Research Article

Forecasting Different Types of Droughts Simultaneously Using Multivariate Standardized Precipitation Index (MSPI), MLP Neural Network, and Imperialistic Competitive Algorithm (ICA)

Pouya Aghelpour and Vahid Varshavian

Agricultural Meteorology, Department of Water Engineering, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran

Correspondence should be addressed to Vahid Varshavian; v.varshavian@basu.ac.ir

Received 11 October 2020; Revised 8 November 2020; Accepted 29 December 2020; Published 19 January 2021

Academic Editor: Shamsuddin Shahid

Copyright © 2021 Pouya Aghelpour and Vahid Varshavian. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Precipitation deficit causes meteorological drought, and its continuation appears as other different types of droughts including hydrological, agricultural, economic, and social droughts. Multivariate Standardized Precipitation Index (MSPI) can show the drought status from the perspective of different drought types simultaneously. Forecasting multivariate droughts can provide good information about the future status of a region and will be applicable for the planners of different water divisions. In this study, the MLP model and its hybrid form with the Imperialistic Competitive Algorithm (MLP-ICA) have been investigated for the first time in multivariate drought studies. For this purpose, two semi-arid stations of western Iran were selected, and their precipitation data were provided from the Iranian Meteorological Organization (IRIMO), during the period of 1988–2017. MSPI was calculated in 5-time windows of the multivariate drought, including MSPI\_3–6 (drought in perspectives of soil moisture and surface hydrology simultaneously), MSPI\_6–12 (hydrological and agricultural droughts simultaneously), MSPI\_3–12 (soil moisture, surface hydrology, and agricultural droughts simultaneously), MSPI\_12–24 (drought in perspectives of agriculture and groundwater simultaneously), and MSPI\_24–48 (socio-economical droughts). The results showed acceptable performances in forecasting multivariate droughts. In both stations, the larger time windows (MSPI\_12–24 and MSPI\_24–48) had better predictions than the smaller ones (MSPI\_3–6, MSPI\_6–12, and MSPI\_3–12). Generally, it can be reported that, by decreasing the size of the time window, the gradual changes of the index give way to sudden jumps. This causes weaker autocorrelation and consequently weaker predictions, e.g., forecasting droughts from the perspective of soil moisture and surface hydrology simultaneously (MSPI\_3–6). The hybrid MLP-ICA shows stronger prediction results than the simple MLP model in all comparisons. The ICA optimizer could averagely improve MLP’s accuracy by 28.5%, which is a significant improvement. According to the evaluations (RMSE = 0.20; MAE = 0.15; R = 0.95), the results are hopeful for simultaneous forecasting of different drought types and can be tested for other similar areas.

1. Introduction

Drought, one of the most complex environmental catastrophes, continuously has an effect on the rest of the world [1]. It occurs naturally in all climatic areas, such as pluvial and arid areas, and causes many economic, environmental, and social costs around the world [1–4]. In recent decades, drought has been one of the costliest natural disasters that has created major challenges in Iranian water resource management. The arid and semiarid climate of Iran has made it highly vulnerable to droughts [5]. During 1998–2000, Iran experienced one of the worst and most damaging drought periods in the last 50 years [6]. During this 3-year period, water shortage in more than 270 cities fell down to below the critical point, and as a result, thousands of villages lost their drinking water, surface water flow decreased to 55%, and Iran’s dams and tanks were forced to act with minimal capacity for water transfer because of low input flow and high temperature [6, 7]. So, during this period, the country faced different types of droughts, such as...
Drought prediction is a major concern for water managers, farmers, and other final users because it limits their decisions. Since droughts have slowly begun, it is possible to present temporal forecasts in order to take measures and develop policies to reduce the effects of droughts [8–10]. A wide range of artificial intelligence (AI) models and modified standalone and hybrid versions have been used for the forecasting of different drought indices. The studies revealed the higher performance of AI models in forecasting drought indices [11–17]. In fact, the AI models can predict the drought events that do not have a good and straightforward mathematical solution and were proven to have the ability to capture the white noise, nonstationary, and nonlinearity in the time series [18]. Multilayer Perceptron (MLP) neural network is the most famous type of AIs which has been widely used in hydrological and meteorological modeling studies [19–30]. Malik and Kumar [31] used the MLP model for meteorological drought prediction based on Effective Drought Index (EDI) in the Uttarakhand state of India and reported the acceptable performance of this model. MLP is also used for predicting the Standardized Precipitation Index (SPI) in Iran as a meteorological drought indicator and was superior compared to the other models such as Adaptive Neuro-Fuzzy Inference System (ANFIS), Radial Basis Function Neural Network (RBFNN), and Support Vector Machine (SVM) [32]. In forecasting agricultural drought based on the Standardized Precipitation-Evapotranspiration Index (SPEI), the MLP model was reported as an acceptable predictor model in Pakistan [33]. This model has been well evaluated for predicting SPI as a meteorological drought index in Awash river basin in Ethiopia [34], Selangor river basin in Malaysia [35], and Santa Ysabel Creek and Leaf rivers in America [36]. The study of Borji et al. [37] based on the Streamflow Drought Index (SDI) confirms the ability of the MLP model for drought forecasting from the perspective of surface hydrology too.

