

## Rainfall Prediction using Artificial Neural Network in Semi-Arid mountainous region, Saudi Arabia

Roohul Abad Khan<sup>1</sup>, Rachida El Morabet<sup>2</sup>, Javed Mallick<sup>3</sup>, Mohammed Azam<sup>4</sup>,  
Viola Vambol<sup>5\*</sup>, Sergij Vambol<sup>6</sup>, Volodymyr Sydorenko<sup>7</sup>

<sup>1</sup>Department of Civil Engineering, King Khalid University, Abha, Saudi Arabia

<sup>2</sup>Department of Geography, University Hassan II of Casablanca, Mohammedia, Morocco

<sup>3</sup>Department of Civil Engineering, King Khalid University, Abha, Saudi Arabia

<sup>4</sup>Department of Mechanical Engineering, King Khalid University, Abha, Saudi Arabia

<sup>5</sup>Department of Applied Ecology and Environmental Sciences, National University “Yuri Kondratyuk Poltava Polytechnic”, Poltava, Ukraine

<sup>6</sup>Department of Life Safety, State Biotechnological University, Kharkiv, Ukraine

<sup>7</sup>Institute of Public Administration in the Sphere of Civil Protection, Kiev, Ukraine

\*corresponding author e-mail: violavambol@gmail.com

Received: 27 August 2021 / Accepted: 15 September 2021

**Abstract:** Rainfall prediction using Artificial Intelligence technique is gaining attention nowadays. Semi-arid region receives rainfall below potential evapotranspiration but more than arid region. However, in mountainous semi-arid region high rainfall intensity makes it highly variable. This renders rainfall prediction difficult by applying normal techniques and calls for data pre-processing. This study presents rainfall prediction in semi-arid mountainous region of Abha, KSA. The study adopted Moving Average (Method) for data pre-processing based on 2 years, 3 years, 4 years, 5 years and 10 years. The Artificial Neural Network (ANN) was trained for a period of 1978-2016 rainfall data. The neural network was validated against the existing data of period 1997-2006. The trained neural network was used to predict for period of 2017-2025. The performance of the model was evaluated against AAE, MAE, RMSE, MASE and PP. The mean absolute error was observed least in 2 years moving average model. However, the most accurate prediction models were obtained from 2 years moving average and 5 year moving average. The study concludes that ANN coupled with MA have potential of predicting rainfall in Semi-Arid mountainous region.

**Keywords:** moving average method, data pre-processing, mean absolute error, root mean square error, mean absolute scaled error, performance parameter.

### 1. Introduction

Global warming and climate changes impacts are evident on rainfall distribution globally (Mislán et al., 2015). Rainfall prediction is essential for water resource management, policy and decision making and sustainability of water resources. Additionally, given fact that all industries rely on water renders its impact on economy of the country also (Xiang et al., 2018; Odnorih et al., 2020; Husain Khan et al., 2020; Ziarati et al., 2021; Sakalova et al.,

2019). Few or none surface water bodies due to high precipitation and less rainfall renders groundwater as primary source of water in arid and semi-arid region (Sihag et al., 2021). This further necessitates rainfall prediction in order to reduce groundwater dependency (Mallick et al., 2018). However, rainfall forecast is a challenge owing to meteorological complexity and non-linearity of generated mathematical models (Kashiwao et al., 2017).

Artificial Neural Network (ANN) can process complex data by employing pre-defined learning process. This renders it with adaptive nature and is highly efficient in addressing complex non-linear processes. (Chattopadhyay & Chattopadhyay, 2018). Various techniques have already been employed for different region for the prediction of rainfall. Rainfall prediction using ANN coupled with Back Propagation technique has been carried out for Monthly rainfall prediction for Indonesia (Mislán et al., 2015). Cluster Analysis was employed for rainfall prediction in Victoria Australia (Bagirov et al., 2017). Random Forests (RF) algorithm coupled with rainfall data from Meteosat Second Generation (MSG) and Spinning Enhanced Visible and Infrared Imager (SEVIRI) was used for rainfall prediction and compared with ANN and SVM results (Ouallouche et al., 2018). Support Vector Regression for short period rainfall prediction and ANN approach for long period prediction was compared for Kunming, Lincang and Mengzi, Yunnan Province, China (Xiang et al., 2018). A local monthly rainfall prediction case study was carried out for Japan using ANN approach (Kashiwao et al., 2017). ANN approach was also employed to model rainfall and watershed flood forecasting in Taiwan (Pan et al., 2013).

