chinaXiv:202011.00129v1

# Precipitation forecasting by large-scale climate indices and machine learning techniques

Mehdi GHOLAMI ROSTAM, Seyyed Javad SADATINEJAD, Arash MALEKIAN[*]

University of Tehran, Tehran 1417466191, Iran

**Abstract:** Global warming is one of the most complicated challenges of our time causing considerable tension on our societies and on the environment. The impacts of global warming are felt unprecedentedly in a wide variety of ways from shifting weather patterns that threatens food production, to rising sea levels that deteriorates the risk of catastrophic flooding. Among all aspects related to global warming, there is a growing concern on water resource management. This field is targeted at preventing future water crisis threatening human beings. The very first stage in such management is to recognize the prospective climate parameters influencing the future water resource conditions. Numerous prediction models, methods and tools, in this case, have been developed and applied so far. In line with trend, the current study intends to compare three optimization algorithms on the platform of a multilayer perceptron (MLP) network to explore any meaningful connection between large-scale climate indices (LSCIs) and precipitation in the capital of Iran, a country which is located in an arid and semi-arid region and suffers from severe water scarcity caused by mismanagement over years and intensified by global warming. This situation has propelled a great deal of population to immigrate towards more developed cities within the country especially towards Tehran. Therefore, the current and future environmental conditions of this city especially its water supply conditions are of great importance. To tackle this complication an outlook for the future precipitation should be provided and appropriate forecasting trajectories compatible with this region's characteristics should be developed. To this end, the present study investigates three training methods namely backpropagation (BP), genetic algorithms (GAs), and particle swarm optimization (PSO) algorithms on a MLP platform. Two frameworks distinguished by their input compositions are denoted in this study: Concurrent Model Framework (CMF) and Integrated Model Framework (IMF). Through these two frameworks, 13 cases are generated: 12 cases within CMF, each of which contains all selected LSCIs in the same lead-times, and one case within IMF that is constituted from the combination of the most correlated LSCIs with Tehran precipitation in each lead-time. Following the evaluation of all model performances through related statistical tests, Taylor diagram is implemented to make comparison among the final selected models in all three optimization algorithms, the best of which is found to be MLP-PSO in IMF.

**Keywords:** backpropagation; genetic algorithms; machine learning; multilayer perceptron; particle swarm optimization; Taylor diagram

# 1　Introduction

There is a growing concentration on sustainable development in water resource management. This field is targeted at preventing future water crisis threatening human beings. The very first stage in such management is to recognize the prospective climate parameters influencing the future water resource conditions. Numerous prediction models, methods and tools, in this case, have been developed and applied so far. Indeed, it is well known within the climate science literature that the teleconnection, which is a valuable context in climatology, has the considerable ability to justify and project climate parameters (Araghinejad and Meidani, 2013).

Teleconnection considers a distant phenomenon to study on regional climate conditions. As these large-scale climate phenomena are repetitive, they are categorized as patterns (Pozo et al., 2005), which can be effective both locally and globally and create much variability in climate parameters. Furthermore, they can result in drought and wet periods worldwide by changing precipitation trends (Hidalgo and Dracup, 2003; Zahraei and Karamouz, 2004; Kampichler et al., 2012; Choubin et al., 2014; Ouyang et al., 2014; Degefu and Bewket, 2017; Xu et al., 2018). Many definitions have been suggested for these patterns, which the main one identifies the teleconnection as a large-scale atmospheric-oceanic pattern, which is constant, repetitive and large-scale oscillated in some parameters, such as pressure (Wallace and Gutzler, 1981). For instance, El Niño, large-scale oceanic warming in the tropical Pacific Ocean, occurs repetitively every few years (Bjerknes, 1969). Its accompanying atmospheric component, Southern Oscillation, is the principal mode of pressure variability in the tropics, which affects the climate of many regions worldwide (Allan et al., 1996). The connections between the stated patterns and the climate parameters throughout different regions have been observed in many studies, e.g., Tyson (1987), Oldenberg et al. (2000), Gong and Ho (2003), Pozo-Vazquez et al. (2005), Canon (2007), and Gaughan and Waylen (2012). A well-known pattern is North Atlantic Oscillation (NAO), that in winter its station-based index is the climate variability mode in North Atlantic Ocean and is defined as the difference in normalized mean winter (December to March of the next year) sea level pressure (SLP) anomalies between Iceland and Portugal (Hurrell, 1995). Most modern NAO indices, however, are according to the simple difference in surface pressure anomalies between various northern and southern locations as well as Gibraltar and Reykjavik sites (Jones et al., 1997; Brandimarte et al., 2011). It is proved that this pattern has a strong impact on winter precipitation in Hungary (Matyasovszky, 2003), Turkey (Karabok et al., 2005) and Mediterranean precipitation trend (Santos et al., 2005). Indian Ocean Dipole (IOD), a coupled ocean-atmosphere phenomenon, is the other remarkable pattern defined in 1999 as a dipole pattern of sea surface temperature (SST) variability in the tropical Indian Ocean (Saji et al., 1999). Anomalous cooling of SST in the southeastern equatorial Indian Ocean and anomalous warming of SST in the western equatorial Indian Ocean normally characterize this pattern (Nigam et al., 1993). In the Mediterranean region, Eastern Mediterranean Pattern (EMP), related to the 500 hPa geopotential height between its east and west sides, is quite dominant (Hatzaki et al., 2007).

