Journal of Research in Environmental and Earth Sciences, 10 (2021) 264-278
Print ISSN: 2356-5799; Online ISSN: 2356-5802
© Knowledge Journals
www.knowledgejournals.com

Research article

Identification of relationships between climate indices and precipitation fluctuation in Peshawar City-Pakistan

Bushra Begum [a], Sapna Tajbar [b,*], Banaras Khan [a], Lubna Rafiq [c]

[a] *Department of Physics, Qurtuba University of Science & Information Technology, Peshawar, Pakistan*
[b] *Department of Climatology, University of Tabriz, Iran*
[c] *Geoinformation, University of Salzburg, Austria*
**Corresponding author. Tel. : +989331128113 (sapnatajbar@gmail.com)*

Article history:
Received 20 January 2021; Received in revised form 17 March 2021.
Accepted 17 April 2021: Available online 17 May 2021.

Abstract

The demand for water has increased in the recent decades in all sectors that are domestic, industrial, and agriculture. Climate signals teleconnection are one of the important factors influencing the oscillations of climate on earth (Ruigar and Golian, 2015). The aim of the present study was to identify the relationship between climate indices such as El Nino Southern Oscillation (Nino-1+2, Nino-3, Nino-3.4, Nino-4, Southern Oscillation Index, Multivariate ENSO Index and Sea Surface Temperature), Quasi-Biennial Oscillation and the precipitation in Peshawar City-Pakistan and predict the precipitation. The Tropical Rainfall Measuring Mission satellite precipitation and National Oceanic and Atmospheric Administration climate indices data for the period of 1982 to 2018 were used and Pearson correlations, cross-correlations and random forest model were applied. The results manifested that on monthly basis, the highest significant correlations were observed for SOI during June (0.48), Nino-1+2 during May (0.47), Sea surface temperature during June (0.45), August (0.42), and October (0.46), Nino-3 and Nino-3.4 in November (0.44 and 0.43, respectively), and Nino-4 during December (-0.44). Seasonal analysis showed positive significant correlations for Southern Oscillation Index (0.44), Quasi-Biennial Oscillation (0.34), Multivariate ENSO Index (0.39) and Nino-3 (0.34) indices during summer, autumn, spring, and autumn seasons, respectively. On annual basis, no significant correlations were noticed. In the antecedent correlation analysis, six climate indices have the maximum lagged correlations. During prediction, the model performed well in training period and the predicted precipitation followed the trend, but its performance was low on the extreme precipitation phases and similarly during the test period. The findings of this study are of importance to help policy makers in decisions making and planning for adaptation to the effects of climate change.

Key words: Precipitation, Climate indices, Pearson correlation analysis, Cross-correlation, Random Forest, Peshawar City-Pakistan.

© 2021 Knowledge Journals. All rights reserved.

1. Introduction

The demand for water has increased in the recent decades in all sectors that is domestic, industrial, and agriculture, particularly in the arid and semi-arid regions of the world which are facing significant issues. The interrelationship of climate factors on the ocean, land and atmosphere has been studied by various hydro-climatological scientists. On global and regional scales, the climate signals teleconnection are one of the important factors effecting the oscillations of climate on earth (Ruigar and Golian, 2015). Precipitation alters both spatially and temporally which is considered as universal complex phenomena therefore to analyze association between precipitation and modes of the climate in different parts of the world, researchers have used various modelling procedures (Mekanik et al., 2013). Globally, strong correlation has been perceived between climate indices and precipitation. Amongst the climatic variables El Nino Southern Oscillation (ENSO) is known for its impacts on precipitation in India, North and South America and Australia (Hossain et al., 2015a) also Quasi-Biennial Oscillation (QBO) is familiar for its impacts on precipitation in Western North Pacific, Eastern Africa and India (Seo et al., 2013; Indeje and Semazzi, 2000; Claud and Terray, 2007; Bhalme et al., 1987). ENSO is a natural oscillation in the tropical Pacific Ocean and atmosphere, involving variations in

sea surface temperature, sea-level pressure, atmospheric circulation, and much more. With an irregular cycle of about 2-7 years, it represents the most important mode of inter-annual climate variability (Lee, 2015; Yan et al., 2018). The QBO is a dominant feature of the equatorial lower stratosphere, consisting of an extremely regular oscillation from westerly to easterly winds, with a period of about 28 months; these changes in circulation regimes are due to the downward propagation of bands of alternate zonal winds from the top of the lower stratosphere to the tropopause. The QBO can affect global stratospheric circulation (Mazzarella et al., 2011). Over years, worldwide many studies conducted to analyze the relationship of climate indices (ENSO and QBO) and precipitation. According to a study (Ihara et al., 2007) in India, the association of the state of the equatorial Indian Ocean, ENSO, and the Indian summer monsoon precipitation was observed. A significant negative correlation was found between the zonal wind variations over the equatorial Indian Ocean and Indian summer monsoon precipitation during the period of El Nino while the similar analysis of the correlation of Indian summer monsoon precipitation and SST index did not exhibit significant correlation. Another study for the Western Australia found strong influences of DMI and ENSO indices on the precipitation (Hossain et al., 2015b).

According to a study carried out for the Bangladesh revealed slow increasing trend of All-Bangladesh summer monsoon precipitation and for ENSO index a very slow decreasing rate. There was no relationship observed between the ENSO index and All-Bangladesh summer monsoon precipitation (Ahmed et al., 2016). A study for western North Pacific found that in comparison with the easterly phase, during the westerly QBO, a mid-latitude spring rainband expanding from southeastern China to the east of the Japanese Islands is shifted southward, and thus, the spring precipitation over Korea and Japan represented a notable decline (Seo et al., 2013). In a study of India statistical analysis showed that in January-February, Zonal winds at 15 hPa were found to have the largest and most significant association with Indian monsoon precipitation (Claud, 2007).

A research was carried out over eastern Africa (Uganda, Kenya and Tanzania) and the correlation analysis indicated that the variability of the long-rains over the eastern part of Africa is associated with the QBO index having the percentage of variation as 36 percent (Indeje and Semazzi, 2000). Several more studies were also carried out in this context throughout the world (Ashok and Saji 2007; As-syakur et al. 2014; Hidayat et al. 2016; Sanabria et al. 2018; Rasmusson and Carpenter 1983; Li and Ting 2015; Fan et al. 2019; Ailikun and Yasunari 2001;Bhalme et al. 1987; Claud and Terray 2007; Indeje and Semazzi 2000; Kane 1989; Yahiya et al. 2009) but such studies over the Pakistan are scant (Naheed and Rasul 2011; Bhutto et al. 2009; Mahmood et al. 2006; Syed et al. 2006; Mahmood et al. 2004; Zawar and Zahid 2013; Sarfaraz, 2007; Adnan et al. 2016; Khan 2004; Salma et al. 2012; Sheikh et al. 2010; Iqbal and Athar 2017; Safdar et al. 2019; Akhtar and Athar 2019) and especially in Peshawar city of Khyber Pakhtunkhwa province (study area) no work is done. Therefore, this study is of particular importance to understand the impacts of climate indices (QBO and ENSO) on precipitation in Peshawar city, and hence to predict the precipitation accordingly. Within this framework, Pearson correlation and cross-correlation analysis will be performed between the climate indices (ENSO and QBO) and precipitation and the Random Forest (RF) modeling will be used for the prediction of precipitation.

2. Material and methods

2.1. Study Area

Peshawar is the capital of the Khyber Pakhtunkhwa province of Pakistan and is situated at the north-west end of Pakistan, about 160 km west of the federal capital Islamabad (Figure 1). On its three borders, it is bounded by tribal agencies. This district's total area is 1,257 square km. Its estimated population in 2020 is 2.20 million and 10 years ago (2010) it was 1.54 million which shows a rapid growth of 3.26% (WPR, 2020). It is located between latitude of 33° 44 to 34° 15 north and longitude of 71° 22 to 71° 42 east. Mid-November is the start of winter which lasts till the end of March. Summer is in May to September months. In summer, mean maximum temperature is over 40 °C (104 °F) and the mean minimum temperature is 25 °C (77 °F). In winter, the mean minimum temperature is 4 °C (39 °F) and maximum is 18.35 °C (65.03 °F) (Bokhari, 2015). In both winter and summer season Peshawar receives rainfall. Western disruptions bring about critical winter rainfall, particularly in the months of February and April. Highest winter and summer precipitation are received in the months of March and August respectively. Comparatively, greater average rainfall is observed in winter than in summer (Mehmood et al., 2017).