This study employed ANN coupled with data pre-processing for rainfall prediction in mountainous region of Abha, Saudi Arabia.

## **2. Method and Data Used**

### **2.1. Study Area**

The Abha mountainous watershed is located in Aseer province, Saudi Arabia (Fig. 1). The city experiences semi-arid climatic conditions and the climatic condition is moderate as compared to normal arid and extreme climatic condition in most of the Saudi Arabia. The rainfall varies between 0-415 mm annually. This renders highly uneven distribution of rainfall in the watershed.

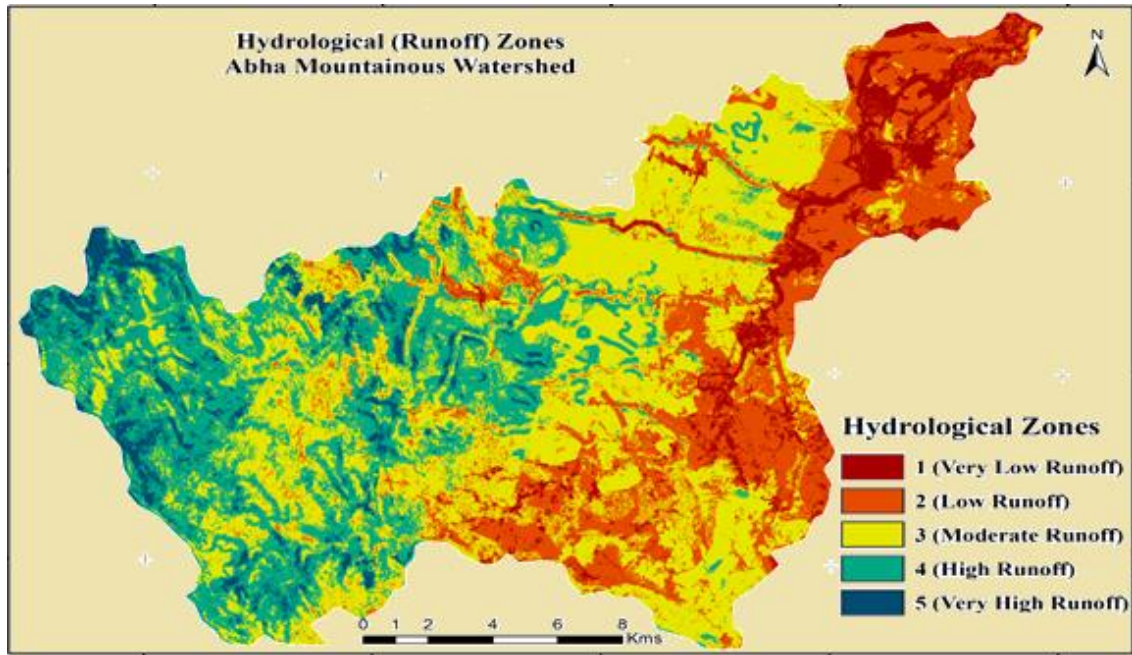


Figure 1. Hydrological Zonation of Abha watershed

## 2.2. Data Collection

The time series data was obtained from Meteorological Subdivision of Abha in Aseer Province, Saudi Arabia. The data was obtained for duration of Year 1978-2016. The total Annual rainfall was considered as it will reduce the impact of climate change impact on rainfall distribution. However, other studies have employed various artificial intelligence approach using total monthly rainfall (Pan et al., 2013; Altunkaynak & Nigussie, 2015; Mislán et al., 2015; Chattopadhyay & Chattopadhyay, 2018; Dash et al., 2018a; Xiang et al., 2018).

## 2.3. Data Pre-processing and Normalization

Data scaling or normalization is a necessary pre-processing step prior to training phase of neural network. The highly uneven rainfall distribution renders usual normalization methods ineffective for ANN. In order to smooth out data unevenness data pre-processing is required before normalization of data. Min-max normalization approach was used here for data normalization using formula presented in Eq. 1. This study employed Moving average technique for reducing unevenness in data for total annual rainfall. The Moving Average method was applied for year interval of 2 year, 3 year 4 year, 5 year and 10 year. Then the obtained values were compared for better performance.

$$a^* = a_{\min} + (a_{\max} - a_{\min}) \times \frac{a - a_{\min}}{a_{\max} - a_{\min}} \quad (1)$$