The connections of these and many other such patterns with climate parameters have been considered in numerous studies in which the methods, especially recent ones, derived from machine learning approaches. The slight difference between machine learning and statistical trajectories is their main points of concentrations, which in statistical methods are more on testing hypotheses, whereas in machine learning ones are more on formulating the process of generalization as a search through possible hypotheses (Wittan and Frank, 2005). Furthermore, machine learning focuses more on a prediction based on known features learned from exposure to datasets over the process known as training (Abbot et al., 2018). In Iran, few studies, accordingly, put their emphases on using such approaches in the area of large-scale climate indices (LSCIs), e.g., Ashrafi et al. (2012) and Choubin et al. (2014).

Iran, located in an arid and semi-arid region, suffers from severe water scarcity caused by mismanagement over years and intensified by climate change (Ghazal et al., 2014), which, consequently, has propelled a large number of population to immigrate towards more developed

cities especially in Tehran province. Sharp population growth in this province over recent years, is an apparent evidence of this fast-growing immigration. According to Statistical Center of Iran (2018) in the last 10 years, the population of Tehran province has increased from less than one million to over 15 million, which means this province, however, covering just 2% of entire Iran area, hosts 20% of country's population, of which 86.5% of the population resides in urban areas, particularly, in Tehran metropolitan, with about 8 million inhabitants. Providing water supplies among all other environmental issues in this metropolitan is of great concern. To tackle this complication an outlook for the future precipitation should be provided and appropriate forecasting trajectories compatible with this region's characteristics should be developed. To this end, this paper aims to compare the prediction capability of LSCIs by focusing on over 39 patterns and putting the emphasis on the application of three optimization algorithms in the multilayer perceptron network.

## 2 Datasets

Tehran metropolitan, facing Alborz mountain range to one side and the Iran central desert to another, varies in the term of elevation from the north to south (900 m a.s.l. in average); this significant difference causes a pronounced influence on the variability of precipitation over the city. As a whole, Tehran is a semi-arid city with mean annual precipitation of 230 mm. Among numerous synoptic stations over this city, the oldest one is Meharabad (35.69°N, 51.31°E; 1191 m a.s.l.), which has the longest climate records since 1951. The monthly precipitation of this station for a 62-a period (1951–2012) was obtained from the IMO (Iran Meteorological Organization). Those teleconnection patterns that had the most qualified datasets and have been suggested in previous papers for the case-study region (e.g., Abdi and Williams, 2010; Ashrafi et al., 2012; Araghinejad and Meidani, 2013; Choubin et al., 2014; Arvin, 2015; Gerkaninezhad and Bazrafshan, 2018), were obtained (Table 1) from National Oceanic and Atmospheric Administration (2015) for the period 1951–2012.

**Table 1** Initial large-scale climate indices (National Oceanic and Atmospheric Administration (2015)

| Row | Index | Row | Index |
|---|---|---|---|
| 1 | The Pacific/North American Pattern (PNA) | 19 | Arctic Oscillation (AO) |
| 2 | North Atlantic Oscillation (NAO) | 20 | Antarctic Oscillation (AAO) |
| 3 | West Pacific Pattern (WP) | 21 | Southern Oscillation Index (SOI) |
| 4 | North Pacific Pattern (NP) | 22 | Central Indian Precipitation |
| 5 | East Pacific Pattern (EP) | 23 | Northeast Brazil Rainfall Anomaly |
| 6 | Pacific Decadal Oscillation (PDO) | 24 | Tropical Northern Atlantic (TNA) |
| 7 | Eastern Pacific Oscillation (EPO) | 25 | Tropical Southern Atlantic (TSA) |
| 8 | North Oscillation Index (NOI) | 26 | Atlantic Meridional Mode (AMM) |
| 9 | El Nino – Southern Oscillation (ENSO) | 27 | Atlantic Multi-decadal Oscillation (AMO) |
| 10 | Multivariate ENSO Index (MEI) | 28 | Western Hemisphere Warm Pool (WHWP) |
| 11 | Extreme Eastern Tropical Pacific SST (Nino 1+2) | 29 | North Tropical Atlantic SST Index (NTA) |
| 12 | Central Tropical Pacific SST (Nino 4) | 30 | Oceanic NINO Index (ONI) |
| 13 | East Central Tropical Pacific SST (Nino 3.4) | 31 | Trans Nino Index (TNI) |
| 14 | Sahel Standardized Rainfall | 32 | Pacific Warm pool (PWP) |
| 15 | Eastern Asia/ Western Russia (EA/WR) | 33 | Indian Ocean Dipole (IOD) |
| 16 | Caribbean Index (CAR) | 34 | Solar Flux |
| 17 | Bivariate ENSO Time series (BEST) | 35 | Monthly totals Atlantic hurricanes and named tropical storms |
| 18 | Quasi-Biennial Oscillation (QBO) | 36 | North Sea-Caspian Pattern (NCP) |

## 3 Methodology

The schematic diagram of the entire study trajectory is shown in Figure 1. As shown, this study

