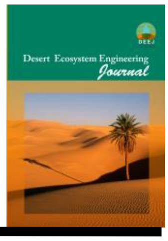
DEEJ (2023) 11 (36) 71-82

University of Kashan

Desert Ecosystem Engineering Journal

Journal homepage: http://deej.kashanu.ac.ir




# Predicting Pan Evaporation in a Hyper-Arid Climate Using Soft Computing Models: A Case Study of Sistan Plain, Sistan-Baluchistan, Iran

**Mojtaba Mohammadi[1]*, Masoomeh Forozanfard[2], Hamid Gholami[3]**

Received: 2022/6/10 Accepted: 2023/1/11

**Expanded abstract**

**Introduction:** Considering the fact that evaporation affects the planning and operations of water resources as a key process in the hydrologic cycle, predicting its magnitude and patterns, particularly in arid, semi-arid, and hyper-arid environments such as Sistan plain (in northern Sistan-Baluchistan Province, Iran) is of great importance. On the other hand, as accurate estimation of Pan evaporation is regarded as one of the main aspects of water management in such regions, it is crucially important to accurately simulate the pan evaporation based on the available regional meteorological parameters. Therefore, this study sought to investigate the capabilities of soft computing techniques for estimating monthly evaporation in Sistan plain. The results of the study could be helpful for the management of water resources in the Sistan area, allowing the policymakers to develop future projects of water resource management/development plans for the region based on Evaporation estimations.

**Materials and Methods:** various meteorological parameters, including maximum, minimum, and average temperature rates, relative humidity, wind speed, and precipitation rate were used to predict monthly Evaporation using the consistent and uninterrupted historical time series data (1994–2021) collected from three meteorological stations (Zabol, Zahak, and Chahnimeh).

The main purpose of this study was to assess the performance of a soft computing model in simulate pan evaporation. To this end, nine soft computing models, including Model Tree (MT), Random Forest (RF), Support Vector Machines (SVM), Bayesian Ridge Regression (BRR), Gaussian Process (GP), Extreme Gradient Boosting (XGB), Artificial Neural Network (ANN), and Multivariate Adaptive Regression Splines (MARS) were used to predict evaporation at the meteorological stations selected for this research.

1. Assistant Professor (Corresponding Author), Department of Desert Management & Control, Faculty of Natural Resource, Higher education complex of Saravan, Iran.
2. Department of Desert Management & Control, Faculty of Natural Resource, Higher education complex of Saravan, Iran.
3 Associate Professor, Department of Watershed Management Engineering, Faculty of Natural Resources, Hormozgan University, Bandar Abbas, Iran.
DOI: 10.22052/deej.2021.11.36.43

University of Kashan

Desert Ecosystem Engineering Journal

Journal homepage: http://deej.kashanu.ac.ir

On the other hand, the model's performance was assessed using statistical measures, including coefficient of determination (R2), root mean square error (RMSE), mean absolute error (MAE), and Taylor diagram. Moreover, to construct predictive models, the dataset was divided into training (70%) and validation (30%) data. Then, eight combinations of input parameters were selected for Zabol and Zahak sites based on the Pearson correlation coefficient between the individual input parameters and evaporation. Finally, the best input combination and the optimal values for different models were determined using the R programming language.

**Results:** The different input combinations were determined for the two sites independently based on the inclusion of the weather parameters with the highest coefficient Pearson with evaporation. Then, each model was run using various fixed sets of parameters. For the Zabol station, the minimum temperature rate indicated the greatest correlation coefficient with (0.96), followed by average temperature (0.95) and smallest by rainfall (-0.42). It was also found that the outputs of Zabol and Zahak stations were very close and similar to each other, which could partly be attributed to the proximity of the two stations and their same topography and climatic conditions in the Sistan plain.

On the other hand, the results of assessing the models' performance indicated that in the validation stage, MT (whose R2 = 0.97 and RMSE=57.3) delivered the best performance in Zabol station under Scenario 1, RF (with its R2 and RMSE being 0.98 61.7, respectively) performed the best under Scenario 2, MT with its R2 and RMSE being 0.97 and 61.27, respectively, showed the best performance under Scenario 3, the ANN and MT (whose R2 and RMSE reported as being 0.96 and 59.9, respectively) put in the best performance under Scenario 4 , and RF with R2 = 0.97 and RMSE=59, 59.24, 58.23, and 58.3 delivered the best performance under scenarios 5 to 8, respectively.

Moreover, the results suggested that adding the number of input variables to the models made no difference in their accuracy level. It was also found that out of eight scenarios investigated in this study, the RF model delivered the performance under five scenarios (2, 5, 6, 7, and 8), and that the MT performed well under two scenarios (1 and 3). Therefore, RF and MT could be introduced as the best soft computing models for simulating and estimating the pan evaporation of the Sistan plain.

**Conclusion:** The results showed that with increasing the input of variables to the model, there was not much difference in the accuracy of the models. For example, R2 of scenario 1 with only the minimum monthly temperature input is equal to scenario 8 with eight inputs equal to 0.98 in the validation stage. Therefore, the findings showed that the model with only the minimum monthly temperature input has the same performance as the model with eight inputs The main contribution of this study was introducing a soft computing model to accurately estimate the pan evaporation of the Sistan plain using a meteorological parameter (minimum monthly temperature).

**Keywords:** Prediction**,** Evaporation, Soft Computing, Hyper-Arid Region, Sistan Plain.