#### 2.4. Data Training and testing

The rainfall time series data of Abha from 1978-2016 was considered for analysis. the data was divided into training data 1978-2003 and testing dataset 2004-2016. This study compared the results of original dataset against 2-10 years average dataset. The input rainfall time series was denoted as  $A = (a_1, a_2, \dots, a_m)$ . The target data was represented as  $T = (t_m)$  for all pre-processed dataset including original observed values. The response of ANN approach was represented as  $R = (R_m)$

#### 2.5. Artificial Neural Network

Artificial Neural Network based on its connected neurons and input-hidden-output layers simulates human brain functioning system is a promising approach for prediction. Back propagation algorithm is generally employed for weight adjustments until the error is reduced within pre specified limits.

This study employs single layer single layer backpropagation with single input, hidden and output layer. Single layer ANN approach is usually employed in hydrological modelling.(Wu & Chau, 2013). The Back-propagation network used in this study uses A inputs and R response neurons.

The training dataset  $\{(a_1t_1), (a_2t_2) \dots, (a_mt_m)\}$  comprised input and output patterns. For a given input dataset of  $(A_i)$  the neural network gives an response  $(R_i)$  which may differ from target  $(T_i)$ . the network is trained using training algorithm in order to reduce error between the response  $(O_i)$  and target  $(T_i)$ . The ANN error was computed using Eq. 1 as the difference between Response and target value obtained (Dash et al., 2018b).

$$ANN_{\text{error}} = \frac{1}{2} \sum_{a=1}^m (A_i - t_i)^2 \quad (2)$$

The network is trained to achieve specified accuracy. Once network is trained new datasets are introduced to network. The network recognizes new patterns based on its previous training. The Back propagation technique is used in various studies for rainfall prediction using ANN. (Mislán et al., 2015; Kashiwao et al., 2017; Xiang et al., 2018) In this study, ANN architecture employed Levenberg-Marquardt for training.

### 3.5 Performance Computation

ANN model performance was computed based on various statistical parameters. The parameters are Absolute Accuracy Error (AAE, %), Mean Absolute Error (MAE, %), Root Mean Square Error (RMSE, %), Mean Absolute Scaled Error (MASE) and Performance Parameter (PP). The lower error in MAE, RMSE and MASE indicates better performance of ANN model. The closer the value of PP to 1 better the performance of forecasting model. The standard deviation (SD) was also calculated for all dataset. The Eq. 3-6 presents the statistical measures used to evaluate performance of ANN model.

$$\text{MAE (\%)} = \frac{100}{b} \sum_{a=1}^b \left[ \frac{O_i - R_i}{O_i} \right] \quad (3)$$

$$\text{RMSE (\%)} = \sqrt{\frac{100}{b} \sum_{a=1}^b \left[ \frac{O_i - R_i}{O_i} \right]^2} \quad (4)$$

$$\text{MASE} = \frac{1}{b} \sum_{a=1}^b \left( \frac{(O_i - R_i)}{\frac{1}{b-1} \sum_{i=2}^b (O_i - R_{i-1})} \right) \quad (5)$$

$$\text{PP} = 1 - \left( \frac{\text{RMSE}}{\text{SD}_O} \right) \quad (6)$$

Original (O) and predicted (R) rainfall values for time t (year) for predicting rainfall for total  $b$  number of years.

## 3. Result and Discussion

This study has investigated artificial neural network (ANN) performance of rainfall prediction in Semi-arid climatic condition for Mountainous area of Abha, Saudi Arabia.

### 3.1. Statics

Statistical parameters comprising of minimum, maximum, mean and Standard Deviation for Original values, 2 years average, 3 year average, 4 year average, 5 year average and 10 year average for 42 years and presented in Table 1. These statistical parameters were calculated for year 1978-2016 for Abha region. The year 2008 has experienced lowest amount of rainfall 84.80 mm in past 42 years owing to deficiency in rainfall as well as impact of climate change. However, when moving average method was applied the minimum rainfall based on

moving average method shifted in back years for 2 year (2008, 90.55 mm), 3 years (2007, 101.47 mm), became stagnant for average of 4 years (2000, 120.50 mm) and 5 years (2000, 126.44mm), while for 10 years it shifted to year 2014 (81.85). However, moving average method is not a specific year value but an average value of intermediate year. Nevertheless, observed statistical values will vary owing to data pre-processing.