In the abovementioned studies, the referred indices can monitor the different drought types separately, for example, SPI individually for meteorological droughts, SDI individually for hydrological droughts [38], and SPEI for agricultural droughts. Among the drought indices, SPI is a different index that can indicate different types of droughts in its different time windows [39]. For example, the 1-month SPI (SPI1) shows the meteorological drought condition. Also, 3-month, 6-month, 12-month, or 24-month SPI (SPI3, SPI6, SPI12, or SPI24) talk about droughts in perspectives of soil moisture, surface hydrology, agriculture, and groundwater, respectively [12, 40]. According to the literature, no model has been evaluated in forecasting different drought types simultaneously. Multivariate drought indices can theoretically monitor different types of droughts simultaneously. So, forecasting a multivariate drought index can theoretically present a perspective about the status of different drought types in the future. Multivariate Standardized Precipitation Index (MSPI) is one of the newest introduced multivariate drought indices. This index was developed by Bazrafshan et al. [1] based on SPI by merging SPI’s different time windows. Also, MSPI was superior in comparison with the Joint Deficit Index (JDI) [41] to monitor multivariate drought in all of the Iranian climates [6]. Up to now, MSPI was used in agricultural drought studies. Bateni et al. [42] used MSPI to develop an agrometeorological drought index. Aghelpour et al. [12] evaluated MSPI in monitoring and forecasting agricultural drought for Iranian climates and found the index acceptable for the issue.

According to the literature, no investigation was carried out on MSPI for different drought types simultaneously (multivariate drought forecasting). Due to MLP’s high capability in drought forecasting and, generally, hydrological modeling studies, this model is used in the current study for forecasting MSPI for multivariate drought forecasting. Two semiarid climate stations located in western Iran have been selected for this issue. Also, in numerical modeling issues of hydrology, the optimization algorithms such as Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and Firefly Algorithm (FA) have been well used recently to optimize the prediction/estimation accuracies of the MLP model. The imperialistic Competitive Algorithm (ICA) is another powerful optimization algorithm developed by Atashpaz-Gargari and Lucas [43], which has been used less in hydrological modeling cases, especially in drought forecasting cases. Therefore, as another innovation in drought studies, in this study, the ICA is merged with MLP to improve its forecasting accuracy.

2. Materials and Methods

2.1. Study Area. According to the extended De-Martonne climatic classification, about 22.91% of the total area of Iran located in a semiarid climatic condition is equal to 377533.4 km² [44]. These semiarid regions receiving appropriate solar radiation are important areas for agricultural production, but they are highly dependent on available water and affected by incoming rainfall and consequent drought events, which directly and indirectly affect various aspects of human life. In fact, precipitation deficit and consequent drought events can be more effective in semiarid regions than humid areas. Therefore, the current study aims to investigate and predict the multivariate drought condition for two ground station about Iran’s semiarid climate. Hamedan and Kermanshah are located in western Iran. Hamedan and Kermanshah having urban area of 19368 and 24998 square kilometers are elevated with 1820 meters and 1400 meters, respectively (Figure 1).

Calculating MSPI needs total monthly precipitation data. The precipitation data was provided from the Iranian Meteorological Organization (IRIMO) for these synoptic stations and were used for MSPI calculation. The stations’ geographical coordinates and statistical characteristics of their monthly precipitation data are shown in Table 1.
2.2 Multivariate Standardized Precipitation Index.

Multivariate standardized precipitation index (MSPI) is derived by applying Principal Components Analysis (PCA) on the $K$ timeseries of SPI (where $K$ is corresponding to the time window $K$). The first major component ($PC_1$) of the PCA analysis is used to describe the percentage changeability in $K$ of the initial variable. Due to the $PC_1$ character, unlike the SPI, which has a mean 0 and standard deviation 1, its values are not comparable between months. Therefore, it is necessary to standardize the $PC_1$ time series using the average and standard deviation of the different months of the year using [1].

---

Table 1: Coordinates of the stations; descriptive statistics of the monthly precipitation time series; climate type based on the extended De-Martonne classification method [44].

| Station     | Latitude (°) | Longitude (°) | Elevation (m) | Mean (mm) | St. Dev (mm) | Min (mm) | Max (mm) | Skew. | Period       | Climate                  |
|-------------|--------------|---------------|---------------|------------|--------------|----------|----------|-------|--------------|--------------------------|
| Hamedan     | +30.87       | +48.53        | +1741.5       | 25.07      | 26.62        | 0.00     | 186.10   | 1.43  | 1988–2017    | Semiarid very cold        |
| Kermanshah  | +34.35       | +47.15        | +1318.6       | 34.21      | 40.05        | 0.00     | 295.40   | 1.63  | 1988–2017    | Semiarid cold             |

---

Figure 1: The geographical location of Hamedan and Kermanshah provinces.
Z_{1ym} = \frac{PC_{1ym} - PC_{1m}}{SD_{1m}} = \frac{PC_{1ym}}{SD_{1m}}, \quad (1)\

where \( Z_{1ym} \) is the standardized value of PC1 in year (y) and month (m), \( PC_{1ym} \) is the value of PC1 in year (y) and month (m), \( PC_{1m} \) is the average PC1 in the month (m), SD_{1m} is the standard deviation of PC1 in month (m), and \( Z_{1ym} \) is taken as the multivariate Standard Precipitation index (MSPI). The value of \( PC_{1m} \) is statistically very small and close to zero; hence, in the case of the fraction of the above equation, it can be neglected [1, 45]. In order to determine the drought severity classes of the MSPI, the time series is arranged in ascending order, and its probability distribution is plotted on a diagram (for example, Figure 2).

Then, the values corresponding to the probability thresholds of different classes of SPI (Table 2) are extracted from the above diagram. The extracted values are taken as the MSPI thresholds, which can be used in the MSPI time series classification for drought severity [1].