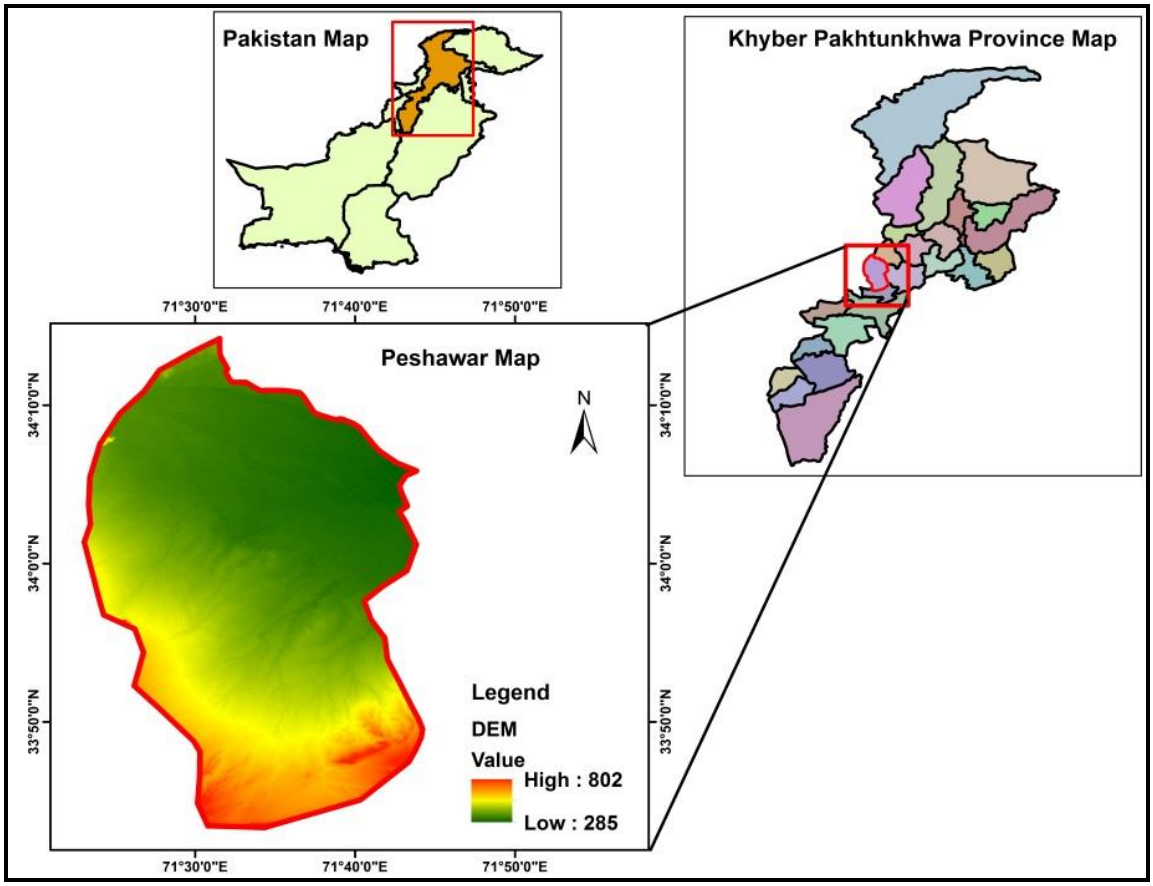

Fig. 1. Study area map of Peshawar

2.2. Data Used and Data Description

2.2.1. TRMM: Precipitation (mm/month)

In this study we used free of cost TRMM precipitation satellite based data sets. TRMM satellite was launched in November, 1997. It is a joint mission between the National Aeronautics and Space Administration (NASA) and Japan Aerospace Exploration Agency (JAXA). It has provided a legacy data set for 17-years of nearly uninterrupted measurements covering the tropics and subtropics (Houze et al., 2015; Hirose et al., 2017). The TRMM satellite records the data using three instruments including Precipitation Radar (PR), TRMM Microwave Image (TMI) and Visible Infrared Scanner (VIRS). Here we used the TRMM monthly precipitation data of 37 years (1982-2018) (Iqbal and Athar, 2017).

2.2.2. NOAA Physical Sciences Laboratory: Climate Indices

ENSO occur in the Pacific Ocean, the two anomalies are the SLP and SST. Its indices included in this study are the Southern Oscillation Index (SOI), Nino-3.4, Nino-3, Nino-4, Nino-1+2, QBO, SST and MEI. SOI is the difference in sea level pressure (SLP) between Darwin and Tahiti in the tropical west Pacific. The West Pacific average sea surface temperature (SST) anomalies are represented by Nino-3.4 covering region from 5°N to 5°S and 170°W to 120°W, respectively. Similarly Nino-3 represents the eastern pacific average sea surface temperature (SST) anomalies covering region from 5ºS to 5ºN and 90 º W to 150 º W while Nino-4 is the Central Pacific average SST anomalies covering area from 5ºN to 5ºS and 150º W to 160ºE (Islam et al. 2017). The Niño-1+2 region is the smallest and eastern-most region of the Nino SST regions from 0ºS to 10ºS and 90ºW to 80oW, and connected with that area of the coastal South America where El Niño was first known by the populations at locality (Trenberth and NCAR Staff, 2020). The Quasi Biennial Oscillation (QBO) index reduces the instability of the equatorial stratosphere (1650km) and is seen as a downward propagation of the eastern and western wind regimes with a variable duration of about 28 months on average (Baldwin et al. 2001). The Sea surface temperature (SST) index represents the temperature of the water closer to the surface of the ocean. MEI combines full information on ENSO i.e. six oceanic and meteorological parameters over the tropical Pacific

(30°S to 30°N and 100°E to 70°W) (Mahmood et al. 2006). All of these climate indices were taken from the NOAA Physical Sciences Laboratory (https://psl.noaa.gov/data/climateindices/list/) except sea surface temperature data. The NOAA monthly Optimum Interpolation (OI) Sea Surface Temperature (SST) version 2.0 data covering 89.5°N-89.5°S and 0.5°E-359.5°E (https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.html) was used for the purpose. All indices data used were for the period of 1982-2018.

2.3. Methodology

Along with simple statistical techniques like Mean and Standard Deviation, the present study investigated the concurrent and antecedent relationship between climate indices and precipitation; Pearson correlation analysis and cross correlation analysis were used in this regard. The formulas of mean and standard deviation are shown below:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \quad (1)$$

$$Std.Dev = \sqrt{\frac{\sum_{i=1}^{n} (\bar{x} - x_i)^2}{n-1}} \quad (2)$$

Where n is the sample size, $\bar{x}$ is the mean of the precipitation and x_i represents the precipitation values (Lee et al. 2015).

To show the concurrent correlations between precipitation and climate indices, the Pearson's product Moment Correlation coefficient was utilized which measures the strength and direction of the link that exists between the two variables (Obilor and Amadi, 2018). The value of R ranges between -1, 0 and 1. On the basis of this statement, R can be either positive, negative or zero. The positive value of R between 0 and 1 indicates positive correlation which means when one variable increases the other also increases and if the value of R is between 0 and -1 then when one variable increases the other decreases. The 0 value of R shows no correlation. Mathematically it can be represented as:

$$R = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}} \quad (3)$$

Where x_i stands for the precipitation data and y_i for the climate indices data (ENSO and QBO), respectively, $\bar{x}$ and $\bar{y}$ are the averages of the precipitation and climate indices respectively.

To show the antecedent correlation between the climate indices and precipitation, we applied cross-correlation technique. In aligning two time series, the cross-correlation is useful, in which one of two time series is delayed with respect to the other, since its peak occurs at the lag at which they are correlated in a best way i.e. the lag at which they match up best (Menke and Menke, 2012). In this study, monthly cross-correlations were calculated between the climate indices and precipitation for the range up to 12 months. In this way the best associations were found at different lags. The formula is:

$$r_m = \frac{\sum (x_i - \bar{x})(y_i - m - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - m - \bar{y})^2}} \quad (4)$$

Whereas r_m represents the m lag time (Taweesin and Seeboonruang, 2019). The R statistical software was used for this analysis to identify and investigate the lag time with maximum correlation value, to utilize in the development of random forest model later.

