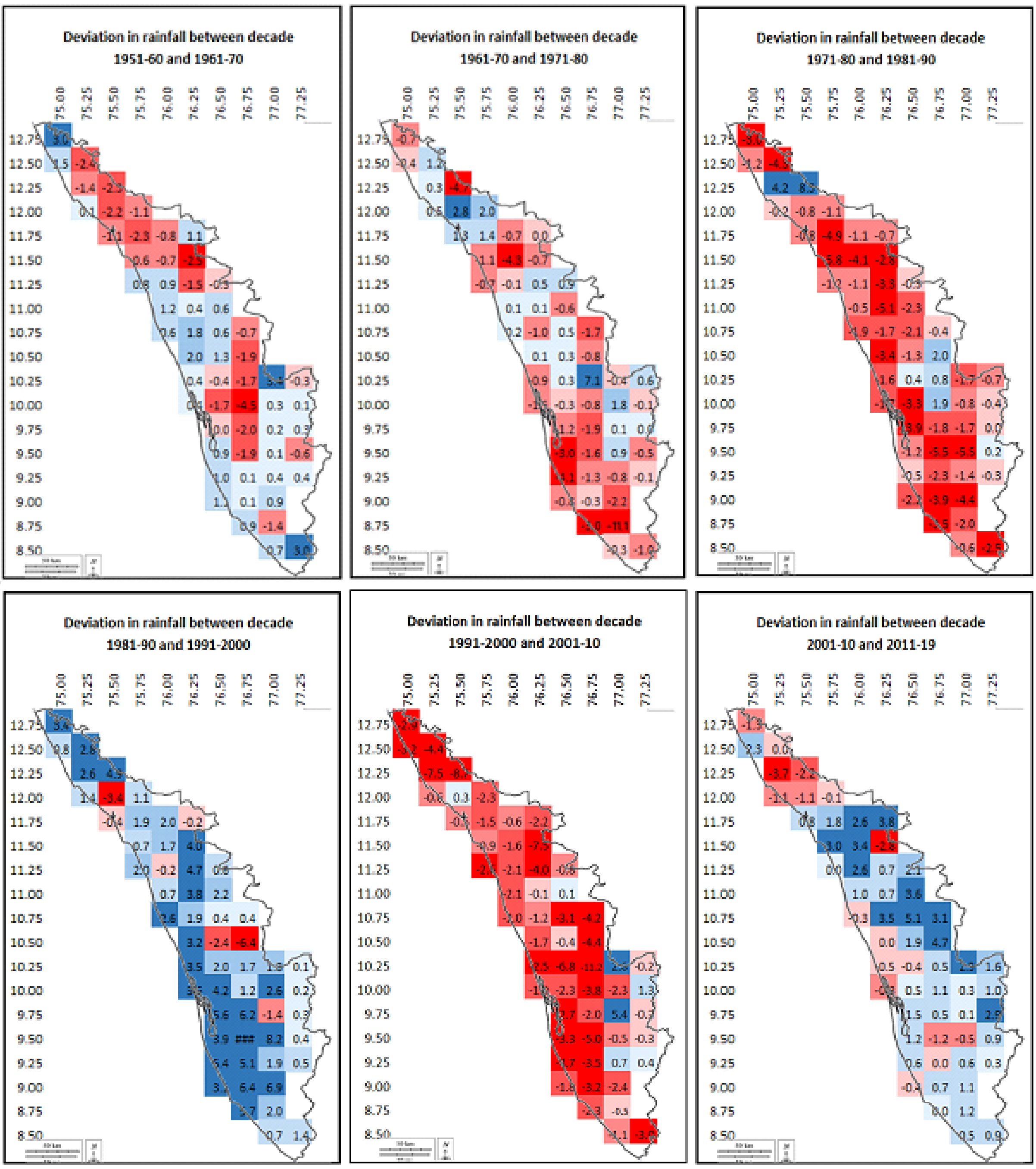


Fig 5. Decade-wise Deviation of Rainfall

Average daily temperature at each gridded point of Kerala was calculated for each decade, shown in Figure 6. At each grid point, the deviation of average temperature of a particular decade from that of the previous decade was calculated. The gridded points with positive deviation indicate a raise in temperature are shades of red with varying intensity. The negative deviation indicates the decrease in temperature and are marked by green shades. Significant increase in temperature was noticed between the decades 1971-1980 and 1981-1990, whereas