According to the main literature of MSPI [1], MSPI can simultaneously monitor different droughts by its five time windows; so, these time windows are getting calculated in the current study. The time windows included are 3–6 months (MSPI_{3–6}), 6–12 months (MSPI_{6–12}), 3–12 months (MSPI_{3–12}), 12–24 months (MSPI_{12–24}), and 24–48 months (MSPI_{24–48}). The mentioned time windows of MSPI are its multivariate drought time windows and theoretically related to more than one type of droughts. SPI’s time windows show the different drought types’ severity and weakness. Extensionally, the time windows SPI_{3}, SPI_{6}, SPI_{12}, and SPI_{24} relate to droughts in perspectives of soil moisture, surface hydrology, agriculture, and groundwater level, respectively, and the larger time windows can be related to the economic and social effects of drought [1, 12, 40]. Therefore, the merged form of these time windows of MSPI can theoretically monitor the different perspectives of droughts simultaneously. For example, the MSPI_{3–6} time window can theoretically monitor drought in perspectives of soil moisture and surface hydrology or MSPI_{6–12} can monitor both hydrological and agricultural droughts, simultaneously. MSPI_{3–12} relates to 3 perspectives of droughts simultaneously, including soil moisture, hydrological, and agricultural droughts. Among the larger time windows, MSPI_{12–24} can simultaneously indicate drought in perspectives of agriculture and groundwater, and MSPI_{24–48} is theoretically related to the socio-economic effects of droughts [1, 6, 12, 39, 47].

2.3. Artificial Neural Networks. An artificial neural network is a parallel information processing system that has a distinct function inspired by the biological structure of the human brain [48]. These systems are able to determine the complexity and nonlinear relationship between the inputs and outputs of a physical system by a network of nodes that are interconnected. In these systems, the activity of each of these connections is set by historical information (learning process).

2.3.1. MLP Neural Network. MLP network is one of the most important structures of artificial neural networks. These networks consist of the layers of sensory units (neurons), the input layer made of one or more hidden layers, and the output layer. The input signal is transmitted through the network in the forward direction to the hidden layer and then to the output layer [48]. The output of each neuron is multiplied by weight coefficients and given as input to a nonlinear excitation function. In the training phase, the training data is given to the perceptron, and then, the grid weights are adjusted to minimize the error between the target and output of the model or to reach the number of training times to the default value. Then, like all modeling processes, different inputs (not present in the training phase) are used for model validation. The training of neural networks is generally very complex and can be stated to be an optimization problem with a large number of variables [49].

2.3.2. Imperialistic Competitive Algorithm (ICA). The imperialistic competitive algorithm was first proposed by Atashpaz-Gargari and Lucas in 2007 [43]. This algorithm, in the first place, with a completely new perspective on optimization, establishes a new link between the humanities and social sciences on the one hand and the technical and mathematical sciences on the other. In particular, this algorithm views the process of colonization as a stage of human socio-political evolution and uses mathematical
modeling as the source of inspiration for a powerful optimization algorithm. With the formation of the early empires, the imperialist rivalry between them began. Any empire that fails to compete for colonial power and increase its power (or at least prevent it from losing its influence) will be removed from the arena of imperialistic competition. Thus, the survival of an empire will depend on its power to attract and control the colonies of rival empires. As a result, during the imperialists’ competitions, the power of the larger empires gradually increased, and the weaker empires were eliminated. To increase their power, empires will have to develop their colonies as well. Over time, the colonies will become closer to the empires in terms of power, and we will see a kind of convergence. The final limit of imperialistic competition is when there is a single empire in the world with colonies that are very close to the imperialist country in terms of position. In the following sections, different parts of the algorithm are presented [43].

Like other evolutionary algorithms, this algorithm starts with several random primary populations, in which each of them is called a "country." A number of the best elements of the population (equivalent to elitism in genetic algorithm or particle in particle swarm optimization) are selected as imperialists. The rest of the population is also considered a colony. Depending on their power, the colonizers will specially colonize these colonies; they pull towards themselves. The total power of an empire depends on both its constituent parts and the imperialist country (as the central nucleus) and its colonies. In mathematical terms, this dependence is modeled by defining the power of the empire as the total power of the imperial state, plus a percentage of the average power of its colonies. For supplementary information about this algorithm and its mathematical equations, the references are suggested [43, 50, 51]. The learning processes of the ICA algorithm is presented step by step in the form of a flowchart (Figure 3).

2.3.3. Combining MLP with Imperialist Competitive Algorithm (MLP-ICA). As a metainnovative neural network, the ICA can be merged with MLP to improve MLP’s modeling accuracy. This approach gets done by optimizing the parameters of MLP by ICA. The makeup of MLP depends on the hidden layers, neurons, and transfer functions, which are found by the trial and error method. The optimizable parameters are the weights and biases of the MLP neural network that these two subjects are optimized by ICA optimization. The schematic form of merging MLP with ICA (MLP-ICA) is shown in Figure 4.

2.4. Evaluation Measures. To ensure the accuracy of modeling and predicting, the outputs of the model should be compared with their actual values. For this purpose, the model performance evaluation criteria are used. The criteria used in this study are the Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Pearson correlation coefficient (R), whose equations (2)–(4) have been described below:

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - f_i)^2},
\]

\[
MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - f_i|,
\]

\[
R = \frac{\sum_{i=1}^{N} (y_i - \bar{y})(f_i - \bar{f})}{\sqrt{\sum_{i=1}^{N} (y_i - \bar{y})^2} \cdot \sqrt{\sum_{i=1}^{N} (f_i - \bar{f})^2}}
\]

In the above equations, \(y_i\) and \(\bar{y}\) are the observational data and their mean and \(f_i\) and \(\bar{f}\) are the models’ predictions and their mean, respectively, and \(n\) is the number of data, respectively. The values closer to zero for RMSE and MSE criteria and the values close to 1 for the Pearson correlation coefficient represent the optimal performance of the model. In this research, coding in MATLAB software has been used to run the MLP and MLP-ICA models.