2.3.1. Random Forest Model Approach

First step was to divide the data in two sets that is training (1982-2008) and testing (2009-2018). Within the decision tree techniques, random forest (RF) proposed by Brieman (2001) was used which is a non-parametric and semi-unsupervised technique and includes a combination of trees that are uncorrelated to give prediction for the tasks of regression and classification. A single decision tree can generate maximum variance and is also prone to noise (James et al. 2013), therefore to address this limitation, RF produces multiple trees in which every tree is created on the training data's bootstrapped sample. A binary split is made every time in a tree known as split node, without replacement a random subset of predictors is considered from the whole set of predictor variables. Out of these predictors, one predictor make the split where in the two resulting nodes, the response variable's expected sum variances are minimized. The process of randomization in producing the features subset impedes one or more especially strong predictor from getting repeatedly selected at every split, resulting in maximum correlated trees (James et al. 2013). After the growth of all trees, every tree casts a vote on a label class for the prediction value for regression or classification task. The result is the average of all the regression values or the most popular class. Waikato Environment for Knowledge Analysis (WEKA) version 3.8.4 software was used for the purpose. The lagged climate indices and precipitation data for training was entered in the software. Because of with replacement sampling, some observations may not be chosen during the boostrap which are known as out-of-bag or OOB, and utilized for the estimation of the error

of the tree on data that is unseen. According to a study, it is estimated that about 37 percent of the samples add up OOB data. OOB is calculated on average for every subsequent added tree to give the performance gain estimate. This OOB is especially sensitive to the number of trees and number of random predictors at every split (Huang and Boutros, 2016). The predictive performance of the model generally ameliorate (or OOB error reduction) as the number of trees increases. The default number of iterations i.e. the number of trees was 100 which were tuned until the performance was improved and in our case 500 trees showed improved performance of the model and when we increased the number of trees after this limit, the performance was slightly improved with slight reduction in OOB error from the previous one. From this observation, we came to know that there is limit for the number of tress after which growing more trees cannot improve the performance of the model which is in accordance with the previous research (Oshiro et al. 2012). In addition to the parameters listed above, a key parameter for the random forest is the number of predictors to consider in every point of split. In Weka this can be changed by the numFeatures attribute, which is set to zero by default and automatically choose the value based on a rule of thumb (Brownlee, 2016).

2.3.2. *Performance Evaluation Metrics*

The performance of the model was evaluated using three metrics i.e. correlation (r), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).

MAE is the sum of the absolute of all error. Mathematically it can be represented as follows:

$$MAE = \frac{\sum_{i=1}^{n} |Y_i - X_i|}{n} \quad (5)$$

The square root of the mean of the square of all of the error is known as the RMSE. To measure the performance of the model, it is considered the standard statistical metric in the studies of meteorology, climate research and air quality (Chai and Drexler, 2014). The closer the value to zero, the best is the performance of the model. Mathematically, it can be represented as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (X_i - Y_i)^2}{n}} \quad (6)$$

Where X_i is the measured precipitation and Y_i is the model predicted precipitation and n is the number of values (Al-Mukhtar and Qasim, 2019).

3. Results

3.1. Statistical Properties of Precipitation

In this research, concurrent correlations between Peshawar precipitation and climate indices such as ENSO (Nino-1+2, Nino-3, Nino-3.4, Nino-4, SOI, MEI, and SST) and QBO on monthly, seasonal and annual basis and antecedent (cross-correlations) on monthly basis were investigated. The statistical significant antecedent correlations of precipitation with the climate indices were further analyzed utilizing random forest (RF) model for the prediction of precipitation.

Table 1 presents the statistical data (mean and standard deviation) of monthly, seasonal and annual measured precipitation in Peshawar city. On monthly basis, the maximum mean precipitation is during the month of March (69.27) while minimum during November (10.44). Generally, the distribution of monthly precipitation is variable. Standard deviation (Std. Dev) is suitable for the representation of variability. The monthly Std. Dev for the region lies between 14.89 and 49.16. The highest value is observed during August (49.16) and minimum (14.89) during November. Similarly, the seasonal Std. Dev lies between 16.09 (winter) and 32.15 (spring) while the annual value is 39.93.

Table 1. Statistical data of monthly, seasonal and annual measured precipitation in Peshawar

Months	Mean Precipitation (mm)	Std. Dev
Jan	29.55	22.48
Feb	56.60	41.68
Mar	69.27	48.93
Apr	53.38	43.34
May	26.07	31.24
Jun	16.17	20.44
Jul	51.60	48.13
Aug	52.37	49.16
Sep	23.81	29.58
Oct	11.22	24.54
Nov	10.44	14.89
Dec	22.02	25.94
Winter	36.00	16.09
Spring	47.92	32.15
Summer	40.05	30.79
Autumn	15.16	17.66
Annual	35.21	39.93

3.2. Concurrent Correlations

3.2.1. Monthly Correlations

To reveal the concurrent correlations, the monthly climate indices (Nino-3.4, SOI, MEI, QBO, Nino-3, Nino-4, SST and Nino-1+2) have been treated against the precipitation (Table 2). Best positive correlation is observed between precipitation and the SOI index during June at the significance level of 0.05. Nino-3.4 index has significant positive correlations during April (0.05 significance level) and November (0.01 significance level). Similarly Nino-3 has positive significant correlations during the April, May (0.05 significance level) and November (0.01 significance level). Nino-4 has significant positive correlations during the months of March and November (0.05 significance level) and negative during December (0.01 significance level). MEI index shows significant positive correlation in the month of November at the 0.05 significance level. SST has significant positive correlations during July and November at the significance level of 0.05 and June, August and October at the 0.01 significance level. The index of QBO did not show significant correlation and the highest are observed during September and December. Nino-1+2 have significant positive correlations during May at the significance level of 0.01 and November at the significance level of 0.05. It can be summarized that all the climate indices except QBO have shown their significant positive or negative correlations with precipitation in different months.

3.2.2. Seasonal and Annual Correlations

The concurrent correlations between the seasonal and annual precipitation and climate indices (Nino-3.4, SOI, MEI, QBO, Nino-3, Nino-4, Nino-1+2 and SST) is shown in Table 2. On annual basis, no significant correlations are observed between all the climate indices and precipitation. During the seasonal analysis, the significant positive correlations of 0.44, 0.34, 0.39 and 0.34 are observed for SOI, QBO, MEI and Nino-3 indices during the summer, autumn, spring and autumn seasons, respectively. The highest correlation is observed for SOI.

Figure 2 shows the annual precipitation and Nino-3.4 index variations for 37 years (1982-2018) in Peshawar. The left side of the graph shows precipitation and the right side shows Nino-3.4 index. Results indicate low positive correlation (0.13) between precipitation and Nino-3.4 index.

Figure 3 depicts the annual precipitation and SOI index variations for 37 years (1982-2018) over Peshawar region. The graph's left side is the precipitation and the right side is the SOI index values. Results manifest very low negative correlation (-0.05) between precipitation and SOI Index.

Figure 4 depicts the annual precipitation and QBO index changes in 37 years (1982-2018) over Peshawar area. The left side of the graph shows the precipitation and right side shows the QBO index values. Results show insignificant low positive correlation (0.01) between precipitation and QBO index.

Figure 5 exhibits the annual precipitation and MEI index fluctuations during the period of 37 years from 1982 to 2018. Precipitation is represented by left side and MEI index by right side. Findings indicate insignificant low positive correlation (0.04) between precipitation and MEI index.

Figure 6 shows annual precipitation and Nino-4 index changes for the period of 37 years (1982-2018) over Peshawar. The graph's left side represent the precipitation and right side the Nino-4 index values. Findings indicate insignificant low positive correlation (0.04) between the precipitation and Nino-4 index.

Figure 7 represents annual precipitation and Nino-3 index alterations for 37 years (1982-2018) over Peshawar. The left side of the plot manifests the precipitation and right side the Nino-3 index. Findings indicate that there is insignificant positive correlation (0.24) between precipitation and Nino-3.

Figure 8 represents annual precipitation and Nino-1+2 index variability for 37 years (1982-2018) over Peshawar. The precipitation is represented by left side and Nino-1+2 index by right side. Findings indicate that there is insignificant positive correlation (0.27) between precipitation and Nino-1+2.

Figure 9 represents annual precipitation and SST variation for 37 years (1982-2018) over Peshawar. The left side of the plot shows the precipitation and right side the SST. Findings indicate that there is insignificant low positive correlation (0.05) between precipitation and SST.