Table 1. Statistical Summary

<b>Moving Average</b>	<b>Min (mm)</b>	<b>Max (mm)</b>	<b>Mean (mm)</b>	<b>SD (mm)</b>
2 years	90.55	529.90	226.2297	105.74284
3 years	101.47	440.53	226.8270	90.66437
4 years	120.50	378.80	227.3574	81.04385
5 years	126.44	355.44	226.2604	74.14263
10 years	81.85	306.65	208.3212	66.12177
Observed	84.80	639.50	227.3189	133.07741

### 3.2. Effect of Moving Average Intervals on accuracy of ANN models

The variation of number of years for moving average method affected the accuracy of results. This study found ANN model based on 2 year moving average provided better results as compared to other average of 3, 4 5, and 10 years. The Absolute Accuracy Error is minimal for 5 year average dataset, but for mean absolute error and root mean square method 10 year average dataset performed only next to 2 year average dataset.

Table 2. Impact of Moving average interval on accuracy of models (2003-2016)

<b>Years</b>	<b>AAE (%)</b>	<b>MAE (%)</b>	<b>RMSE (%)</b>
2 Year	4.01	1.36	1.33
3 Year	9.08	4.80	3.92
4 Year	4.50	13.26	5.27
5 Year	2.68	8.73	5.97
10 Year	8.3	3.23	2.97
Observed	11.4	4.9	3.46

The results of this study infer that increasing the number of average years for ANN modelling the accuracy decreases (Table 2). The average of 2 year dataset provided more accurate

results as compared to others which can also be observed in fig. 2 where 2 year moving average rainfall values closely follows original observed values.

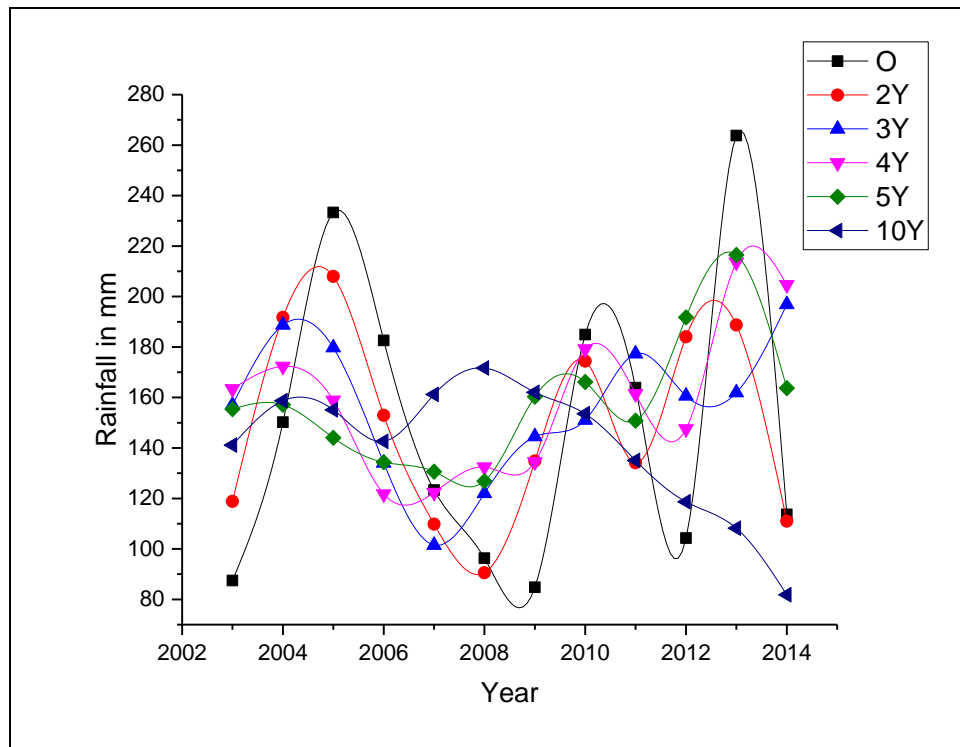


Figure 2. Rainfall observed along with Rainfall values obtained from moving Average method for interval of 2,3,4,5 and 10 years

### 3.3. Prediction accuracy of ANN models

Artificial Neural Network prediction accuracy for all the dataset is presented in Table 3. Mean original, mean predicted, Standard deviation (SD) Original, predicted standard deviation (SD), mean absolute error (%), root mean square error (%), mean absolute scaled error (%) and performance parameter were computed using ANN technique.