3. Results

3.1. Input Selection. According to the main reference [1], the MSPI was calculated in five time windows including MSPI\(_{3-6}\), MSPI\(_{6-12}\), MSPI\(_{12-24}\), MSPI\(_{12-24}\) and MSPI\(_{24-48}\) which resulted 5 time series for each station. For predicting the index, the model’s input must be the previous amounts of the index (time lags of MSPI). For this, the Autocorrelation Function (ACF) plot was used to determine the correlation signification of the time lags (Figure 5).

ACF plots (Figure 5) can show the significance of the variable’s correlation with its time lags, and due to this, the significantly correlated time lags can be used as the models’ inputs for prediction. Figure 5 shows the ACF plots of Hamedan station, drawn for the 5 mentioned time windows of MSPI. As it is observable in the ACF plot of MSPI\(_{3-6}\), the time lags of 1, 2, and 3 have significant correlations, and so, they are selected as the predictor inputs. In both MSPI\(_{12-24}\) and MSPI\(_{12-24}\) time windows, the 1, 2, 3, 4, and 5 time lags are significantly correlated. For MSPI\(_{12-24}\), the significant time lags are 1, 2, ..., 12, and for MSPI\(_{24-48}\) time window, the significant correlations belong to the 1, 2, ..., 17 time lags. This method was similarly implemented for Kermanshah station, and the total results are shown in Table 3.

3.2. Predicting Multivariate Drought and Assessment. After selecting the input variables for each of the time windows, the input-target samples were divided into two phases: 75% for the training phase and 25% for the testing phase. Data with specified inputs were entered into the two artificial intelligence models for modeling and evaluation. The parameters of the MLP neural network model includes the number of hidden layers, number of neurons in hidden layers, and type of transfer functions within the neurons, which were selected by the trial and error method. Consequently, the best suitable number of hidden layers was up to 3 hidden layers (1, 2, and 3 hidden layers), best number of neurons was from 6 neurons to 18 neurons in each hidden
layer, and best-fitted transfer function was the saturating linear transfer function (satlin). In the MLP-ICA model, the weights and biases of MLP (with similar makeup) were optimized by the ICA algorithm. The evaluation is carried out by the evaluation criteria of RMSE, MAE, and $R$, and the results are shown in Table 4.

In Table 4, the models are evaluated separately for stations and time windows. The predictive section is actually
Multilayer perceptron neural network

Imperialistic competitive algorithm

MSPIₜ, MSPIₜ₋₁, MSPIₜ₋₂, ..., MSPIₜ₋ₙ

Inputs Modeling phase Outputs

Figure 4: The interaction of the MLP model with Imperialist Competitive Algorithm in a schematic structure.

Figure 5: ACF plots of the five time windows of MSPI for Hamedan station.
The test section discussed here. At first glance, it is clear that, in both stations, as the time windows get bigger, the accuracy of the models also increases. In both stations, the accuracy of the models from low to high is reported in the time windows MSPI3–6, MSPI3–12, MSPI6–12, MSPI12–24, and MSPI24–48, respectively. In all cases, the comparison between the two models, which includes 10 series (5 time windows for each of the stations), and the MLP-ICA hybrid model has higher accuracy in predicting MSPI. This shows that ICA is effective in optimizing the MLP model and is able to improve its performance. The lowest percentage of MLP performance improvement by ICA at Hamedan station belongs to the MSPI3–12 time window (about 1.9%) and the highest, which is about 60%, was reported for the MSPI24–48 time window (where the model error was reduced from 0.32 to 0.20). At Kermanshah station, the MLP-ICA hybrid model has performed much better compared to the simple MLP model. In MSPI12–24 and MSPI24–48 time windows, it was able to improve MLP forecasting accuracy by more than 50% (in MSPI12–24, the model error was reduced from 0.32 to 0.20), and in MSPI24–48, the model error was reduced from 0.30 to 0.20. The lowest performance improvement from ICA merging with MLP at Kermanshah station belonged to the MSPI3–6 time window, where the model error was reduced from RMSE = 0.73 to RMSE = 0.64 (about 14.1% improvement in prediction accuracy). It can be said that, on average, the ICA algorithm in Hamedan and Kermanshah stations caused 25% and 32% increase in accuracy for the MLP model, respectively. The strongest and weakest performance in Hamedan station was reported in MSPI24–48 window by the MLP-ICA model (with RMSE = 0.20, MAE = 0.15, and $R = 0.95$) and MSPI3–6 time window by the MLP model, respectively. In Kermanshah station, the best forecast was reported by the MLP-ICA model and MSPI24–48 time window (RMSE = 0.20, MAE = 0.15, and $R = 0.92$) and the weakest was obtained by the simple MLP model in the MSPI3–12 time window (RMSE = 0.80, MAE = 0.60, and $R = 0.60$). These overlays of the models’ predictions and their observations are shown in time series plots (Figure 6).