Figure 10 shows annual precipitation and all ENSO indices combined variation during 37 years (1982-2018) over Peshawar region. The left side of the graph is the precipitation and right side manifests the combination of all ENSO indices. The relationship between precipitation and all ENSO indices on annual basis is insignificant while the highest is observed for Nino-1+2 and Nino-3.

Table 2. Seasonal and annual correlation coefficients between climate indices and precipitation data

Peshawar Station					
Indices/Seasons	Annual	DJF (Winter)	MAM (Spring)	JJA (Summer)	SON (Autumn)
Nino-3.4 Index	0.13	0.17	0.03	-0.14	-0.14
SOI Index	-0.05	-0.14	0.02	0.44**	0.07
MEI Index	0.04	0.05	0.20	-0.30	0.39*
QBO Index	0.01	0.14	0.01	0.26	0.34*
Nino-3 Index	0.24	-0.15	0.34*	0.09	-0.15
Nino-4 Index	0.04	-0.28	-0.27	-0.08	-0.09
Nino-1+2	0.27	-0.03	0.15	-0.08	0.07
SST	0.05	0.25	0.05	0.05	-0.15

**Correlation is significant at the 0.01 level (2-tailed).
*Correlation is significant at the 0.05 level (2-tailed).

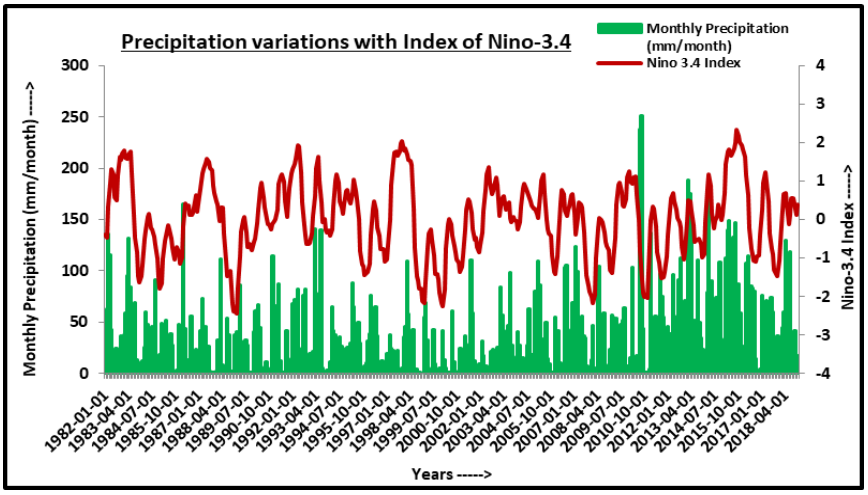

Fig. 2. Variation in precipitation and Nino-3.4 index over Peshawar