Table 3. Prediction performance for Moving Average Intervals (2003-2016)

Performance	Moving Average					Observed
	2 years	3 years	4 Years	5 years	10 years	
Mean Original in mm	149.916	156.272	159.337	158.132	140.809	149.5
Mean Predicted in mm	131.163	358.486	225.105	154.409	298.156	109.986
SD Original in mm	37.069	26.741	28.640	24.478	25.114	55.911
SD Predicted in mm	1.264	40.32	4.74	0.029	25.46	0.078
MAE (%)	1.36	4.80	13.26	8.73	3.23	4.90
RMSE (%)	1.339	3.923	5.274	5.975	2.975	3.462

MASE (%)	0.198	0.185	0.076	0.004	0.221	0.001
PP	0.943	0.851	0.829	0.871	0.897	0.853

It can be observed that 2-year average dataset was closest in mean original to original observed values of rainfall. Also, with respect to mean absolute Error and root mean square error its performance is better than other average year's dataset. Predicted outcomes are shown in in Fig. 3 and Fig. 4. From Fig 3 it can be observed that minimal MAE (%) was for 2 years dataset in green for all predicted year. From Fig. 4 it can be inferred that 2 years moving average was closest to prediction values against the ariginal values.

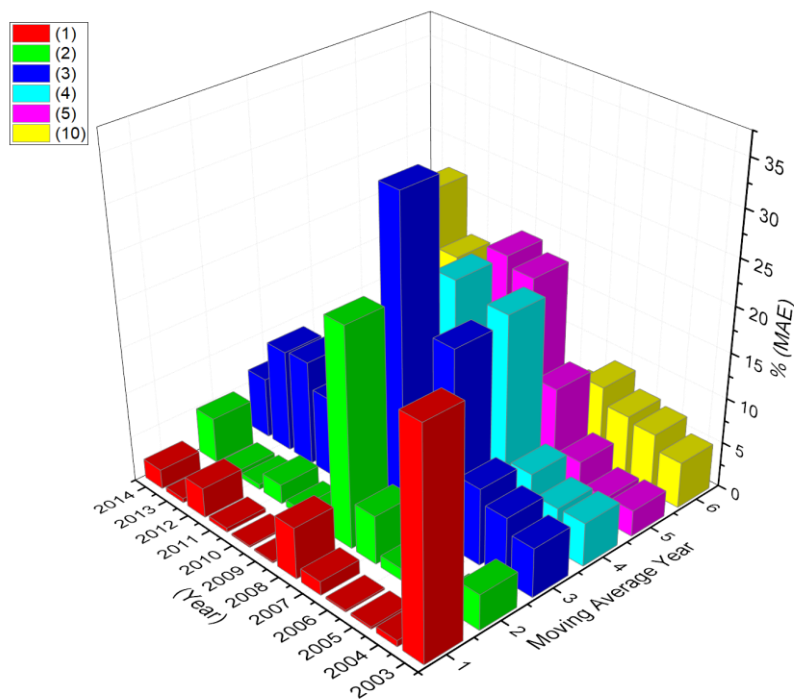


Figure 3. Mean Absolute Error (%) For Moving Average method values for interval 2 year, 3 year, 4 year, 5 year and 10 year

Rainfall over Abha mountainous region was predicted for year 2017-2025. The prediction error of model cannot be calculated as the data is not available for the interval. It will be real time test of the accuracy and performance of the model when data will be available.

The computational time for all the dataset was computed and 2 years average required least time for computation. The approximate time of computation required by all the dataset are: 25.312 seconds (2 years average dataset), 45.36 seconds (3 years average dataset), 41.324 seconds (4 year average dataset), 34.761 seconds (5 years average dataset) and 28.719 seconds (10 year average dataset). The ANN techniques requires more time as compared to



other artificial intelligence technique owing to its iterative weight updating technique followed by back propagation mechanism (Dash et al., 2018b).