significant decrease in temperature was observed on comparing the decades 1981-1990 and 1991-2000. The number of red shaded grid points are 58, 58, 58, 17, 56 and 11. The initial three decadal percentage deviation showed increase in temperatures with only red shades over all grid points. By the last decade of 20st century, most of the grid points showed decrease in temperature than the previous decade. The same scenario followed in the next decade and finally ended with more number of green shaded grid points in the last decade.

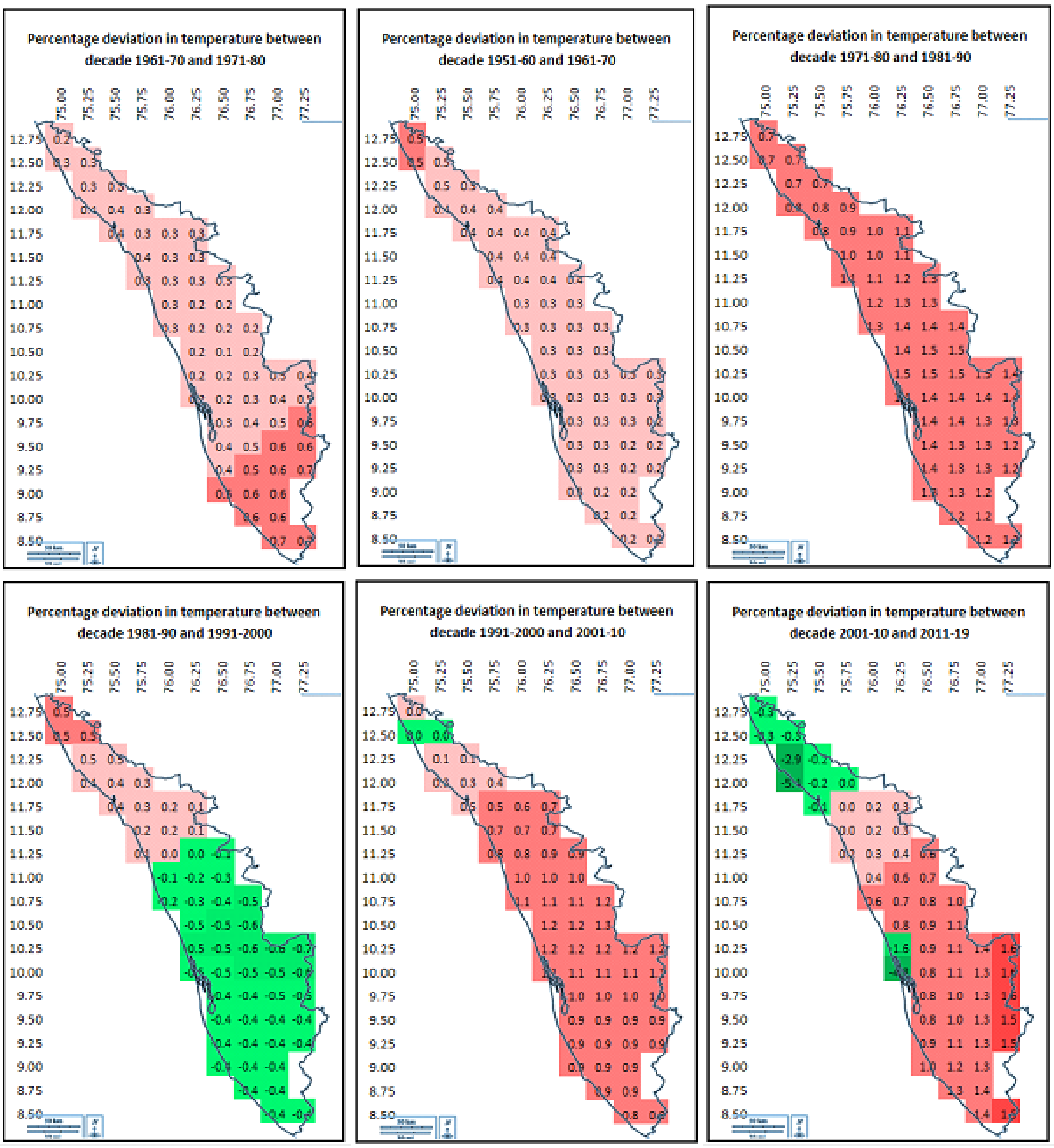
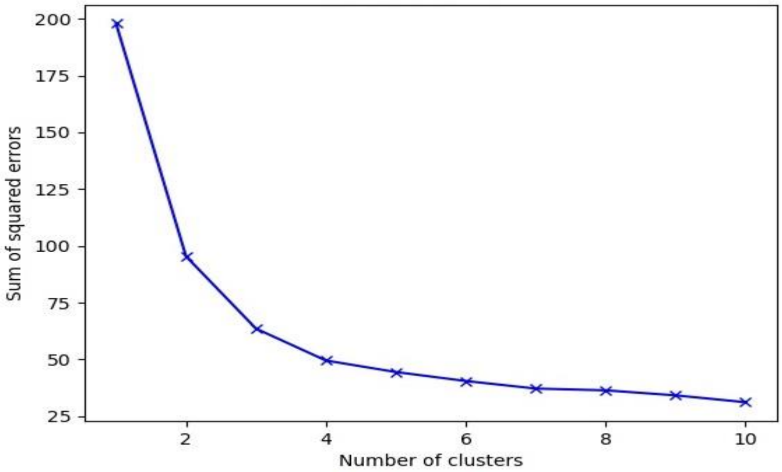