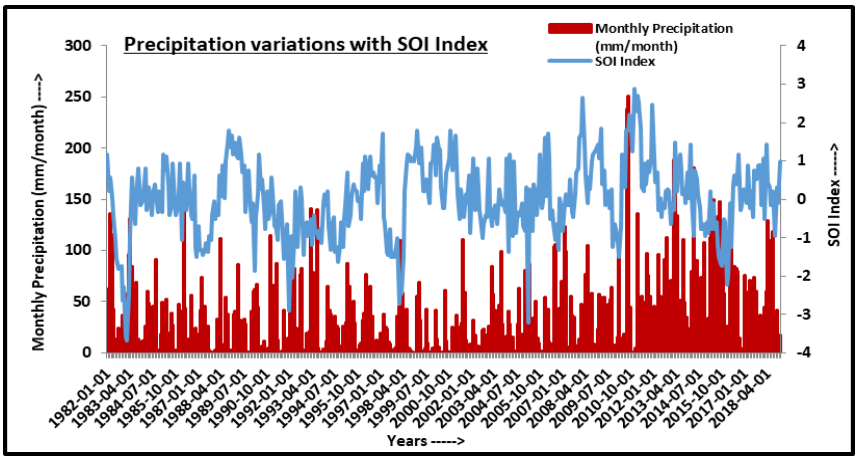

Fig. 3. Precipitation variations and SOI index over Peshawar

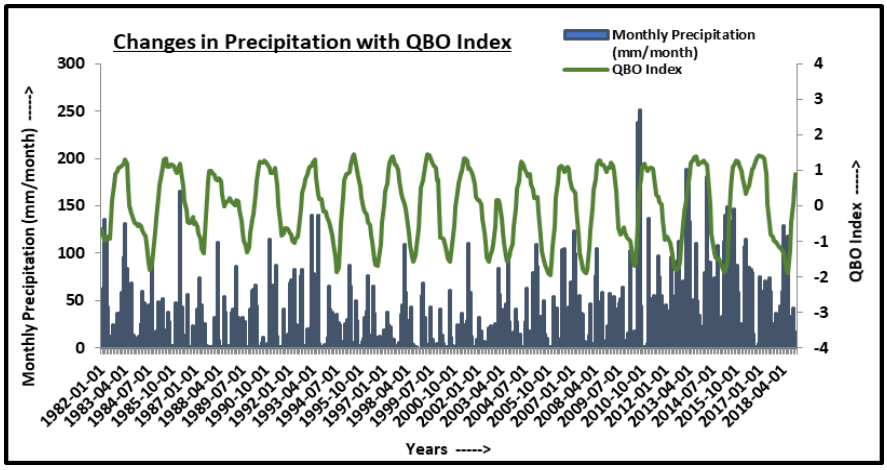

Fig. 4. Changes in precipitation with index QBO over Peshawar

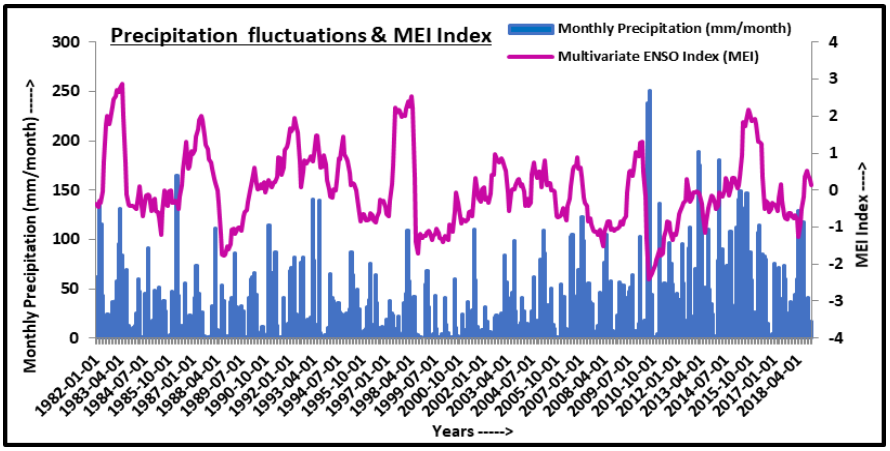

Fig. 5. Precipitation fluctuations with index of MEI over Peshawar

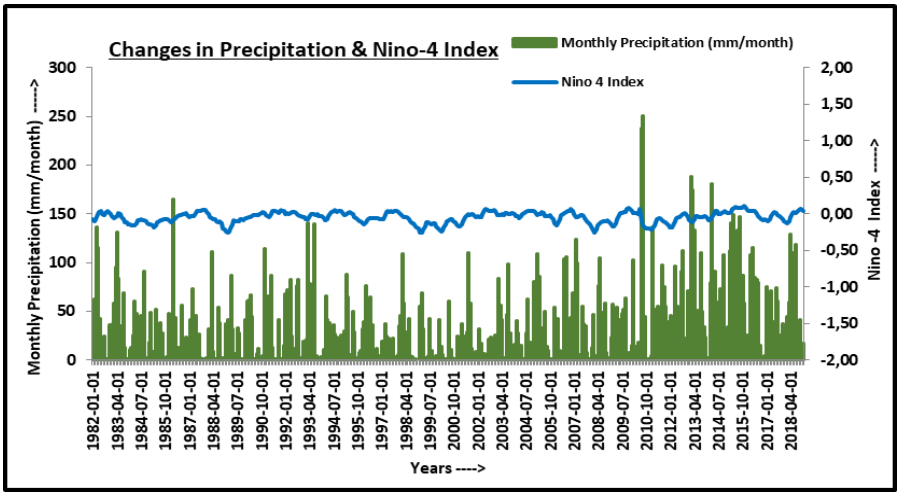

Fig. 6. Nino-4 index and precipitation changes over Peshawar

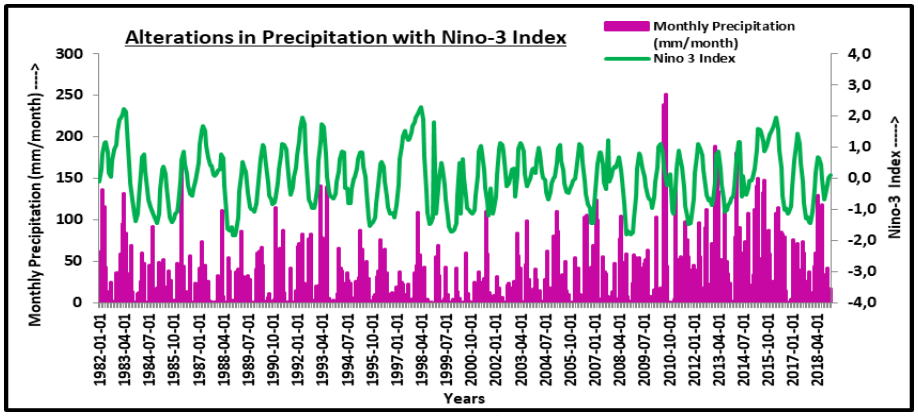

Fig. 7. Nino 3 index and alterations of precipitation over Peshawar

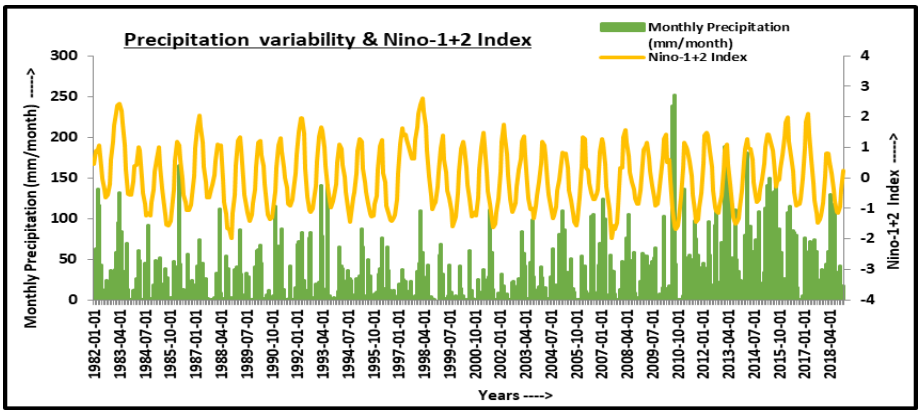

Fig. 8. Nino-1+2 index and precipitation variability over Peshawar

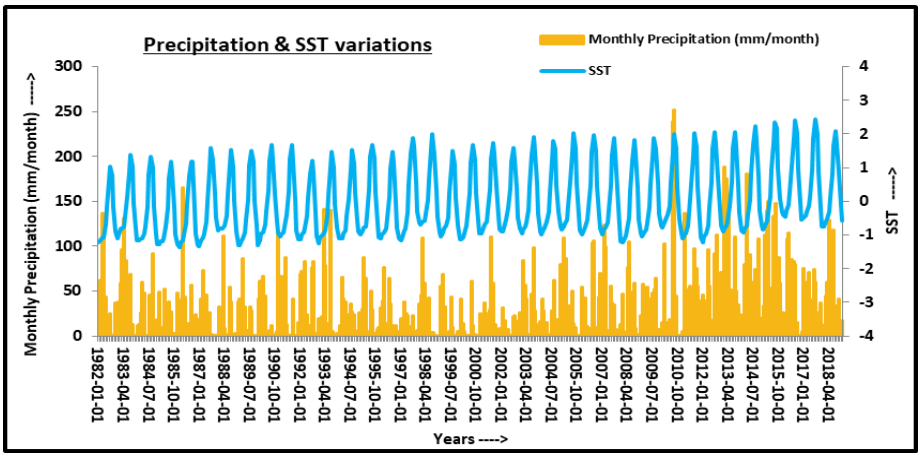

Fig. 9. Precipitation variations with SST over Peshawar

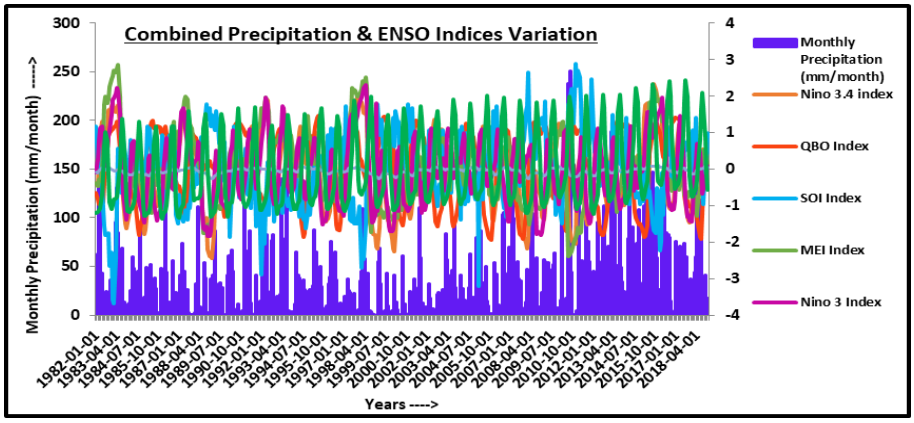

Fig. 10. Combined changes in precipitation with ENSO indices over Peshawar city

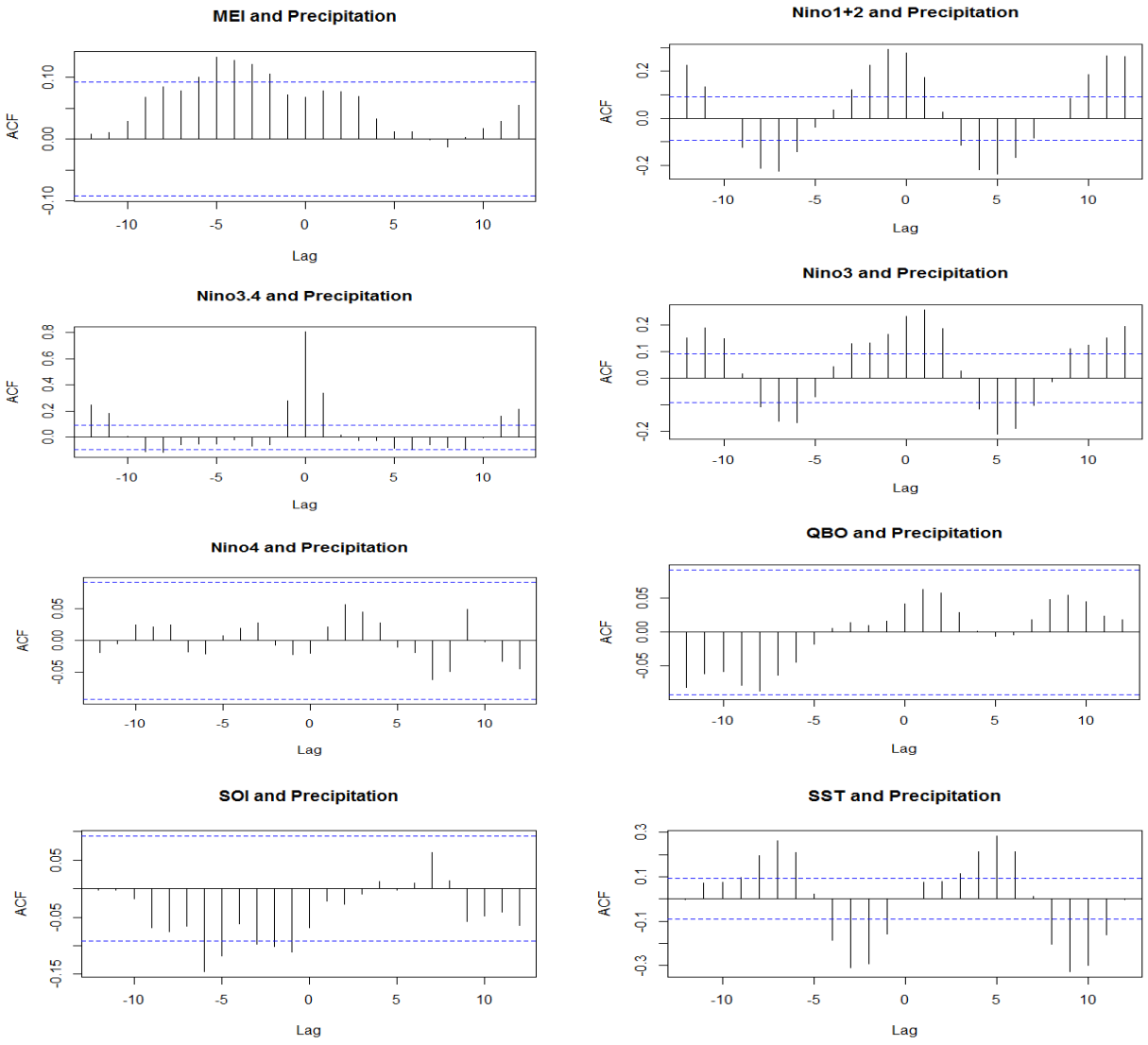

Fig. 11. The characteristics of cross-correlation among the MEI, Nino1+2, Nino3.4, Nino3, Nino4, QBO, SOI and SST index and the monthly rainfall