consisted of four main steps. The first of was data collecting and choosing the most appropriate LSCIs, followed by stating the main two frameworks (i.e., Concurrent Model Framework (CMF), and Integrated Model Framework (IMF)). Pearson correlation, then, was used to define the most correlated LSCIs in each time lag before having the IMF. The third step was generating the three modelling methods, i.e., BP-based (backpropagation-based) MLP (multilayer perceptron), GA-based (genetic algorithm-based) MLP and PSO-based (particle swarm optimization-based) MLP. In the end, the most appropriate case was introduced by the Taylor diagram.
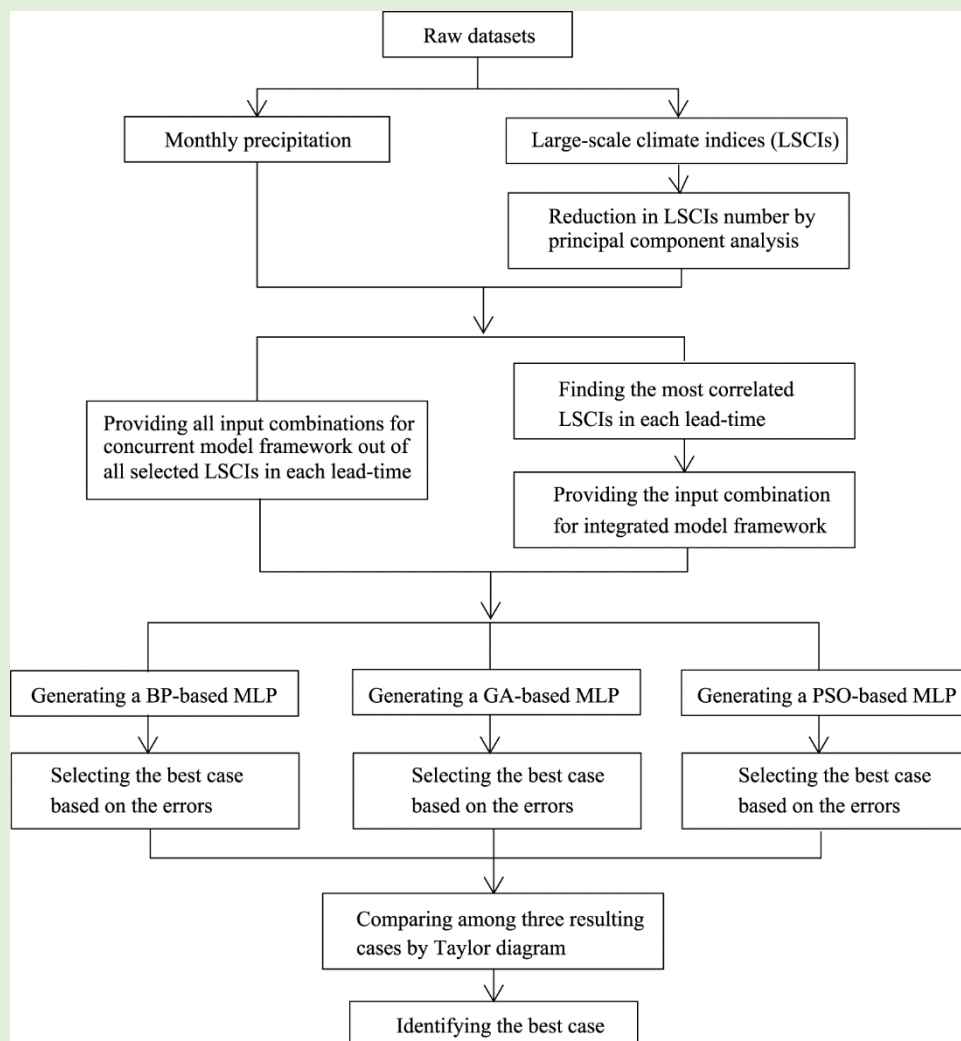
**Fig. 1**  Schematic study trajectory diagram. As can be seen, after providing the inputs of two frameworks (i.e., concurrent model framework and integrated model framework), three methods, namely, backpropagation-based (BP-based) multilayer perceptron (MLP), genetic algorithm-based (GA-based) MLP and particle swarm optimization-based (PSO-based) MLP were targeted.

### 3.1   Principal component analysis (PCA)

The idea of PCA is to reduce the dimensions in a dataset in which there are a large number of interrelated variables (Pasini, 2017). PCA is a useful tool to investigate simultaneously correlations between a large number of parameters, for finding subsets in data and defining outliers. Linear combinations of the principal components can be applied to reproduce parameters characterizing objects in the dataset (Abdi and Williams, 2010). In the present study, several LSCIs (i.e., 36 ones) were entered into the PCA process to be reduced. As the first step, the correlation matrix was preferred rather than the covariance matrix since its application is more

common than that of the covariance matrix (Jolliffe, 2002). The most controversial part of PCA is to define the optimized number of principal components (PCs) preventing serious loss of information. Among all different criteria and algorithms for choosing the PC numbers, the commonly used Kaiser's rule (Kaiser, 1960), which suggests retaining only those principal components whose variances exceed 1, was applied. After implementing the varimax rotation on loading matrix (Jolliffe, 2002) (Eq. 1), LSCIs with high rotated loadings were retained. As a conclusion, 9 LSCIs (i.e., Southern Oscillation Index, Multivariate ENSO Index, East Central Tropical Pacific SST, Central Tropical Pacific SST, Bivariate ENSO Time series, Oceanic NINO Index, Atlantic Meridional Mode, North Tropical Atlantic SST Index, and Tropical Northern Atlantic) were selected out of 36 ones.

$$\text{Loading matrix} = V \times L^{1/2},  \tag{1}$$

where $V$ is the eigenvectors matrix and $L^{1/2}$ is the diagonal square respective eigenvalues.