In the MSPI4–12 time window of Hamedan station, the models have relatively close estimates of the actual value so
(a) Figure 6: Continued.
that, in most of the months, there are overlaps. During the severe drought from May–October 2013 and March–July 2015, both of the models were overestimated but the MLP-ICA’s overestimation is milder so that this model could present an acceptable prediction for the classes of these severe drought months. In the same time window (MSPI₁₂–₂₄) at Kermanshah Station, the forecasts are also suitable, but compared to Hamedan Station, the overlap is somewhat weaker. The error was enormous during the months of January through October 2016. The months are in wet conditions, so this underestimation in forecasting does not cause problems. Another remarkable point in MSPI₁₂–₂₄ of Kermanshah is the severe and extreme drought status during April–November 2015. In this case, MLP was weak and overestimated, but MLP-ICA could have a highly accurate prediction for these important months.

The prediction of models for the MSPI₂₄–₄₈ time window of the two stations was very accurate in most of the months, especially in the diagnosis of the drought classes. At the Hamedan station (Figure 6), the MLP model in all months of 2013 (unlike the MLP-ICA) had a large difference with actual values. However, in the course of May–July 2015 (with a little difference), it could have forecast a long-term severe drought class as well as the MLP-ICA. In this diagram (MSPI₂₄–₄₈ of Hamedan), the proximity of the MLP-ICA prediction curve is clearly visible with the actual value curve compared to the MLP. This comparison between the two models at Kermanshah station, due to the significant difference in the forecast of two more convenient models, shows the superiority of the MLP-ICA hybrid model. The MLP’s error in forecasting the MSPI₂₄–₄₈ time window is more than MLP-ICA in most of the months. However, in March–December 2012, February–September 2016, and April–Nov 2017, it is too clear that MLP-ICA is more accurate than MLP, with a large difference. The ability of the models in predicting drought severity classes of MSPI has been evaluated in Table 5. In this table, the number of months of the test period that predicted drought severity correctly was divided by the total number of test periods. This method is done for both models in each of the time windows of both stations, and the resulted likelihood is shown by percentage.

Observing all of the comparison cases (a specific time window in a station) between the two models shows the superiority of MLP-ICA’s likelihood against the MLP simple form. The error decreasing by increasing the time window size is reported here (evaluation of the class prediction) too. The least likelihood percentage is reported for MSPI₁₂–₆₀ of Kermanshah station, which is 65.05%, and resulted in the
Table 5: The models’ likelihood in determining MSPI drought classes.

| Time window | Likelihood percentage of the drought severity classes (%) |
|-------------|----------------------------------------------------------|
|             | Hamadan MLP | MLP-ICA | Kermanshah MLP | MLP-ICA |
| MSPI3–6     | 69.90       | 73.79   | 65.08          | 68.93   |
| MSPI6–12    | 68.32       | 77.23   | 70.30          | 75.25   |
| MSPI12–24   | 72.28       | 75.25   | 67.33          | 69.31   |
| MSPI24–48   | 80.00       | 90.00   | 86.67          | 88.89   |
| MSPI3–6     | 79.52       | 86.75   | 95.78          | 96.39   |

MLP model. The largest likelihood percentage is 96.39%, which is reported for the MLP-ICA model in predicting MSPI24–48 of Kermanshah station. This can show the good capability of AI methods in forecasting the severity classes of different types of droughts simultaneously. The improvement of drought class prediction by the new hybrid model (MLP-ICA) is clearer in Hamadan station compared to Kermanshah station. The biggest improvement reached by combining ICA with MLP is 10%, which is reported for the MSPI12–24 time window of Hamadan station. This case can make another confirmation on the ability of the new hybrid model, MLP-ICA. In Figure 7 (regression diagrams), the correlation of the output of the two models with actual values is investigated. It is clear that the distribution of points around the axis of the regression line is greater in smaller time windows, and in larger time windows, the points are closer and more concentrated to the models’ regression lines.

This indicates a higher correlation and therefore a more accurate prediction in larger time windows, which is consistent with the results in Table 4. At Hamadan station, MSPI24–48 forecasts have the highest concentration around their fitted regression lines and as a result, have the highest correlation between the time windows. The $R^2$ coefficients in these series for the MLP and MLP-ICA models are 0.827 and 0.928, respectively, which shows a very good correlation in predicting socio-economic droughts. The MSPI3–6 time window, which represents drought from the perspectives of soil moisture and surface hydrology simultaneously, had the weakest forecasts in Hamadan ($R^2$ is equal to 0.605 and 0.654 for MLP and MLP-ICA, respectively). The reason for this difference in accuracy can be the sudden jumps in the time series of the index. Sudden jumps occur more frequently among smaller time windows, but in larger time windows that reflect the long-term effects of lack of precipitation, monthly changes are gradual (see Figure 6 for a better understanding). The index with gradual changes will always have a more accurate prediction, while the presence of sudden jumps in the time series naturally increases the prediction error. Among the time windows smaller than 12 months (MSPI3–6, MSPI6–12, and MSPI12–24), the MSPI6–12 time window, which is associated with both hydrological and agricultural droughts, had the strongest predictions at both stations. From the theoretical point of view, the MSPI6–12 time window simultaneously represents 3 types of drought perspectives (drought from soil moisture perspective, hydrological drought, and agricultural drought), while MSPI12–24 simultaneously represents two types of droughts (hydrological drought and agricultural drought). Therefore, it is logical that the simultaneous forecasting of two types of droughts will be more accurate than three types of droughts. At Kermanshah station, the best forecasts for the MSPI12–24 time window (which could theoretically be a simultaneous indicator of drought from agricultural and groundwater perspectives) and MSPI24–48 time window were presented. The $R^2$ value of the MLP-ICA output for the MSPI12–24 and MSPI24–48 time windows of Kermanshah are 0.849 and 0.841, respectively. In predicting all cases (all 10 series investigated), the slope of the MLP-ICA regression line is closer to the 1:1 line, and the $R^2$ value is higher than the simple MLP model. This demonstrates the reliable performance of the ICA algorithm in optimizing the MLP model, which results in a more accurate prediction of simultaneous droughts. The highest increased correlation by the MLP-ICA hybrid model is observed in the MSPI3–12 time window of Kermanshah station, where the $R^2$ value has increased from 0.355 to 0.601. Distribution and concentration of the model error around zero can be another measure of predictability. In this section, the violin plot (Figure 8) is used to examine the distribution of prediction errors.