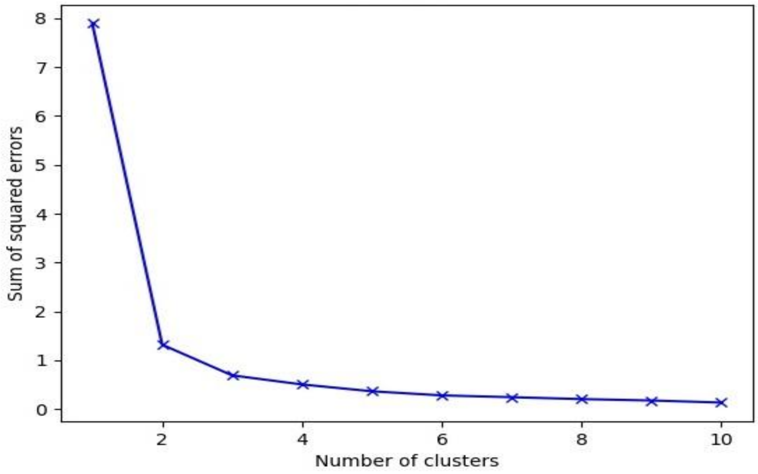
Fig 6. Decade-wise Deviation of Temperature

3.4 k-means clustering:

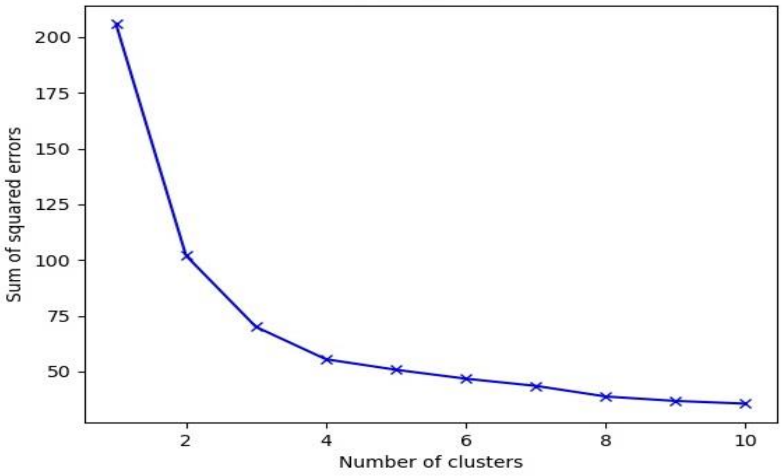
The initial step in k means clustering is to determine the optimal number of clusters, which is obtained using Elbow method. The Elbow charts for rainfall, temperature and rainfall-temperature is shown in Figure 7. The graph is drawn by taking the number of clusters along the horizontal axis and the sum of squared distances from the cluster centroids along the vertical axis. It is observed that, the elbow chart begins to flatten out at the point 4. Thus, the number of clusters is chosen to be k=4.