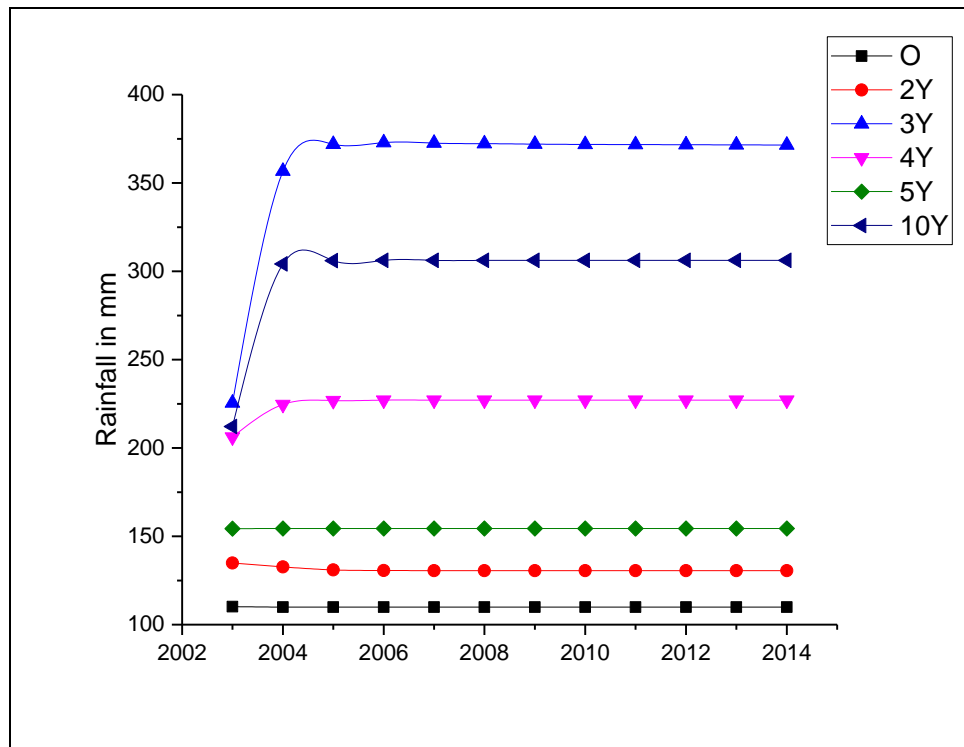


Figure 4. Rainfall prediction using Artificial Neural network for observed values against values obtained from Moving Average Method for interval of 2 year, 3 year, 4 year, 5 year and 10 years

#### 4. Conclusion

Performance of artificial neural network (ANN) was evaluated based on moving average year's dataset of 2 years, 3 years, 4 years, 5 years and 10 years for rainfall prediction in mountainous region of Abha. The performance was gauged based on various statistical parameters. The 2 year annual average dataset provided with better performance as compared to others in prediction. However, when average dataset year was increased to 3 and 4 years the error increased and the performance of model decreased. When it was increased to 5 and 10 years average dataset, model performance increased and prediction error decreased. This was also observed in statistical results presented in Table 1. This study concludes that data pre-processing is necessary for ANN modelling for rainfall prediction in mountainous region of Abha experiencing semi-arid climatic condition. Moving Average method can be successfully employed for Data pre-processing. However other artificial Intelligence methods should also be used for better comparison and obtain an optimized model for rainfall

prediction based on other parameters. The increased number of average years decreasing modelling error should be investigated further with larger datasets.

### Acknowledgement

The authors' team is grateful to the universities administration, which they represent for the opportunity to organize scientific research.