## 3.2 Multilayer perceptron (MLP)

The study of artificial neural networks is motivated by their similarity to biological systems consisting of very simple but numerous nerve cells (neurons) working massively in parallel and linked to each other in a weighted way (Kriesel, 2007). Equation 2 shows the formulated neural neuron operation.

$$y = f(\sum w_i x_i + b),  \tag{2}$$

where $y$ is the scalar output (Kriesel, 2007); $f$ is the neuron activation or transfer function; and $w_i$, $x_i$ and $b$ are weight, input and bias of the $i^{th}$ member, respectively. The MLP is the most frequently used neural networks (Popescu et al., 2009), which is in the feed-forward artificial neural network class consisting of at least three layers of input, hidden and output layer (Rosenblatt, 1961). Figure 2 demonstrates the schematic MLP topology used in this study. A three-layer MLP including one hidden layer was implemented.
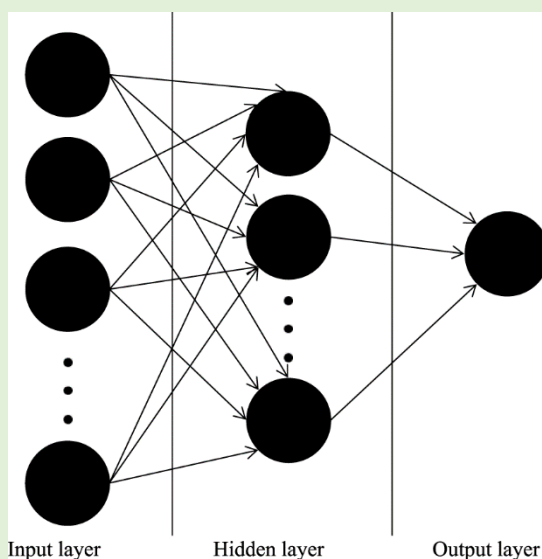


**Fig. 2** Schematic topology of a three-layer MLP (Multilayer perceptron) applied in this study. In the input layer for each input, one neuron should be considered; the number of the neurons of the hidden layer in each model was selected based on trial and error and represented in its relating section; the number of the neurons in the output layer, as can be seen, is the modeled precipitation measure for each month.

## 3.3 Optimization methods

The training methods that this study benefited from were BP, GAs (genetic algorithms) and PSO. The BP is an approach to calculate the gradient of the loss function of an MLP network concerning its weights (Ashrafi et al., 2012). It, however, is expensive in term of computational

cost but was found out within the academic community how simply weights of its hidden layer can be optimized. This method, basically, applies the gradient descent algorithm to minimize the network error. Regularly, a gradient descent algorithm is used to adapt the weights according to a comparison between the desired and actual network response (Seiffert, 2001); however, it is likely to become captured in local minimum when moving across a rugged error surface. Therefore, many suggestions have been made to avoid this failure, like GAs and PSOs. In this study, the selected function to train the BP-based MLP network was Levenberg-Marquardt, since the accuracy of results, in this case, was better than those of other training approaches. The applied activation functions in each neuron of the hidden and output layers were the sigmoid and linear functions, respectively. Moreover, the number of neurons in the hidden layer experimentally has been chosen in every model to reach the lowest error value.

Genetic algorithm (GA) is an evolutionary algorithm, in which a population of individuals evolves based on a set of bio-inspired theories such as selection, mutation and crossover (Srinivasan et al., 2003) to generate high-quality solutions for optimization problems (Mitchell, 1996). These algorithms encode a potential solution to a specific issue on a simple chromosome-like data structure (Whitley, 1994) competing with each other to achieve increasingly better results (Seiffert, 2001). Concerning the mentioned problem of gradient descent algorithm, implementing of GA as a complete substitution might lead the model to more precise conclusions.

Observations of social and collective behaviour of biological organisms are the main inspirations in PSO (Garro et al., 2015), specifically the movement of the best member of the population at the same time on their own experience. This seeking behaviour was associated with that of an optimization search for solutions to non-linear equations in a real-valued search space (Bratton et al., 2007). The population is considered as a cumulus of particles $i$ where each has a position $x_i \in R^D$, ($i=1, …, M$), in a multidimensional space. These particles are evaluated in a particular optimization function to recognize their fitness value and save the best solution. All the particles change their positions in the search space according to a velocity function $v_i$ which takes into account the best position of a particle in a population $P_g$ (social component) as well as their own best position $P_i$ (cognitive component) (Jiang et al., 2007). The particles will, repetitively, move to different positions until they find themselves in an optimum one (Garro and Vazquez, 2015).

The datasets, for all optimization methods should be divided into three sets: training, validation and test sets. In the present study, the first 70%, the next 15% and the last 15% of the entire dataset were, respectively, dedicated to the training, the validation and the test sets. As the MLP was aimed to be implemented for time series, randomly data splitting and assigning to each mentioned sets were inappropriate. Therefore, a solid division of datasets was considered.

### 3.4 Frameworks

This study benefited from two different model frameworks varying in the term of inputs: CMF) and IMF. In CMF, all contemporary LSCIs were entered in 12 individual models in each lead-time from 0 to 11 months in advance (Eq. 3), whereas in IMF the best LSCIs in each time lag, based on Pearson correlation, were entered in just one model (Eq. 4).

$$P_n = f(\text{Nino3.4}_n, \text{Nino4}_n, \text{AMM}_n, \text{MEI}_n, \text{SOI}_n, \text{TNA}_n, \text{NTA}_n, \text{ONI}_n, \text{BEST}_n)$$
$$P_{n-11} = f(\text{Nino3.4}_{n-11}, \text{Nino4}_{n-11}, ..., \text{BEST}_{n-11}) \qquad (3)$$

$$P_n = f(\text{LSCI}_n, \text{LSCI}_{n-1}, \text{LSCI}_{n-2}, \text{LSCI}_{n-3}, ..., \text{LSCI}_{n-12}), \qquad (4)$$

where $P_n$ is the amount of monthly precipitation in $n$ month. In Equation 3 all of the LSCIs were at the same time lags, which concluded 12 models from $P_n$ to $P_{n-11}$, whereas in Equation 4, precipitation in $n$ month was associated with a combination of the most correlated LSCIs in each time lags covering $n$ to $n–12$.