(a) Rainfall



(b) Temperature



(c) Rainfall-Temperature

Fig 7. Elbow charts for (a) Rainfall (b) Temperature (c) Rainfall-Temperature

The K- means clustering is applied by setting the optimal number of clusters as 4, found using Elbow method. That is, a total of four clusters or groups will be obtained. The K- means clustering divided the rainfall grid points of Kerala into 4 groups, which equivalently divided the rainfall range 2 mm – 28 mm into 4 sub-ranges, shown in Figure 8. The divisions named as, Extreme Rainfall Areas with average daily rainfall ranging from 19 to 28 mm, Heavy Rainfall Areas with average daily rainfall 13 to 18 mm, Moderate Rainfall Areas ranging from 7 to 12 and Low Rainfall Areas with 2 to 6 mm. Majority of the places receive heavy rainfall in Kerala. Out of the 58 gridded points,8 gridded points receive extreme, 22 receives heavy rain, 18 receives average rainfall and 10 receives low rainfall. In cluster analysis, it was observed that the extreme rainfall is occurring strictly on the northern side of Kerala, while Low rainfall areas are distributed over the south-east part of Kerala. Central Kerala receives heavy or moderate rainfall, where heavy rainfall areas are on the western side and moderate rainfall areas are on the eastern side.

The cluster properties of rainfall is shown in Table 3. The minimum daily average rainfall in the red cluster is 19.47 mm and the maximum is 28.18 mm. The average daily rainfall is 22.804 mm and the standard deviation is 3.08 in cluster 1. A total of 8 grid points in this cluster lies in Kasaragod and Kannur districts. The minimum, maximum and average daily rainfall in cluster 2 (blue) is 12.91 mm, 18.39 mm and 15.578 mm, while the standard deviation is 1.593. The 22 grid points in this cluster lies in the 8 districts namely, Kozhikode, Malappuram, Thrissur, Ernakulam, Alappuzha and Kottayam. The third cluster (yellow) covers the greatest number of districts. The maximum and minimum daily rainfall in the third cluster are 7.85 mm and 12.68 mm, with standard deviation of 10.098. A total of 18 grid points in the third cluster covered districts Kannur, Wayanad, Palakkad, Idukki, Pathanamthitta, Alappuzha, Kollam, Trivandrum. The fourth cluster has the lowest average daily rainfall of 4.143 mm and a standard deviation of 1.490. The grid points lie within the districts Wayanad, Malappuram, Idukki, Pathanamthitta and Trivandrum.