This diagram was drawn separately for each station, and the power of the two models was compared in each time window. The results of this plot also confirm the strength of the MLP-ICA model compared to the simple MLP model. For example, in Kermanshah’s MSPI3–6 time window, the curvature of the MLP-ICA violin around the error = 0 axis is greater than that of the MLP (the so-called wider violin). This indicates a higher percentage of zero errors in MLP-ICA, which shows its superiority against the simple MLP model. This is true of all time windows and is more noticeable in larger time windows. In the MSPI24–48 time window, the width of the MLP-ICA violin is about twice that of the MLP. Therefore, in predicting larger MSPI time windows, the use of the MLP-ICA hybrid model is highly recommended. The situation is similar at Hamadan station, and MLP-ICA violins are wider than MLP, at error = 0 axis. Also, as the size of the time window increases, the width of the violins increases, the distance between the first and third quartiles (IQR) decreases, and the upper and lower tails of the frequency distributions gradually disappear. This indicates that the prediction error is centered on zero, which is minimized in the largest time window (MSPI24–48).

4. Discussion

MSPI forecasts were so far only reported in the Aghelpour et al. [12] study, which used the ANFIS model and its hybrid models. In this study, 31 stations from different climates of Iran have been studied, and MSPI has been monitored and predicted from the perspective of agricultural drought (the time window of MSPI1–12). In Aghelpour et al. [12] study, the Kermanshah station was also examined, which in its most accurate forecast was RMSE equal to 0.356. While in the present study, the RMSE rate for Kermanshah varies between 0.2 and 0.6. This difference is due to the different dimensions
Figure 7: Continued.
Figure 7: Scatter plots to evaluate the correlations between the models’ outputs and their observations in the test period.
of the time windows of the present study (MSPI_{3–6}, MSPI_{6–12}, MSPI_{3–12}, MSPI_{12–24}, and MSPI_{24–48}). This difference also conceptually creates a different perspective for MSPI. In other words, the MSPI_{6–12} time window can only comment on one type of drought (agricultural drought), while the time windows examined in the current study are multivariate. These current time windows of MSPI can discuss several different types of droughts simultaneously. The ICA algorithm has been used in drought studies in only one study of Hosseini-Moghari et al. [50] to predict SPI, which has not studied drought in the multivariate type. Also, the MLP-ICA hybrid model has not been compared to its simple form to evaluate the improvement of the hybrid model.

The reason for the discrepancy between the results obtained in the two stations studied in the current study could be due to their microclimatic and topographic differences. Also, the difference in the type and number of atmospheric systems that affect these areas could be another reason for the difference in the accuracy of forecasts [52]. The Mediterranean low-pressure system approaches Iran from the northwest and the Sudanese low-pressure system approaches from the southwest. Due to having the Zagros Mountains in the west, Iran has favorable conditions for intensifying and expanding rainfall and sometimes flooding. Also, the role of the Red Sea is to provide more precipitation moisture to Kermanshah compared to Hamedan. In fact, the existence and role of higher mountains in the Hamedan region can be the main controlling factor of atmospheric systems and fronts. This orographic factor decreases the effects of the systems and consequently decreases the irregularity in precipitation time series in Hamedan compared to Kermanshah. So, while the MSPI is originated from the precipitation regime, this can cause weaker autocorrelation and finally weaker predictions in Kermanshah compared to Hamedan.

5. Conclusion

The results show that the simultaneous forecasting of different drought types can have acceptable accuracies for semi-arid climates of western Iran. The best performance will be resulted in predicting the MSPI_{24–48} time window, which is related to socio-economic drought. Also, the weakest accuracy belongs to the predictions of the short-term effects of precipitation deficit, such as soil moisture and surface hydrological droughts simultaneously (MSPI_{3–6}). Evaluating the models shows the significant capability of an imperialistic competitive optimization algorithm in improving MLP’s prediction accuracy, which is reported 28.5% on average for the current study. As the first use of the MLP models and its hybrid form MLP-ICA in multivariate drought forecasting, it has been reported to be promising and is suggested for other similar climatic areas. Also, using the other previous well-used optimizers of MLP-like genetic algorithm, particle swarm, and firefly are suggested for future researchers to be compared with the imperialistic competitive algorithm and to choose the best optimizer of MLP in multivariate drought forecasting. One of the most important points in this area is that the current results are theoretically acceptable, and to have an actual or applicable investigation, it would be better to consider some more points. For example, in the simultaneous investigation and consequent prediction of agricultural and hydrological droughts, it is better suggested to investigate the relation of the index with the natural events, such as dam reservoir water content’s variations (to validate the index in hydrological drought monitoring) or vegetation cover variations (to validate the index in agricultural drought monitoring). This subject can be another suggestion for future studies. Consequently, according to the acceptable results of the current investigated climate (semi-arid and cold climate type), the work has research value for other different climates around the world.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.
Acknowledgments

The work was supported by the Bu-Ali Sina University Deputy of Research and Technology (Grant no. 98-941).