### 3.5 Model performance evaluation

Root mean square error (RMSE) and mean absolute error (MAE) were employed in the model

evaluation stage, both of which have been used as standard statistical metrics to measure model performance in a great variety of fields such as meteorology, air quality and climate research. These metrics were applied to select the most appropriate model in each optimization method based on the lowest error value. Besides, Z-test that indicates the accuracy of the selected models in the term of similarity between the means of desired and actual model responses was investigated. In Z-test, as long as Z (Eq. 5) remains below the critical value, null hypothesis (H0: $\mu_1=\mu_2$, where $\mu_1$ and $\mu_2$ are the means of the two comparing populations) is fulfilled and the model keeps the dataset mean approximately constant (Mann, 1997). The Z value formula is as follows:

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}}, \tag{5}$$

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}, \tag{6}$$

where $\mu$ is the mean of each population; $\bar{x}$ is the mean of the sample drawn from each population; $\sigma$ is the standard deviation; $n$ is the size of sample drawn from each population (Mann, 1997); and the subscript number 1 and 2 respectively represent observed and predicted. The advantage of using Z-test for the final three models led the results of the present study to be re-evaluated and more reliable.

### 3.6 Taylor diagram

Taylor diagram was applied to compare the performance of the final selected models. This diagram is particularly useful in assessing the relative merits of competing models and in monitoring overall performance as a model evolves (Taylor, 2001). Its overview is shown in Figure 3, in which the correlation coefficient between simulated and observed datasets is given by the azimuthal angle and the standard deviation of the simulated set is proportionally related to the radial distance from the origin (Xu et al., 2016).
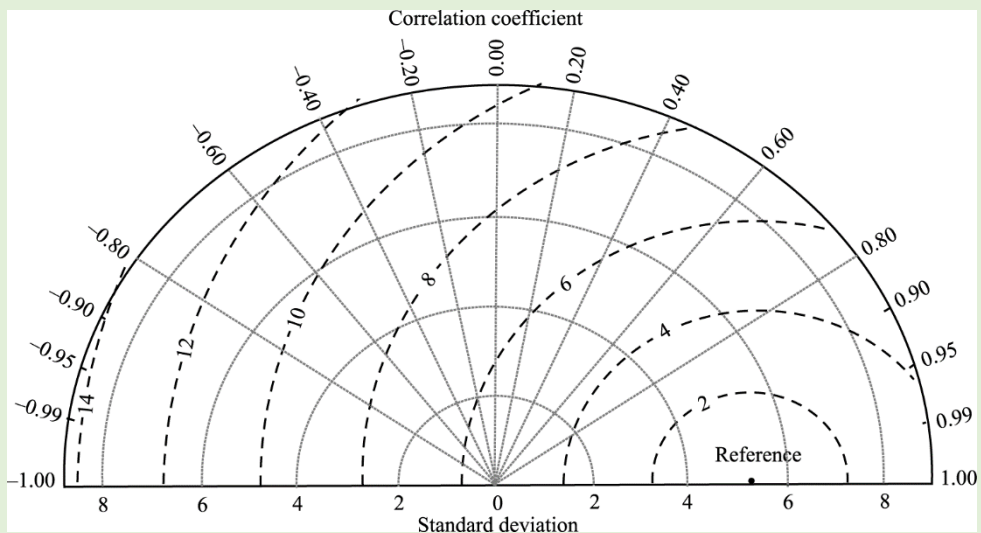


**Fig. 3** Diagram of statistical comparison among the models. The reference point on *x*-axis is the standard deviation of the observed dataset and the radial distance from this point (dashed contour) represents the RMSE (root mean square error) of the model; the dotted contours, which are the radial distance from the origin, shows the standard deviation of the simulated sets. It is clear that the most precise model is the one closer to the reference point (Taylor, 2001).

## 4   Results and discussion

For all 13 cases (12 ones generated in CMF and 1 in IMF), BP-based MLP, GA-based MLP and

PSO-based MLP were generated. Among all 9 selected LSCIs, SOI, Nino3.4 and Nino 4 were highly correlated with Tehran monthly precipitation; therefore, the input for IMF case in all three optimization methods was the combination of these three indices in each lead-time (Eq. 7).

$$P_n = f(\text{SOI}_n, \text{Nino3.4}_{n-1}, \text{Nino3.4}_{n-2}, \text{Nino3.4}_{n-3}, \text{Nino3.4}_{n-4}, \text{Nino3.4}_{n-5}, \text{Nino3.4}_{n-6},$$
$$\text{Nino3.4}_{n-7}, \text{Nino3.4}_{n-8}, \text{Nino3.4}_{n-9}, \text{Nino3.4}_{n-10}, \text{Nino4}_{n-11}, \text{Nino4}_{n-12}) \qquad (7)$$