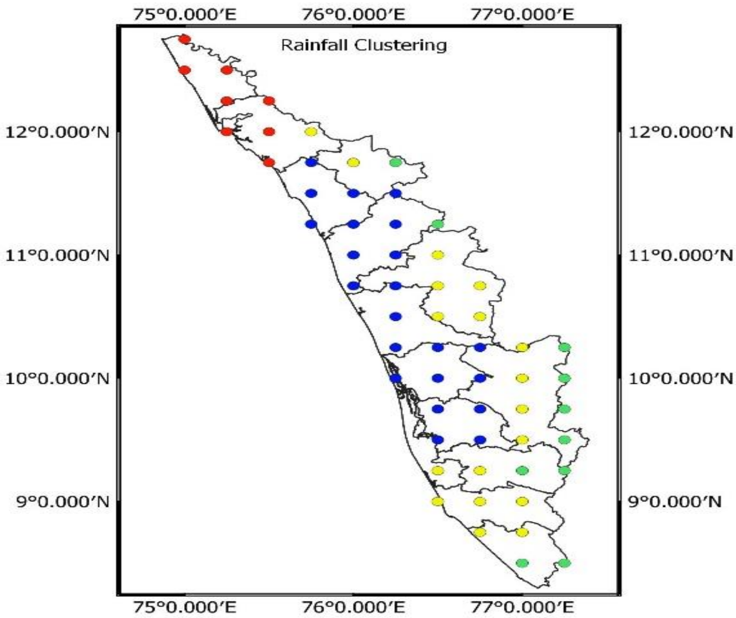


Fig 8. K-means clustering of rainfall in Kerala grid points

Table 3 Cluster properties of rainfall

Rainfall Cluster Analysis	Minimum Average Daily Rainfall (mm)	Maximum Average Daily Rainfall(mm)	Average Daily Rainfall (mm)	Standard Deviation	Districts shared by the cluster	No. of Grid Points
Cluster 1 (Red)	19.47	28.18	22.804	3.08	Kasaragod, Kannur	8
Cluster 2 (Blue)	12.91	18.39	15.578	1.593	Kozhikode, Malappuram, Thrissur, Ernakulam, Alappuzha, Kottayam	22
Cluster 3 (Yellow)	7.85	12.68	10.098	1.737	Kannur, Wayanad, Palakkad, Idukki, Pathanamthitta, Alappuzha, Kollam, Trivandrum	18
Cluster 4 (Green)	2.09	6.20	4.143	1.490	Wayanad, Malappuram, Idukki, Pathanamthitta, Trivandrum	10

The temperature gridded points of Kerala is divided into 4 groups using K-means clustering, shown in Figure 9. The clusters are named as Extreme Temperature Areas with average daily temperature ranging from 26.33°C to 27.77°C, High temperature Areas with ranges from 25.13°C to 26.15°C, Moderate Temperature Areas from 24.1°C to 24.94°C and Low temperature Areas with 23.53°C to24.07°C. It is observed that the extreme temperature areas are strictly on the southern side of Kerala and very low temperature areas are distributed over the south-eastern part of the upper (northern) half of Kerala. The northern part of Kerala experiences a moderate temperature, while the high temperature grid points are just below the central Kerala. That is, starting from south to north Kerala, the average temperature keeps on decreasing; and slight increase is seen in northern most districts.

The cluster properties of temperature is shown in Table 4.The minimum average daily temperature in the Red cluster is 26.33°C and the maximum is 27.77 °C. The average daily temperature is 26.83°C and the standard deviation is 0.38 in cluster 1. A total of 16 grid points in this cluster lies in Idukki, Pathanamthitta, Kottayam, Alappuzha, Kollam, Trivandrumdistricts. The minimum, maximum and average daily temperature in cluster 2 (blue) is 25.13°C, 26.15°C and 25.63°C, while the standard deviation is 0.36. The 9 grid points in this cluster lies in districts Idukki, Ernakulam and Kottayam. The third cluster (yellow) covers the greatest number of districts. The maximum and minimum daily temperature in the third cluster is 24.10°C and 24.94°C, with standard deviation of 0.23. A total of 18 grid points in the third cluster covers9 districts namely, Kasaragod, Kannur, Kozhikode, Malappuram, Palakkad, Thrissur, Idukki Ernakulam and Alappuzha. The fourth cluster has the lowest average daily temperature of 23.87°Cand a standard deviation of 0.15. The grid points lie within the districts Kannur, Kozhikode, Wayanad, Malappuram, Palakkad and Idukki.