References

[1] J. Bazarfashan, S. Hejabi, and J. Rahimi, "Drought monitoring using the multivariate standardized precipitation index (MSPI)," Water Resources Management, vol. 28, no. 4, 2014.
[2] D. A. Wilhite, "Planning for drought: a methodology," in Drought assessment, Management, and Planning: Theory and Case Studies, pp. 87–108, Springer, Berlin, Germany, 1993.
[3] A. K. Mishra and V. P. Singh, "A review of drought concepts," Journal of Hydrology, vol. 391, no. 1-2, pp. 202–216, 2010.
[4] A. H. Ahmed, T. A. Awchi, M. Al-mola, and S. Shahid, "Evaluation of remotely sensed precipitation sources for drought assessment in Semi-Arid Iraq," Atmospheric Research, vol. 242, Article ID 105007, 2020.
[5] S. Nabaei, A. Sharafati, Z. M. Yaseen, and S. Shahid, "Copula based assessment of meteorological drought characteristics: regional investigation of Iran," Agricultural and Forest Meteorology, vol. 276-277, Article ID 107611, 2019.
[6] J. Bazarfashan, M. Nadi, and K. Ghorbani, "Comparison of empirical copula-based joint deficit index (IDI) and multivariate standardized precipitation index (MSPI) for drought monitoring in Iran," Water Resources Management, vol. 29, no. 6, pp. 2027–2044, 2015.
[7] M. Abbaspour and A. Sabetrafar, "Review of cycles and indices of drought and their effect on water resources, ecological, biological, agricultural, social and economical issues in Iran," International Journal of Environmental Studies, vol. 62, no. 6, pp. 709–724, 2005.
[8] W. Pozzi, J. Sheffield, R. Stefanski et al., "Toward global drought early warning capability: expanding international cooperation for the development of a framework for monitoring and forecasting," Bulletin of the American Meteorological Society, vol. 94, no. 6, pp. 776–785, 2013.
[9] R. S. Pulwarty and M. V. K. Sivakumar, "Information systems in a changing climate: early warnings and drought risk management," Weather and Climate Extremes, vol. 3, pp. 14–21, 2014.
[10] E. E. Moreira, C. L. Pires, and L. S. Pereira, "SPI drought class predictions driven by the North Atlantic Oscillation index using log-linear modeling," Water (Switzerland), vol. 8, no. 2, 2016.
[11] Z. M. Yaseen and S. Shahid, "Drought index prediction using data intelligent analytic models: a review," Intelligent Data Analytics for Decision-Support Systems in Hazard Mitigation, Springer, Berlin, Germany, 2020.
[12] P. Aghelpour, H. Bahrami-Pichagchi, B. Mohammadi, O. Kisi, and D. Zhang, "Using the MODIS sensor for snow cover modeling and the assessment of drought effects on snow cover in a mountainous area," Remote Sensing, vol. 12, no. 20, pp. 3437–3453, 2020.
[13] R. K. Deo, M. A. Ghorbani, S. Samadianfard, T. Maraseni, M. Bilgili, and M. Biazar, "Multi-layer perceptron hybrid model integrated with the firefly optimizer algorithm for wind speed prediction of target site using a limited set of neighboring reference station data," Renewable Energy, vol. 116, pp. 309–323, 2018.
[14] S. M. Biazar, V. Rahmani, Q. Bao Pham, D. Nguyen Khoi, and N. Thi Thuy Linh, "Implementing novel hybrid models to improve indirect measurement of the daily soil temperature: elman neural network coupled with gravitational search algorithm and ant colony optimization," Measurement, vol. 165, Article ID 108127, 2020.
[15] A. Ashrafzadeh, O. Kisi, P. Aghelpour, S. M. Biazar, and D. M. Asoule, "Comparative study of time series models, support vector machines, and GMDH in forecasting long-term evapotranspiration rates in northern Iran," Journal of Irrigation and Drainage Engineering, vol. 146, no. 6, 2020.
[16] P. Aghelpour and V. Varshavian, "Evaluation of stochastic and artificial intelligence models in modeling and predicting of river daily flow time series," Stochastic Environmental Research and Risk Assessment, vol. 34, no. 1, pp. 33–50, 2020.
[17] P. Aghelpour, B. Mohammadi, and S. M. Biazar, "Long-term monthly average temperature forecasting in some climate types of Iran, using the models SARIMA, SVR, and SVR-FA," Theoretical and Applied Climatology, vol. 138, no. 3-4, pp. 1471–1480, 2019.
[18] R. C. Deo, M. A. Ghorbani, S. Samadianfard, T. Maraseni, M. Bilgili, and M. Biazar, "Multi-layer perceptron hybrid model integrated with the firefly optimizer algorithm for wind speed prediction of target site using a limited set of neighboring reference station data," Renewable Energy, vol. 116, pp. 309–323, 2018.
[19] S. M. Biazar, V. Rahmani, M. Isazadeh, O. Kisi, and Y. Dinapshoh, "New input selection procedure for machine learning methods in estimating daily global solar radiation," Arabian Journal of Geosciences, vol. 13, p. 431, 2020.
[20] A. Ashrafzadeh, M. A. Ghorbani, S. M. Biazar, and Z. M. Yaseen, "Evaporation process modeling over northern Iran: application of an integrative data-intelligence model with the krill herd optimization algorithm," Hydrological Sciences Journal, vol. 64, no. 15, pp. 1843–1856, 2019.
water with more efficient input variables,” Pure and Applied Geophysics, vol. 177, pp. 5599–5619, 2020.