3.2.3. Association between climate indices and precipitation by cross-correlation

The cross-correlation of eight indices such as MEI, Nino-1+2, Nino-3, Nino-3.4, Nino-4, SOI, QBO and SST and the precipitation at the Peshawar city were applied in order to find the relationship of two time series while utilizing a range of possible alignments. In this process, significant applications include correlation and the finding of delay times between the two measured variables. Figure 11 shows the association between the variables for 12 months. Here, the negative lag values mean that rainfall is leading while the positive lags mean that the climate index is leading. We hypothesized that the actual association should be that the climate index is leading and in the same direction correlated with the precipitation. The climate index was manifested to have a correlation to the precipitation, and on the lag axis, the cyclic pattern is because of signals periodicities. The blue lines in all graphs of figure 11 show the lines of 95% significance and any lines of lagged correlations crossing them are considered as significant lagged correlations.

Figure 12 manifests the selected lagged times highest significant correlations between the precipitation and the climate indices. Six out of eight climate indices have the maximum lagged correlations. MEI has 0.13 highest correlation at the lag (-5), SOI has -0.15 at lag (-6), Nino-3.4 has 0.34 at lag (1), Nino-3 has 0.26 at lag (1), Nino-1+2 has 0.29 at lag (-1), SST has -0.33 at lag (9) while QBO and Nino-4 have very low insignificant correlations and are therefore not considered further in the analysis.

3.2.4. Random Forest (RF) Model

Figure 13 shows the TRMM precipitation and predicted precipitation by Random Forest (RF) model for Peshawar city during the training period (1982-2008). The model's predicted precipitation followed the trend of TRMM precipitation and performed well during the training period but its performance was low at the extreme precipitation phases. Similarly during the test period (2009-2018), the model followed the trend of TRMM precipitation but at the extreme phases of precipitation its performance was low (Figure 14).

We next evaluated the overall performance of the RF model using three metrics r, MAE and RMSE (Table 3). We observe similar trend in Figure 13, when RF performs better on training set (higher r, lower MAE and RMSE). The correlation (r), MAE and RMSE in case of training period are 0.97, 0.03 and 0.04 respectively which are pretty low showing the best performance of the model and in case of test period they are 0.40, 0.14 and 0.21 showing a bit low performance.

Table 3. Performance evaluation of the Random Forest (RF) Model

Station	Period	Correlation (r)	MAE	RMSE
Peshawar	Training (1982-2008)	0.97	0.03	0.04
	Test (2009-2018)	0.40	0.14	0.21

4. Discussion

Some studies carried out in this context in Pakistan are mentioned. In a study to analyze the influence of North Atlantic Oscillations (NAO) and Southern Oscillation (SO) on winter precipitation in Pakistan, it was concluded that statistical correlation of SOI with the winter precipitation remained better contributor towards the activity of rainfall in Hindukush, Karakoram and Himalaya (HKH) areas. Its negative phase was responsible for the above normal rainfall and the positive one for the low rainfall (Afzal et al. 2013). Another research conducted to investigate the effect of El Nino event on the summer monsoon rainfall over Pakistan, where the monthly correlation between MEI and percent summer monsoon rainfall departure was calculated. In all Pakistan, during the monsoon months (July, August and September), July and September month correlation was significant while during August it was insignificant. In Punjab province, during all (July, August and September) months it was significant and in Balochistan province no significant correlation was found. In case of Sindh and NWFP (presently Khyber Pakhtunkhwa) provinces, the correlation was significant during August and July respectively (Mahmood et al. 2006).

A research carried out in Pakistan for analyzing the influence of El Nino Southern Oscillation (ENSO), North Atlantic Oscillation (NAO) and Indian Ocean Dipole (DMI) on precipitation in Pakistan found that the positive DMI values and El-Nino years were the reason for deficient rainfall over the whole Pakistan and the monsoon region. It was also discussed that the La-Nina years were generally positively correlated with the precipitation. They found that the emerging monsoon depressions from the Bay of Bengal which moves towards the Pakistan did not reach and instead dispersed over either India or Bangladesh.

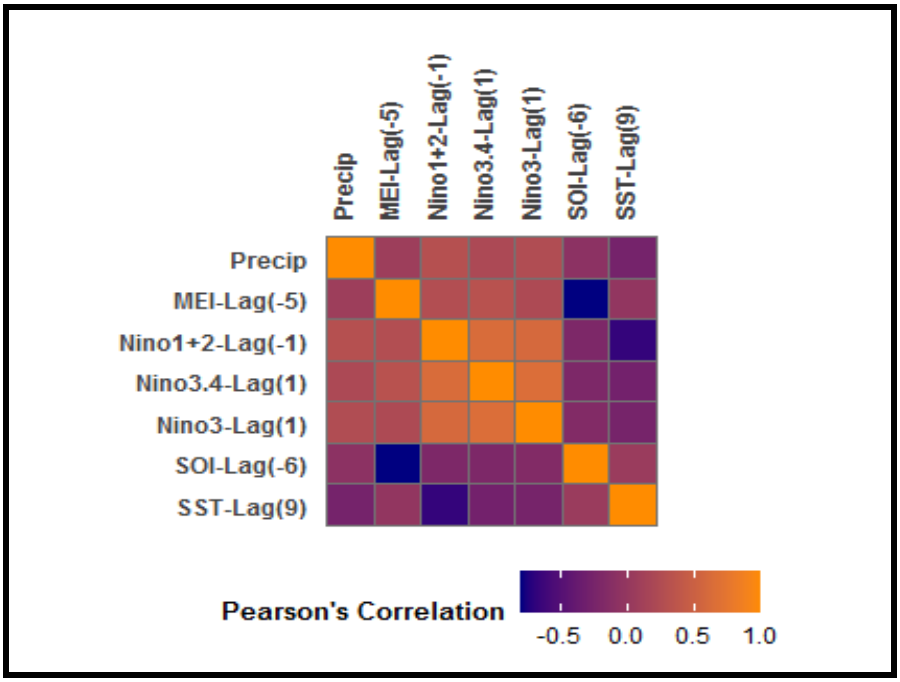

Fig. 12. The selected highest lagged correlation indices

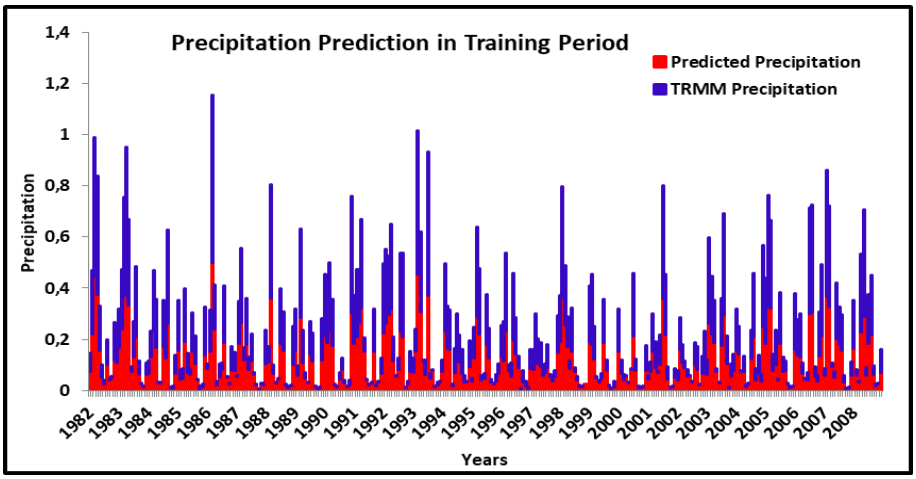

Fig. 13. Prediction of precipitation in Peshawar city through Random Forest model (Training Period)