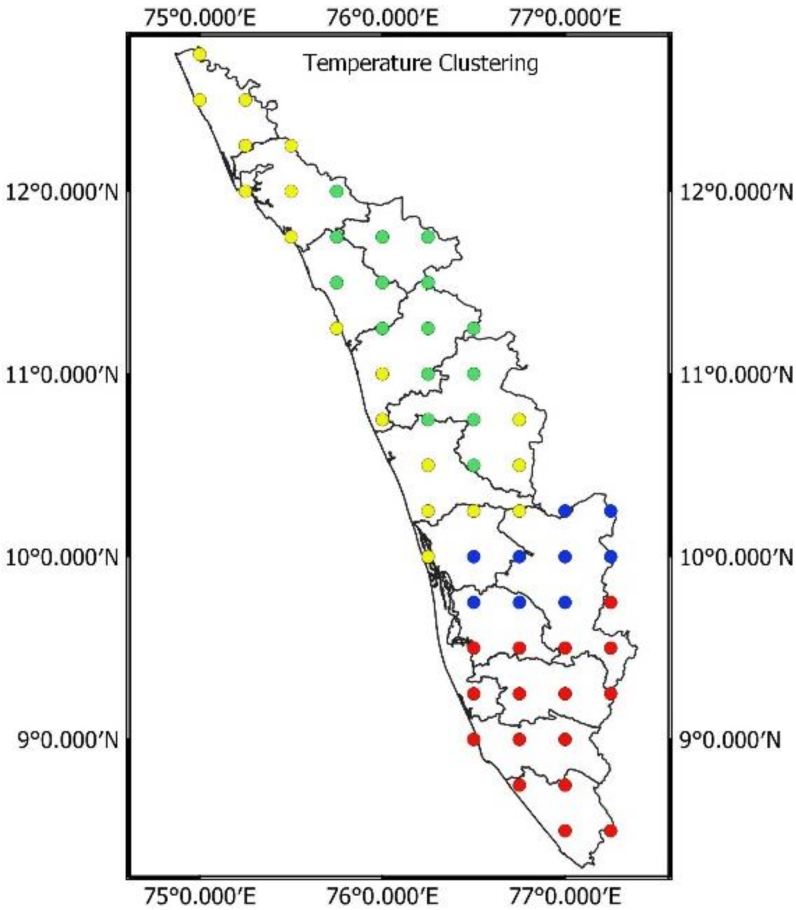


Fig 9. K-means clustering of temperature in Kerala grid points

Table 4 The cluster properties of temperature

Temperature Cluster Analysis	Minimum Daily Average Temperature (°C)	Maximum Daily Average Temperature (°C)	Average Temperature (°C)	Standard Deviation	Districts shared by the cluster	Number of Gridded Points
Cluster 1 (Red)	26.33	27.77	26.83	0.38	Idukki, Pathanamthitta , Kottayam, Alappuzha, Kollam, Trivandrum	16
Cluster 2 (Blue)	25.13	26.15	25.63	0.36	Idukki, Ernakulam, Kottayam	9
Cluster 3 (Yellow)	24.10	24.94	24.38	0.23	Kasaragod, Kannur, Kozhikode,	18

					Malappuram, Palakkad, Thrissur, Idukki Ernakulam, Alappuzha.	
Cluster 4 (Green)	23.53	24.07	23.87	0.15	Kannur, Kozhikode, Wayanad, Malappuram, Palakkad, Idukki	15

A comparative study of rainfall clustering and temperature clustering lead to the conclusion that the rainfall at a particular gridded point is not significant with the temperature at that place, whereas the geography has more impact on the rainfall. For example, the places that receive extreme rainfall has only a moderate temperature, not low. The places with extreme temperature receive only moderate or very low rainfall, not heavy. But, certain gridded points with very low temperature receive very low rainfall. The places with extreme temperature receive moderate or low rainfall.

The K-means clustering is also performed by taking the rainfall and temperature data together at the gridded points of Kerala. The cluster result obtained is same as that of rainfall clustering. This is because of the wide range of the rainfall when compared to the temperature data range. The average daily rainfall at the gridded points of Kerala ranges from 0.1782 mm to 50.2481 mm with variance 51.5348, whereas the average temperature ranges from 21.08°C to 28.55°C with variance1.6182. Since the percentage of deviation from the average is very less for temperature data, it has not made any significant effect on clustering.