RMSE, MAE, $Z$-test, and Taylor diagram were used to identify the most appropriate model among all 36 outcomes. The results of all three optimization methods through CMF and IMF are presented in Figure 5, in which all lead-time can be visually compared. $x$-axis shows the cases used in this study (12 cases from simultaneous to 11-month lead-time in CMF and 1 case in IMF). According to Figure 5a, the best proposed monthly model in BP-based MLP according to its error values belonged to 3-month lead-time in CMF. In this case, RMSE and MAE were 19.37 and 12.38 mm respectively, both of which were lower than those of the other cases. The number of its hidden layer neurons was 8 concluded on the basis of trial and error. Except for this case, IMF outperformed other cases in CMF, proving its prediction ability.
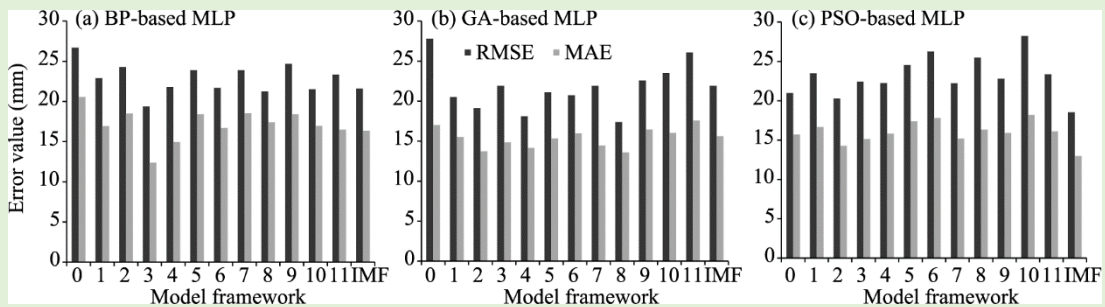


**Fig. 5** Performance of all generated cases in three optimization methods. The 0–11 of $x$-axis shows the framework of CMF (concurrent model framework). IMF, integrated model framework. MAE, mean absolute error; RMSE, root mean square error.

The same information related to GA-based MLP model is shown in Figure 5b. It should be noted that this algorithm was applied as the complete substitution in the training stage in the MLP network. The performance of this algorithm is highly dependent on gamma and mutation rate (0.2 and 0.1, respectively). Case model, 8-month lead-time, seemed quite satisfactory according to its RMSE (17.4 mm) and MAE (13.6 mm). It can be highlighted that for GA-based MLP, in contrast with BP-based MLP, there is not any remarkable superiority sign of IMF over other CMF. Figure 5c indicates PSO-based MLP results. As stated earlier, cognitive and social components are the key parameters in the performance of the PSO algorithm, both of which based on error and trial were considered, respectively, as 1 and 2 in this study. According to Figure 5c, IMF (RMSE, 18.5 mm; MAE, 12.9 mm) was the case that the results remarkably outperformed the other cases, while 2-month lead-time seemed to be more accurate throughout other cases in CMF. The results of $Z$-test for all selected cases in three methods are presented in Table 2.

**Table 2** $Z$-test for selected cases in each method

| Time scale | Model framework | Lead-time (month) | $Z$ |
|---|---|---|---|
| Monthly | CMF | 3 | 0.84 |
| Monthly | CMF | 8 | 0.62 |
| Monthly | IMF | - | 0.56 |

Note: CMF, Concurrent Model Framework; IMF, Integrated Model Framework; -, no lead-time in IMF.

As the critical $Z$ in 5% significance level in two-tailed case is 1.94, which is greater than the calculated $Z$ (Table 2) in all selected cases, the variation between the means of observed and estimated values was not derived to be notable; therefore, the results of these models could be considered as reliable. The comparison between BP-based MLP, GA-based MLP and PSO-based MLP, respectively, in 3- and 8-month lead-time and IMF framework are demonstrated in Table 3,

which clearly expresses the vagueness of distinguishing the best model among the three final selected cases based on RMSE or MAE. Since the judgment on the most accurate performance might be challenging due to some close results, Taylor diagram could help with interpreting. This diagram, under consideration of RMSE of the cases, standard deviation of the simulated datasets and correlation coefficient between the simulated and observed datasets (Taylor, 2001), precisely presents the best-generated model (Fig. 6). The reference point on *x*-axis, which indicates the standard deviation of the observations, was 23.54 mm. According to Figure 6 the PSO-based MLP was the best case, as it was the closest one to the reference point.

**Table 3** Selected models and their RMSE (root mean square error) and MAE (mean absolute error) under training, validation and test steps

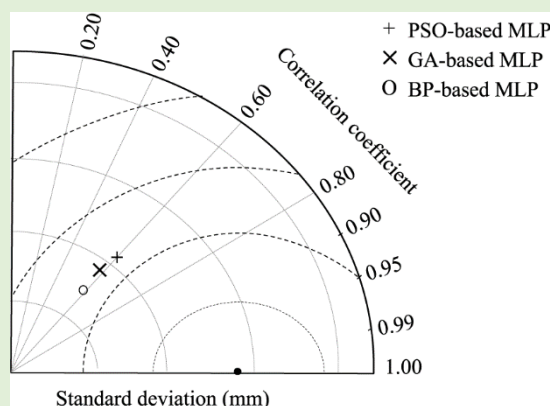| Method | Model framework | Lead-time (month) | Training | | Validation | | Test | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | | | RMSE | MAE | RMSE | MAE | RMSE | MAE |
| | | | (mm) | (mm) | (mm) | (mm) | (mm) | (mm) |
| BP-based MLP | CMF | 3 | 18.53 | 12.10 | 20.43 | 14.90 | 19.37 | 12.38 |
| GA-based MLP | CMF | 8 | 19.34 | 14.61 | 20.36 | 14.26 | 17.39 | 13.59 |
| PSO-based MLP | IMF | - | 21.21 | 15.12 | 19.17 | 13.69 | 18.54 | 12.97 |



**Fig. 6** Result of Taylor diagram. As can be seen, the performance of all cases are in an approximately same range; however, the MLP-PSO is the closest case to the reference point.