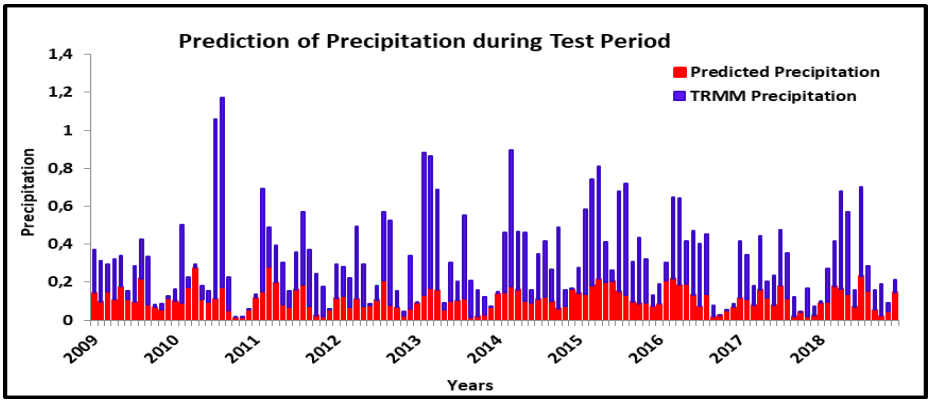

Fig. 14. Random Forest Model's prediction of precipitation in Peshawar city (Test Period)

The correlations of NAO with the winter rains were found positive (Adnan et al. 2016). In another study, Iqbal and Athar (2017) studied the variability, trends and teleconnections of observed precipitation in the Pakistan region (Gilgit-Baltistan, Khyber Pakhtunkhwa, Balochistan, Punjab, Sindh and Azad Jammu and Kashmir) while taking into account the circulation indices (NAO, IOD, ENSO, AO, PDO, QBO and AMO) with the precipitation. The summer season received maximum precipitation (45%), whereas the pre-monsoon (post-monsoon) and winter seasons contributed 30 and 20% (5%) respectively, towards the whole annual precipitation. Also during the monsoon and post-monsoon seasons, significant effect of ENSO was noticed in all four seasons in the KPK, AJK, Punjab, and Balochistan.

5. Conclusions

For Peshawar which is a rapidly growing city, the behavior of precipitation particularly the variation of it with the climate indices are important for the designing, planning and development of water resources management strategies. The present study identify the concurrent and antecedent relationship between climate indices such as El Nino Southern Oscillation (Nino-1+2, Nino-3, Nino-3.4, Nino-4, SOI, MEI and SST), Quasi-Biennial Oscillation and the precipitation in Peshawar City-Pakistan and predict the precipitation in the region. The results manifested that on monthly basis, the highest significant concurrent correlation is observed for SOI during June (0.48), Nino-1+2 during May (0.47), SST during June (0.45), August (0.42), and October (0.46), Nino-3 and Nino-3.4 in November (0.44 and 0.43, respectively), and Nino-4 during December (-0.44). During the seasonal analysis, the significant positive concurrent correlations are observed for SOI (0.44), QBO (0.34), MEI (0.39) and Nino-3 (0.34) indices during the summer, autumn, spring, and autumn seasons, respectively. The highest concurrent correlation is observed for SOI. On annual basis, no significant concurrent correlations are noticed. In the antecedent correlation analysis, six out of eight climate indices have the maximum lagged correlations (MEI at lag (-5), SOI at lag (-6), Nino-3.4 at lag (1), Nino-3 at lag (1), Nino-1+2 at lag (-1), SST at lag (9)). During prediction, the random forest model's predicted precipitation followed the trend of TRMM precipitation and performed well in the training period but its performance at the extreme precipitation phases was low and similarly during the test period. The results obtained from this study are of importance to help policy makers in decisions making and planning for adaptation to the effects of climate oscillations. Moreover, the findings can provide a support for the better management of water resources in Peshawar.

Acknowledgements

Authors wish to thank the National Oceanic and Atmospheric Administration (NOAA) for climate indices data and National Aeronautics and Space Administration (NASA) and Japan Aerospace Exploration Agency (JAXA) and also to the NASA Goddard Earth Sciences Information Services Center (DISC) for the provision of TRMM satellite precipitation data free of cost for usage in the study.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Ahmed R., Alam M.S., Rahman M.M. (2016). Long Term Trend of All-Bangladesh Summer Monsoon Rainfall, and its Association with the ENSO Index. Journal of Environment and Earth Science 6(4).

Ashok K., Saji N.H. (2007). On the impacts of ENSO and Indian Ocean dipole events on sub-regional Indian summer monsoon rainfall. Natural Hazards 42, 273–285.

As-syakur A.R., Adnyana I.W.S., Mahendra M.S., Arthana I.W., Merit I.N., Kasa I.W., Ekayanti N.W., Nuarsa I.W., Sunarta I.N. (2014). Observation of spatial patterns on the rainfall response to ENSO and IOD over Indonesia using TRMM Multisatellite Precipitation Analysis (TMPA). Internationa Journal of Climatology 34, 3825–3839.

Ailikun B., Yasunari T. (2001). ENSO and Asian Summer Monsoon: Persistence and Transitivity in the Seasonal March. Journal of Meteorological Society of Japan 79, 145–159.

Akhtar I.ul.H., Athar H. (2019). Contribution of changing precipitation and climatic oscillations in explaining variability of water extents of large reservoirs in Pakistan. Scientific Reports 9, 1–14.

Al-Mukhtar M., Qasim M. (2019). Future predictions of precipitation and temperature in Iraq using the statistical downscaling model. Arabian Journal of Geosciences 12(2), 12-25.

Afzal M., Haroon M. A., Rana A. S., Imran A. (2013). Influence of North Atlantic oscillations and Southern oscillations on winter precipitation of Northern Pakistan. Pakistan Journal of Meteorology 9(18).

Adnan M., Rehman N., Khan A.A, Mir K.A., Khan M.A. (2016). Influence of Natural Forcing Phenomena on Rainfall of Pakistan. Pakistan Journal of Meteorology 12, 23–35.

Adnan M., Rehman N., Sheikh M.M., Khan A.A., Mir K.A., Khan M.A. (2016). Influence of Natural Forcing Phenomena on Precipitation of Pakistan. Pakistan Journal of Meteorology 12(24), 23-35.

Bokhari S.M.U.H. (2015). Pakistan Emergency Situational Analysis District Peshawar. Alhasan Systems Private limited, Islamabad, Pakistan. ISSN: 2410-8820.

Breiman L. (2001). Random forests. Machine learning 45, 5–32.

Baldwin M.P., Gray L.J., Dunkerton T.J., Hamilton K., Haynes P.H., Randel W.J., Holton J.R., Alexander M.J., Hirota I., Horinouchi T., Jones D.B.A., Kinnersley J.S., Marquardt C., Sato K.,

Takahashi M. (2001). The quasi-biennial oscillation. Reviews of Geophysics 39, 179–229.

Bhalme H.N., Rahalkar S.S., Sikder A.B. (1987). Tropical Quasi-Biennial Oscillation of the 10-mb wind and Indian monsoon rainfall-implications for forecasting. Journal of Climatology 7, 345–353.

Brownlee J. (2016). How to Use Ensemble Machine Learning Algorithms in Weka. Machine Learning Mastery. Available at: https://machinelearningmastery.com/use-ensemble-machine-learning-algorithms-weka/.

Bhutto A., Wei M., Liu Y.A., Li N. (2009). Impact of ENSO on summer monsoon in southern parts of Pakistan. In: 2009 1st International Conference of Information Science and Engineering ICISE, 4903–4906.

Claud C., Terray P. (2007). Revisiting the possible links between the quasi-biennial oscillation and the Indian summer monsoon using NCEP R-2 and CMAP fields. Journal of Climate 20, 773–787.