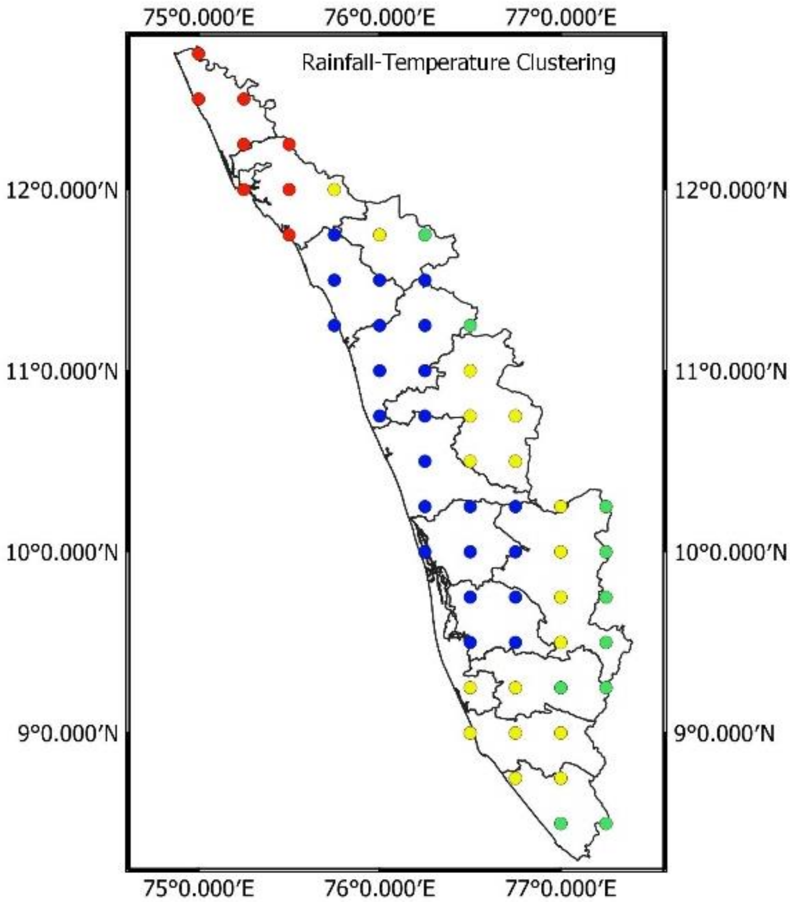


Fig 10. K-means clustering of rainfall-temperature in Kerala grid points

4. CONCLUSIONS:

Kerala state do not suffer large variations in temperature and has a balanced temperature, not too hot and not too cold when compared to the human body temperature. But the rainfall distribution has large variance at certain places and for certain years. It is beneficial for the Kerala state to have detailed rainfall information at the district level for disaster management, preparedness, water resource management and flood control measures, which would in turn would help to solve the environmental and socio-economic issues.

In this study, we have analyzed the general, monthly, annual and decadal trend of rainfall and temperature at the major gridded points of Kerala during the period from 1951 to 2019. The study is mainly based on the temperature and rainfall series at the grid points of Kerala for the last 68 years obtained from India Meteorological Department. The k-means clustering was performed to group the gridded points with homogeneous rainfall and temperature properties. This paper analyses temperature and rainfall changes that the state has witnessed during the last century and the selected location points of the state has been clustered according to temperature, rainfall separately and taken together.

1. Seasonal Trend of Rainfall was analyzed for the period 1951 to 2019 which showed that the seven months from May to November contributes significantly to the annual rainfall, of which July receives the highest rainfall throughout these years.
2. Deviations of rainfall and temperature anomaly showed a steady increase in the temperature profile, whereas the rainfall pattern fluctuates with neither a steady increase nor a steady decrease.
3. Decade-wise analysis of change in rainfall trend at each gridded points showed that there is a significant change in the rainfall pattern of Kerala during the decade 1991-2000 compared to the previous decades, which indicates a climatic change.
4. The deviation of rainfall and temperature at each gridded point from the annual average showed that 48% places receive rainfall more than the average rainfall of Kerala and 48% places has temperature more than the average temperature of Kerala.
5. The analysis of percentage of deviation of rainfall from the annual average at each gridded points of Kerala showed that the rainfall pattern followed by the areas in Kerala is similar for almost all the years except for two or three years, which makes sense for the clustering of homogeneous areas in Kerala.
6. The gridded points of Kerala have been grouped into 4 clusters of homogeneous rainfall characteristics by applying k-means clustering which showed that 14% as extreme rainfall areas and 17% of the places as very low rainfall areas. The extreme rainfall areas are on the northern end, whereas very low rainfall areas are distributed over the south-east part of the lower half of Kerala.
7. When the regions in Kerala have been classified into 4 clusters according to temperature using k-means clustering, it is observed that about 26% of the places in Kerala has high temperature and another 26% of the places has very low temperature. The extreme temperature areas are on the southern side of Kerala, whereas the low temperature areas are distributed over the north-east part of Kerala.
8. When the parameters rainfall and temperature were both taken together for k-means clustering, the clusters obtained were same as the rainfall clusters which means that in the presence of highly variant rainfall, the temperature does not play a significant role in classifying the places in Kerala, as the 1.8 variance in temperature at different grid points in Kerala is low.
9. Analysis of rainfall clusters showed that the northern end of Kerala particularly, Kasaragod district and the north of Kannur district belong to the extreme rainfall areas as per the rainfall data during the period 1951 to 2019.
10. From the temperature clustering it is observed that the southern districts Trivandrum, Kollam and Pathanamthitta; and the southern part of the districts Alappuzha, Kottayam and Idukki are the extreme temperature areas in Kerala.