[29] A. Ashrafzadeh, A. Malik, V. Jothiprakash, M. A. Ghorbani, and S. M. Biazar, “Estimation of daily pan evaporation using neural networks and meta-heuristic approaches,” ISH Journal of Hydraulic Engineering, vol. 26, no. 4, pp. 421–429, 2020.

[30] M. R. Khalid, M. Isazadeh, S. M. Biazar, and Q. B. Pham, “Simulating Caspian Sea surface water level by artificial neural network and support vector machine models,” Acta Geophysica, vol. 68, pp. 553–563, 2020.

[31] A. Malik and A. Kumar, “Meteorological drought prediction using heuristic approaches based on effective drought index: a case study in Uttarakhand,” Arabian Journal of Geosciences, vol. 13, 2020.

[32] S. Mohamadi, S. Sh, and S. Fatemeh, Zoning Map for Drought Prediction Using Integrated Machine Learning Models with a Nomadic People Optimization Algorithm, Springer, Dordrecht, Netherlands, 2020.

[33] N. Khan, D. A. Sachindra, S. Shahid, K. Ahmed, M. Sanusi, and N. Nawaz, “Advances in water resources prediction of droughts over Pakistan using machine learning algorithms,” Advances in Water Resources, vol. 139, 2020.

[34] A. Belayneh and J. Adamowski, “Standard precipitation index drought forecasting using neural networks, wavelet neural networks, and support vector regression,” Applied Computational Intelligence and Soft Computing, vol. 2012, Article ID 794061, 2012.

[35] D. Hong and K. A. Hong, “Drought forecasting using MLP neural networks,” in Proceedings of the 2015 8th International Conference on U-And E-Service, Science and Technology (UNESST), pp. 62–65, IEEE, Jeju, South Korea, November 2015.

[36] P. Maca and P. Pech, “Forecasting SPEI and SPI drought indices using the integrated artificial neural networks,” Computational Intelligence and Neuroscience, vol. 2016, Article ID 3868519, 17 pages, 2016.

[37] M. Borji, A. Malekian, A. Salajegheh, and M. Ghadimi, “Multi-time-scale analysis of hydrological drought forecasting using support vector regression (SVR) and artificial neural networks (ANN),” Arabian Journal of Geosciences, vol. 9, no. 19, p. 725, 2016.

[38] P. Aghelpour, B. Mohammadi, S. M. Biazar, O. Kisi, and Z. Sourmirinezhad, “A theoretical approach for forecasting different types of drought simultaneously, using entropy theory and machine-learning methods,” ISPRS Int. J. Geo-Information, vol. 9, no. 12, p. 701, 2020.

[39] M. Svboda, M. Hayes, and D. Wood, Standardized Precipitation Index User Guide, World Meteorological Organization, Geneva, Switzerland, 2012.

[40] S.-C. Kao and R. S. Govindaraju, “A copula-based joint deficit index for droughts,” Journal of Hydrology, vol. 380, no. 1-2, pp. 121–134, 2010.

[41] M. M. Bateni, A. Behmanesh, C. De Michele, J. Bazrafshan, and H. Rezaie, “Composite agrometeorological drought index accounting for seasonality and autocorrelation,” Journal of Hydrologic Engineering, vol. 23, no. 6, 2018.

[42] E. Atashpaz-Gargari and C. Lucas, “Imperialist competitive algorithm: an algorithm for optimization inspired by imperialistic competition,” in Proceedings of the 2007 IEEE Congress on Evolutionary Computation, pp. 4661–4667, Singapore, September 2017.

[43] J. Rahimi, M. Ebrahimpour, and A. Khalili, “Spatial changes of Extended De Martonne climatic zones affected by climate change in Iran,” Theoretical and Applied Climatology, vol. 112, no. 3–4, 2013.

[44] J. A. Keyantash and J. A. Dracup, “An aggregate drought index: assessing drought severity based on fluctuations in the hydrologic cycle and surface water storage,” Water Resources Research, vol. 40, no. 9, 2004.

[45] T. B. McKee, N. J. Doesken, and J. Kleist, “The relationship of drought frequency and duration to time scales,” Applied Climatology, vol. 17, no. 22, pp. 179–183, 1993.

[46] M. Svboda and B. Fuchs, Handbook of Drought Indicators and Indices, World Meteorological Organization, Geneva, Switzerland, 2016.

[47] S. Haykin, Neural Networks: A Comprehensive Foundation, Prentice-Hall, Inc., Upper Saddle River, NY, USA, 2007.

[48] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, Learning Internal Representations by Error Propagation, University of California, San Diego, CA, USA, 1985.

[49] S.-M. Hosseini-Moghari, S. Araghinejad, and A. Azarnivand, “Drought forecasting using data-driven methods and an evolutionary algorithm,” Modeling Earth Systems and Environment, vol. 3, no. 4, pp. 1675–1689, 2017.

[50] Z. Beheshti, M. Firouzi, S. M. Shamsuddin, M. Zibarzani, and Z. Yusop, “A new rainfall forecasting model using the CAPSO algorithm and an artificial neural network,” Neural Computing and Applications, vol. 27, no. 8, pp. 2551–2565, 2016.

[51] M. Ahmadi, S. Salimi, S. A. Hosseini, H. Poorantiyosh, and A. Bayat, “Iran’s precipitation analysis using synoptic modeling of major teleconnection forces (MTF),” Dynamics of Atmospheres and Oceans, vol. 85, pp. 41–56, 2019.