Chai T., Draxler R.R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?- Arguments against avoiding RMSE in the literature. Geoscientific Model Development 71247-1250.

Fan L., Xu J., Han L. (2019). Impacts of onset Time of El Niño events on summer rainfall over southeastern australia during 1980-2017. Atmosphere (Basel) 10.

Hossain I., Rasel H.M., Imteaz M.A., Pourakbar S. (2015a). Effects of climate indices on extreme rainfall in Queensland, Australia. In: Proc - 21st International Congress on Modelling and Simulation, Gold Coast, Australia, 29 November-4 December.

Hossain I., Rasel H.M., Imteaz M.A., Moniruzzaman M. (2015b). Statistical correlations between rainfall and climate indices in Western Australia. In: 21stInternational Congress on Modelling and Simulation, Gold Coast, Australia, 29 November-4 December.

Hidayat R., Ando K., Masumoto Y., Luo J.J. (2016). Interannual Variability of Rainfall over Indonesia: Impacts of ENSO and IOD and Their Predictability. IOP Conference Series: Earth and Environmental Science 3.

Huang B.F., Boutros P.C. (2016). The parameter sensitivity of random forests, BMC bioinformatics 17, 331.

Houze R.A., Rasmussen K.L., Zuluaga M.D., Brodzik S.R. (2015). The variable nature of convection in the tropics and subtropics: a legacy of 16 years of the Tropical Rainfall Measuring Mission satellite. Reviews of Geophysics 53, 994–1021.

Hirose M., Takayabu Y.N., Hamada A., Shige S., Yamamoto M.K. (2017). Impact of long-term observation on the sampling characteristics of TRMM PR precipitation. Journal of Applied Meteorology and Climatology 56, 713–723.

Islam F., Imteaz M.A., Rasel H.M. (2017). Analysing the Effect of Lagged Climate Indices on Rainfall Predictability for Western Australia's North Coast Region. In: 22nd International Congress on Modelling and Simulation, Hobart, Tasmania, Australia, 3-8 December.

Indeje M., Semazzi F.H.M. (2000). Relationships between QBO in the lower equatorial stratospheric zonal winds and East African seasonal rainfall. Meteorology and Atmospheric Physics 73, 227–244.

Ihara C., Kushnir Y., Cane M.A., Pena V.H.D.L. (2007). Indian summer monsoon rainfall and its link with ENSO and Indian Ocean climate indices. International Journal of Climatology 27, 179-187.

Iqbal M.F., Athar H. (2017). Variability, trends, and teleconnections of observed precipitation over Pakistan. Theoretical and Applied Climatology 134, 613–632.

James G., Witten D., Hastie T., Tibshirani R. (2013). An Introduction to Statistical Learning, 103 XIV.

Kane R.P. (1989). Relationship between the southern oscillation/El Nino and rainfall in some tropical and midlatitude regions. Proceedings of the Indian Academy of Sciences – Earth and Planetary Sciences 98, 223–235.

Khan A.H. (2004). The influence of La-Nina phenomena on Pakistan's precipitation. Pakistan Journal of Meteorology 1, 23–31.

Lee H.S. (2015). General Rainfall Patterns in Indonesia and the Potential Impacts of Local Seas on Rainfall Intensity. Water 7, 1751-1768.

Li X., Ting M. (2015). Recent and future changes in the Asian monsoon-ENSO relationship: Natural or forced? Geophysical Reserach Letters, 42, 3502–3512.

Lee D.K., In J., Lee S. (2015). Standard deviation and standard error of the mean. Korean Journal of Anesthesiology 68(3), 220-223.

Mekanik F., Imteaz M.A., Gato-Trinidad S., Elmahdi A. (2013). Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes. Journal of Hydrology 503, 11–21.

Mazzarella A., Giuliacci A., Liritzis I. (2011). QBO of the Equatorial-Stratospheric Winds Revisited: New methods to verify the dominance of 28-month cycle. International Journal of Ocean and Climate Systems 2(1), 19-26.

Mahmood A., Khan T.M.A., Faisal N. (2006). Relationship between El Nino and Summer Monsoon Rainfall over Pakistan. Pakistan Journal of Marine Sciences 15(2), 161-178.

Mahmood A., Khan T., Faisal N. (2004). Correlation between multivariate ENSO Index (MEI) and Pakistan's summer rainfall. Pakistan Journal of Meteorology 1, 53–64.

Mehmood R., Butt M.A., Mahmood S.A., Ali F. (2017). Appraisal of Urban Heat Island and Its Impacts on Environment Using Landsat TM in Peshawar, Pakistan. Advances in Remote Sensing 6(3), 192-200.

Menke W., Menke J. (2012). Environmental Data Analysis with MatLab. Elsevier, pp: 288.

Naheed G., Rasul G. (2011). Investigation of Rainfall Variability for Pakistan. Pakistan Journal of Meteorology 7, 25–32.

Oshiro T.M., Perez P.S., Baranauskas J.A. (2012). How many trees in a random forest?, In: International workshop on machine learning and data mining in pattern recognition, Springer, pp: 154–168.

Obilor E.I., Amadi E.C. (2018). Test for Significance of Pearson's Correlation Coefficient (r). International Journal of Innovative Mathematics, Statistics & Energy Policies 6(1), 11-23.

Ruigar H., Golian S. (2015). Assessing the correlation between climate signals and monthly mean and extreme precipitation and discharge of Golestan Dam Watershed. Earth Sciences Research Journal 19(1), 65-72.

Rasmusson E., Carpenter T. (1983). The Relationship between Eastern Equatorial Pacific Sea Surface Temperatures and Rainfall over India and Sri Lanka. Monthly Weather Review 111, 517–528.

Seo J., Choi W., Youn D., Park D-S.R., Kim J.Y. (2013). Relationship between the stratospheric quasi-biennial oscillation and the spring rainfall in the western North Pacific. Geophysical Research Letters 40, 5949–5953.

Syed F.S., Giorgi F., Pal J.S., King M.P. (2006). Effect of remote forcings on the winter precipitation of central southwest Asia part 1: Observations. Theoretical and Applied Climatology 86, 147–160.

Sanabria J., Bourrel L., Dewitte B., Frappart F., Rau P., Solis O., Labat D. (2018) Rainfall along the coast of Peru during strong El Niño events. International Journal of Climatology 38, 1737–1747.

Sarfaraz. (2007). Monsoon dynamics: Its behavioral impact in Pakistan's perspective. Pakistan Journal of Meteorology 4, 55–73.

Salma S., Rehman S., Shah M.A. (2012). Rainfall trends in different climate zones of Pakistan. Pakistan Journal of Meteorology 9(17), 37-47.

Sheikh M.M., Manzoor N., Adnan M. (2010). Precipitation Related Disasters in Pakistan , Linkage to Climate Change , Risk Reduction and Possible Adaptation Measures.

Safdar F., Khokhar M.F., Arshad M., Adil I.H. (2019). Climate Change Indicators and Spatiotemporal Shift in Monsoon Patterns in Pakistan. Advances in Meteorology 2019.

Taweesin K., Seeboonruang U. (2019). The relationship between the climate indices and the rainfall fluctuation in the lower central plain of Thailand. International Journal of Innovative Computing, Information and Control 15(1), 107-127.

Trenberth K., NCAR Staff (Eds). Last modified 21 Jan, 2020. The Climate Data Guide: Nino SST Indices (Nino 1+2, 3, 3.4, 4; ONI and TNI). Retrieved from https://climatedataguide.ucar.edu/climate-data/nino-sst-indices-nino-12-3-34-4-oni-and-tni.

WPR. (2020). World Population Review. Available at: https://worldpopulationreview.com/world-cities/peshawar-population. Accessed on: 14th Dec, 2020.

Yan X., Konopka P., Ploeger F., Tao M., Muller R., Santee M.L., Bian J., Riese M. (2018). El Nino Southern Oscillation influence on the Asian summer monsoon anticyclone. Atmospheric Chemistry and Physics 18, 8079-8096.

Yahiya Z., Chandimala J., Siriwardhana M., Zubair L. (2009). Sri Lankan Rainfall Climate and its Modulation by El Nino and La Nina Episodes. Engineer: Journal of the Institution of Engineers, Sri Lanka 42, 11.

Zawar M., Zahid M. (2013). Rainfall Patterns over Punjab, Khyber Pakhtunkhwah and Kashmir during El Nino and La Nina Years 1960-2008. Pakistan Journal of Meteorology 10, 83–90.