5. REFERENCES:

1. Ratcliffe, J. (1978). Social justice and the demographic transition: Lessons from India's Kerala State. International Journal of Health Services, 8(1), 123-144.
2. Krishnakumar, K. N., & Prasad Rao, G. S. L. H. V. (2008). Trends and variability in northeast monsoon rainfall over Kerala. J. Agromet, 10(2), 123-126.

3. Soman, M. K., Kumar Krishna, K., & Singh, N. (1988). Decreasing trend in the rainfall of Kerala. Current science, 97(1), 7-12.
4. Nair, A., Joseph, K. A., & Nair, K. S. (2014). Spatio-temporal analysis of rainfall trends over a maritime state (Kerala) of India during the last 100 years. Atmospheric Environment, 88, 123-132.
5. Hunt, K. M., & Menon, A. (2020). The 2018 Kerala floods: a climate change perspective. Climate Dynamics, 54(3), 2433-2446
6. Jain, S. K., & Kumar, V. (2012). Trend analysis of rainfall and temperature data for India. Current Science, 37-49
7. Preenu, P. N., Joseph, P. V., & Dineshkumar, P. K. (2017). Variability of the date of monsoon onset over Kerala (India) of the period 1870–2014 and its relation to sea surface temperature. Journal of Earth System Science, 126(5), 1-19.
8. Dash, S. K., Jenamani, R. K., Kalsi, S. R., & Panda, S. K. (2007). Some evidence of climate change in twentieth-century India. Climatic change, 85(3), 299-321.
9. Kothawale, D. R., & Rupa Kumar, K. (2005). On the recent changes in surface temperature trends over India. Geophysical Research Letters, 32(18).
10. Krishnakumar, K. N., Rao, G. P., & Gopakumar, C. S. (2009). Rainfall trends in twentieth century over Kerala, India. Atmospheric environment, 43(11), 1940-1944.
11. Simon, A., & ohankumar, K. (2004). Spatial variability and rainfall characteristics of Kerala. Journal of Earth System Science, 113(2), 211-221
12. Gadgil, S., & Narayana Iyengar, R. (1980). Cluster analysis of rainfall stations of the Indian peninsula. Quarterly Journal of the Royal Meteorological Society, 106(450), 873-886.
13. Shirin, A. S., & Thomas, R. (2016). Regionalization of rainfall in Kerala state. Procedia Technology, 24, 15-22.
14. Parmar, H., & Saket, S. (2017). Overview of clustering algorithm for weather data. International Journal of Advance Research and Innovative Ideas in Education, 3(6), 1023-1029.
15. Varsha, K.S. and Pai, M.L., 2018. Rainfall prediction using fuzzy C-mean clustering and fuzzy rule-based classification. International Journal of Pure and Applied Mathematics, 119(10), pp.597–605.
16. Kulkarni, A. (2017). Homogeneous clusters over India using probability density function of daily rainfall. Theoretical and Applied Climatology, 129(1), 633-643.
17. Yovan Felix, A., Vinay, G. S. S., & Akhik, G. (2019). K-Means Cluster Using Rainfall and Storm Prediction in Machine Learning Technique. Journal of Computational and Theoretical Nanoscience, 16(8), 3265-3269.
18. Dabbura, I. (2018). K-means clustering: Algorithm, applications, evaluation methods, and drawbacks. Towards Data Science.