## 5    Conclusions

In this study, three optimization methods, BP, GAs and PSO, on the platform of a multilayer perceptron (MLP) network were considered. These methods were applied in the training stage of the MLP and in turn, were completely substituted. Two frameworks, CMF and IMF, were denoted. Through these two frameworks, 13 cases were generated with 12 cases within CMF, each of which contained all selected LSCIs in the same lead-times, and one case within IMF that was constituted from the combination of the most correlated LSCIs with Tehran precipitation in each lead-time. In each optimization method, the most accurate case, based on the RMSE and MAE, was recognized and then assessed by the application of *Z*-test. Since the RMSE and MAE of the selected cases were close to each other, it might be concluded that all three cases performed in the same accuracy; however, through the Taylor diagram, it could be denoted that PSO-based MLP in IMF had the best performance for its lowest distance to the reference point in the diagram. Therefore, Equation 7, as the input combination under PSO-based MLP algorithm, can be applied for precipitation forecasting reliably in Tehran metropolitan. The overall results, however, can be interpreted that in all optimization methods the IMF could conclude more precise outcomes and outperform CMF.

In previous studies, the optimization methods can promote the performance of a multilayer perceptron network (Choubin et al., 2014; Tskiran et al., 2015; Prabhu and Karthikeyan, 2018;

Qui et al., 2018; Bensingh et al., 2019). The results of this study were in agreement with Choubin et al. (2014), which suggested that PSO algorithm could evolve a multilayer perceptron network to generate better results. Although, the errors, resulting in this research, seemed considerable, the selected model could track the future trends more reliably rather than predicting the exact amounts of precipitation (Bjerknes, 1969; Jones et al., 1997; Matyasovszky, 2003; Karabok et al., 2005; Hatzaki et al., 2007; Kakapour, 2011; Ashrafi et al., 2012; Choubin et al., 2014). Despite that, the present study with the application of machine learning reached an optimum forecast model, in which up to 34% of Tehran monthly precipitation could be explained. As a conclusion, it is recommended to take surface parameters coupled with LSCIs under consideration to investigate the probability of further uncertainty reduction.

## References

Abbot J, Marohasy J. 2018. Forecasting of medium-term rainfall using artificial neural networks: case studies from eastern Australia. doi: 10.5772/intechopen.72619. [2017-12-27]. https://www.intechopen.com.

Abdi H, Williams L J. 2010. Principal component analysis. Wiley Interdisciplinary Reviews: Computational Statistics, 2: 433–459.

Allan R J, Beard G S, Close A, et al. 1996. Mean sea level pressure indices of the El Nino-Southern oscillation: relevance to stream discharge in south-eastern Australia. Divisional report. Canberra: CSIRO Division of Water Resources, 96/1.

Araghinejad S, Meidani E. 2013. A review of climate signals as predictors of long-term hydro-climatic variability. Climate Variability. doi: http://dx.doi.org/10.5772/56790.

Arvin A. 2015. Relationship between El-Nino-southern oscillation (ENSO) and total ozone variations in Iran. Geography and Development Iranian Journal, 12: 165–180. (In Farsi)

Ashrafi K, Shafiepour M, Ghasemi L, et al. 2012. Prediction of climate change induced temperature rise in regional scale using neural network. International Journal of Environmental Research, 6(3): 677–688.

Bensingh R J, Machavaram R, Boopathy S R, et al. 2019. Injection molding process optimization of a bi-aspheric lens using hybrid artificial neural networks (ANNs) and particle swarm optimization (PSO). Measurement, 134: 359–374.

Bjerknes J. 1969. Atmospheric teleconnections from the equatorial Pacific. Monthly Weather Review, 97(3): 163–172.

Brandimarte L, di Baldassarre G, Bruni G, et al. 2011. Relation between the North-Atlantic oscillation and hydroclimatic conditions in Mediterranean areas. Water Resource Management, 25: 1269–1279.

Bratton D, Kennedy J. 2007. Defining a standard for particle swarm optimization. Proceedings of the 2007 IEEE Swarm Intelligence Symposium, 120–127, doi: 10.1109/SIS.2007.368035.

Canon J, Gonzalez J, Valdez J. 2007. Precipitation in the Colorado River basin and its low frequency associations with PDO and ENSO signals. Journal of Hydrology, 333(2–4): 252–264.

Choubin B, Khalighi-Sigaroodi S, Malekian A, et al. 2014. Drought forecasting in a semi-arid watershed using climate signals: a neuro-fuzzy modeling approach. Journal of Mountain Science, 11: 1593–1605.

Degefu M A, Bewket W. 2017. Variability, trends, and teleconnections of stream flows with large-scale climate signals in the Omo-Ghibe river basin, Ethiopia. Environmental Monitoring and Assessment, 189(4): 142.

Garro B, Vazquez R. 2015. Designing artificial neural networks using particle swarm optimization algorithms. Computational Intelligence and Neuroscience, ID 369298.

Gaughan A E, Waylen P R. 2012. Spatial and temporal precipitation variability in the Okvango-Kwando-Zambezi catchment, southern Africa. Journal of Arid Environments, 82: 19–30.

Gerkaninezhad M Z, Bazrafshan O. 2018. Impact of climatic signals on the wet and dry season precipitation (case study: Persian Gulf and Oman Sea watersheds). Journal of the Earth and Space Physics, 44: 333–349. (In Farsi)

Ghazal R, Ardeshir A, Zahedi Rad I, 2014. Climate change and storm-water management strategies in Tehran. Procedia Engineering, 89: 780–787.

Gong D, Ho C. 2003. Detection of large-scale climate signals in spring vegetation index (normalized difference vegetation index) over the Northern Hemisphere. Journal of Geophysical Research, 108(D16): 4498.

Hatzaki M, Flocas H, Asimakopoulos D, et al. 2007. The eastern Mediterranean teleconnection pattern. International Journal of Climatology, 27(6): 727–737.