### References

- Altunkaynak A. & Nigussie T.A., 2015, Prediction of daily rainfall by a hybrid wavelet-season-neuro technique', *Journal of Hydrology*. Elsevier B.V. 529(P1): 287–301. <https://doi.org/10.1016/j.jhydrol.2015.07.046>.
- Bagirov A.M., Mahmood A. & Barton A., 2017, Prediction of monthly rainfall in Victoria, Australia: Clusterwise linear regression approach. *Atmospheric Research* 188: 20–29. <https://doi.org/10.1016/j.atmosres.2017.01.003>.
- Chattopadhyay S. & Chattopadhyay G., 2018, Conjugate gradient descent learned ANN for Indian summer monsoon rainfall and efficiency assessment through Shannon-Fano coding. *Journal of Atmospheric and Solar-Terrestrial Physics*. Elsevier Ltd 179(August): 202–205. <https://doi.org/10.1016/j.jastp.2018.07.015>.
- Dash Y., Mishra S.K., Sahany S. & Panigrahi B.K., 2018a, Indian summer monsoon rainfall prediction: a comparison of iterative and non-iterative approaches. *Applied Soft Computing* 70: 1122–1134. <https://doi.org/10.1016/j.asoc.2017.08.055>.
- Dash Y., Mishra S.K. & Panigrahi B.K., 2018b, Rainfall prediction for the Kerala state of India using artificial intelligence approaches. *Computers & Electrical Engineering* 70: 66–73. <https://doi.org/10.1016/j.compeleceng.2018.06.004>.
- Husain Khan A., Abdul Aziz H., Khan N.A., Ahmed S. Mehtab M.S., Vambol S., Vambol V., Changani F. & Islam S., 2020, Pharmaceuticals of emerging concern in hospital wastewater: removal of Ibuprofen and Ofloxacin drugs using MBBR method. *International Journal of Environmental Analytical Chemistry*, <https://doi.org/10.1080/03067319.2020.1855333>.
- Kashiwao T., Nakayama K., Ando S., Ikeda K., Lee M. & Bahadori , 2017, A neural network-based local rainfall prediction system using meteorological data on the Internet: A case study using data from the Japan Meteorological Agency. *Applied Soft Computing* 56: 317–330. <https://doi.org/10.1016/j.asoc.2017.03.015>.
- Mallick J., Singh C.K., AlMesfer M.K., Kumar A., Khan R.A., Islam S. & Rahman A., 2018, Hydro-geochemical assessment of groundwater quality in Aseer Region, Saudi Arabia. *Water* 10(12): article 1847. <https://doi.org/10.3390/w10121847>
- Mislan M., Haviluddin H., Hardwinarto S., Sumaryono S. & Aipassa M., 2015, August, Rainfall monthly prediction based on artificial neural network: a case study in Tenggara Station, East Kalimantan-Indonesia. *The International Conference on Computer Science and Computational Intelligence (ICCSCI 2015)-Procedia Computer Science* 59: 142–151. <https://doi.org/10.1016/j.procs.2015.07.528>
- Odnorih Z., Manko R., Malovanyy M. & Soloviy K., 2020, Results of surface water quality monitoring of the western Bug river Basin in Lviv Region. *Journal of Ecological Engineering* 21(3): 18–26. <https://doi.org/10.12911/22998993/118303>.

- Ouallouche F., Lazri M. & Ameer S., 2018, Improvement of rainfall estimation from MSG data using Random Forests classification and regression. *Atmospheric Research* 211: 62–72. <https://doi.org/10.1016/j.atmosres.2018.05.001>
- Pan T.Y., Yang Y.T., Kuo H.C., Tan Y.C., Lai J.S., Chang T.J., Lee C.S. & Hsu K.H., 2013, Improvement of watershed flood forecasting by typhoon rainfall climate model with an ANN-based southwest monsoon rainfall enhancement. *Journal of Hydrology* 506: 90–100. <https://doi.org/10.1016/j.jhydrol.2013.08.018>
- Sakalova H., Malovanyy M., Vasylynych T. & Kryklyvyi R., 2019, The Research of Ammonium Concentrations in City Stocks and Further Sedimentation of Ion-Exchange Concentrate. *Journal of Ecological Engineering* 20(1): 158–164. <https://doi.org/10.12911/22998993/93944>
- Sihag P., Sadikhani M.R., Vambol V., Vambol S., Prabhakar A.K. & Sharma N., 2021, Comparative study for deriving stage-discharge–sediment concentration relationships using soft computing techniques. *Journal of Achievements in Materials and Manufacturing Engineering* 104(2): 57–76. <http://dx.doi.org/10.5604/01.3001.0014.8489>
- Wu C.L. & Chau K.W., 2013, Prediction of rainfall time series using modular soft computing methods. *Engineering Applications of Artificial Intelligence* 26(3): 997–1007. <http://dx.doi.org/10.1016/j.engappai.2012.05.023>
- Xiang Y., Gou L., He L., Xia S. & Wang W., 2018, A SVR–ANN combined model based on ensemble EMD for rainfall prediction. *Applied Soft Computing* 73: 874–883. <http://dx.doi.org/10.1016/j.asoc.2018.09.018>
- Ziarati P., Kozub P., Vambol S., Vambol V., Khan N.A., Kozub S. & Tajik S., 2021, Kinetics of Cd, Co and Ni Adsorption from Wastewater using Red and Black Tea Leaf Blend as a Bio-adsorbent. *Ecological Questions* 32(2): 59–70. <http://dx.doi.org/10.12775/EQ.2021.014>