Hidalgo H, Dracup J. 2003. ENSO and PDO effects on hydroclimatic variations of the upper Colorado River basin. Journal of Hydrometeorology, 4(1): 5–23.

Hurrell J W. 1995. Decadal trends in the north Atlantic oscillation: regional temperatures and precipitation. Science, 269(5224): 676–679.

Jiang M, Luo Y, Yang S. 2007. Particle swarm optimization-stochastic trajectory analysis and parameter selection. In: Felix T S

C, Tiwari M K. Swarm Intelligence, Focus on Ant and Particle Swarm Optimization. doi: 10.5772/5104.

Jolliffe I T. 2002. Principal Component Analysis (2$^{nd}$ ed.). New York: Springer, 2.

Jones P, Jonsson T, Wheeler D. 1997. Extension to the north Atlantic oscillation using early instrumental pressure observations from Gibraltar and south-west Iceland. International Journal of Climatology, 17(13): 1433–1450.

Kaiser H. 1960. The application of electronic computers to factor analysis. Educational and Psychological Measurements, 20: 141–151.

Kampichler C, van Turnhout C, Devictor V, et al. 2012. Large-scale changes in community composition: determining land use and climate change signals. PLoS ONE, 7(4): e35272.

Karabok M, Kahya E, Karaca M. 2005. The influences of the Southern and North Atlantic oscillations on climatic surface variables in Turkey. Hydrological Processes, 19(6): 1185–1211.

Kriesel D. 2007. A Brief Introduction to Neural Networks. [2020-01-20]. http://www.dkriesel.com.

Mann P. 1997. Introductory Statistics (3$^{rd}$ ed.). New York: Wiley, 405.

Matyasovszky I. 2003. The relationship between NAO and temperature in Hungary and its nonlinear connection with ENSO. Theoretical and Applied Climatology, 74: 69–75.

Mitchell M. 1996. An Introduction to Genetic Algorithms. Cambridge: MIT Press, 3.

Nigam S, Shen H. 1993. Structure of oceanic and atmospheric low-frequency variability over the tropical Pacific and Indian Oceans. Journal of Climatology, 6(4): 657–676.

Oldenberg-van G, Burgers G, Tank A. 2000. On the El-Nino teleconnection to spring precipitation in Europe. International Journal of Climatology, 20(5): 565–574.

Ouyang R, Liu W, Fu G, et al. 2014. Linkages between ENSO/PDO signals and precipitation, stream-flow in China during the last 100 years. Hydrology and Earth System Science, 18(9): 3651–3661.

Pasini G. 2017. Principal component analysis for stock portfolio management. International Journal of Pure and Applied Mathematics, 115(1): 153–167.

Popescu M, Balas V, Popescu L, et al. 2009. Multilayer perceptron and neural networks. WSEAS Transactions on Circuits and Systems, 8(7): 579–588.

Pozo-Vazquez D, Gamiz-Fortis S R, Tovar-Pescador J, et al. 2005. El Nino-Southern oscillation events and associated European winter precipitation anomalies. International Journal of Climatology, 25(1): 17–31.

Prabhu M V, Karthikeyan R. 2018. Comparative studies on modelling and optimization of hydrodynamic parameters on inverse fluidized bed reactor using ANN-GA and RSM. Alexandria Engineering Journal, 57(4): 3019–3032.

Qui S, Chen B, Wang R, et al. 2018. Atmospheric dispersion prediction and source estimation of hazardous gas using artificial neural network, particle swarm optimization and expectation maximization. Atmospheric Environment, 178: 159–163.

Rosenblatt F. 1961. Principles of Neurodynamics: Perceptron and the Theory of Brain Mechanisms. Washington DC: Spartan Books, 245.

Saji N, Goswami B, Vinayachandran P, et al. 1999. A dipole mode in the tropical Indian Ocean. Nature, 401: 360–363.

Santos J, Corte J, Leite S. 2005. Weather regimes and their connection to the winter rainfall in Portugal. International Journal of Climatology, 25(1): 33–50.

Seiffert U. 2001. Multiple layer Perceptron training using genetic algorithms. Bruges: Proceedings of European Symposium on Artificial Neural Networks, 159–164.

Srinivasan D, Seow T H. 2003. Particle swarm inspired evolutionary algorithm (PS-EA) for multi-objective optimization problem. Acta Biomaterialia, 4: 2292–2297.

Taylor K E. 2001. Summarizing multiple aspects of model performance in a single diagram. Journal of Geophysical Research, 106(D7): 7183–7192.

Tyson P. 1987. Climate change and variability in southern Africa. The Quarterly Journal of Royal Meteorological Society, 8: 552–562.

Wallace J, Gutzler D. 1981. Teleconnections in the geopotential height field during the northern hemisphere winter. Monthly Weather Review, 109(4): 748–812.

Whitley D. 1994. A genetic algorithm tutorial. Statistics and Computing, 4: 65–85.

Wittan I H, Frank E. 2005. Data Mining: Practical Machine Learning Tools and Techniques (2$^{nd}$ ed.). San Francisco: Elsevier, 29.

Xu L, Chen N, Zhang X. 2018. A comparison of large-scale climate signals and the North American multi-model ensemble (NMME) for drought prediction in China. Journal of Hydrology, 557: 378–390.

Xu Z, Hou Z, Han Y, et al. 2016. A diagram for evaluating multiple aspects of model performance in simulating vector fields. Geoscientific Model Development, 9: 4365–4380.

Zahraei B, Karamouz M. 2004. Seasonal precipitation prediction using large scale climate signals. Salt Lake City: World Water and Environmental Resources Congress.