Evaluating the performance of a hybrid machine learning method for modeling evapotranspiration

Paulina Limón [a], Jorge Gálvez [a], Gerardo Nuñez-González [a], Alma Y. Alanís [a]

[a]Universidad de Guadalajara, CUCEI, Guadalajara, Jal, Mexico.
Av. Revolución 1500, Guadalajara, Jal, México
{jorge.galvez, gerardo.nunezg, alma.alanis}@academicos.udg.mx
paulina.limon@cucsur.udg.mx

Abstract. Evapotranspiration (ET) is a hydrological variable that is very important in hydrology and agriculture. Traditionally, ET estimation process is based on empirical models and conducted over monthly and daily data. Technological developments have increased the collection of climate data related to ET. Commonly, processing data with a finer temporal resolution using classical methods is complex. However, artificial intelligence has become an alternative tool to model ET, when large datasets are available. One of the most used techniques to model ET is neural networks. Models developed using neural networks have shown good performance in ET estimation, mainly when a multilayer perceptron is employed. However, it is recognized that multilayer perceptron architectures could present difficulties related to their learning mechanisms. Thus, this paper presents the implementation of a hybrid machine learning method based on multilayer perceptron and metaheuristic algorithms. The experimental results show that hybridized methods can improve the ET estimation process. In this study, the performance comparison is based on several state-of-the-art metaheuristic methods including the Runge Kutta Optimization Algorithm, the Crow Search Algorithm, the Fuzzy Logic, and the Grey Wolf Optimizer Algorithm.

Keywords: Machine learning, Metaheuristic Algorithms, Multilayer Perceptron.

1 Introduction

Evapotranspiration refers to the quantity of water that returns to the atmosphere due to evaporation from water bodies and vegetation transpiration [1]. Estimating evapotranspiration is considered an essential task in hydrological studies and water resources management in agriculture [2]. Commonly, evapotranspiration estimation is based on the calculation of reference evapotranspiration, which is considered a climatic parameter that expresses the evaporating power of a specific location and time [3]. However, due to the complexity of their measurement, empirical methods were formulated through time mainly based on the relationship between evapotranspiration and climatological variables. Among these methods are Priestley-Taylor [4], Hargreaves [5] Penman [6], and the most used and recommended by the Food and Agriculture Organization (FAO), the FAO-56 Penman-Monteith method [3]. One feature of these methods

https://doi.org/10.61728/AE20255060

is that the calculations are performed commonly on a monthly and daily basis. This was due to the lack of information on the climatological variables used in the calculations in the past. Nonetheless, it is currently possible to use technological devices to collect meteorological data with a finer temporal resolution, but now, a task to solve consists of processing this information. In this sense, computational models derived from artificial intelligence have become a good option mainly due to their capacity to capture complex dynamics and non-linear patterns. Among these kinds of models, Neural Networks are one of the most used and tested to model evapotranspiration [7]. For example, Kumar et al. [8] use a Multilayer Perceptron to model reference evapotranspiration based on information on six meteorological variables, obtaining good results compared with the FAO 56 Penman-Monteith method and lysimeter measurements. Rahimikhoob [9] also uses a multilayer perceptron to model evapotranspiration but uses only information about maximum and minimum temperature and solar radiation. Their results show a good agreement with the reference evapotranspiration calculated based on FAO 56 Penman-Monteith method. Recently, Jain and Gupta [10] compared four different architectures to model evapotranspiration according to the climate conditions of India. Their results found that it is possible to model evapotranspiration with good results based only on temperature, solar radiation, and wind speed information. Nevertheless, the results recognized that the models based on multilayer perceptron could have problems with the global weight adjustment in their learning mechanisms. To solve this problem, new architecture has been proposed. The hybridization method has been recognized as a good option to improve the performance of the models because, in some research, precision improvement and a reduction in computing time have been observed. However, the need to study the behavior of these methods is recognized. Thus, this research aimed to implement a hybrid machine-learning method based on multilayer perceptron and metaheuristic algorithms and evaluate their performance to predict reference evapotranspiration.

2 Methodology

To analyze the performance of the optimization algorithm, a multi-layer perceptron was built based on a network architecture 5-6-1, which means the input layer is composed of five input variables, the hidden layer comprises six neurons, and an output layer which returns evapotranspiration predictions. The number of parameters or function weights according to the proposed architecture was 42. The iteration number to evaluate the performance of the optimization techniques was 1000. A schematic description of the multi-layer perceptron is shown in Figure 1. A list of the metaheuristic methods applied in the hybridization as well as their initial parameter configuration is shown in Table 1. The proposed algorithm was implemented using MATLAB software. The input data were obtained from a weather station located in Autlán, Jalisco, México. This weather station registers hourly data for temperature, solar radiation, wind velocity, relative humidity, and barometric pressure.

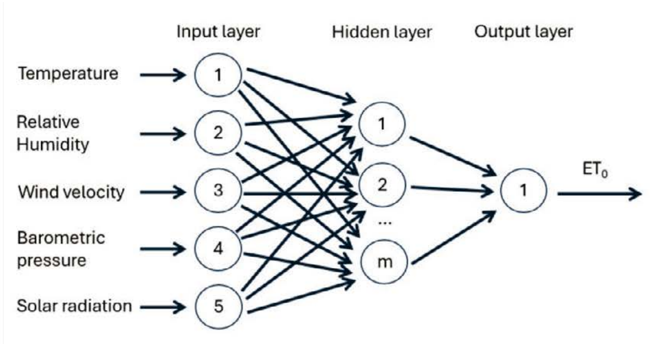

Fig. 1. Multi-layer perceptron architecture. The input layer uses the climatic variables commonly used to calculate evapotranspiration with traditional models.

Table 1. Metaheuristic algorithms used for the hybridization.

ID	Name	Parameters
PSO	Particle Swarm Optimization	C_1=2 C_2=2 *Wdamp*=0.99 *W*=1
SCA	Sine Cosine Algorithm	r = 2
GWO	Grey Wolf Optimizer	r_1 = rand(0,1) $r2$ = rand(0,1)
CSA	Crow Search Algorithm	AP = 0.1 Fl = 2
CMA-ES	Covariance Matrix Adaptation Evolution Strategy	σ = 0.3 α = 2
Fuzzy	Fuzzy Logic	β = 3.5 β_1 = 0.15 *Per* = 1.5 *Tun* = 0.75
RUN	Runge Kutta Optimization Algorithm	h = 2

3 Results and discussion

Figure 2 shows examples of the results obtained while running the seven metaheuristic algorithms used to optimize the model's parameters. The first subgraph shows the results for all the models tested. In this subgraph, apparently, all the models show similar behavior. However, when the results of each model were plotted individually, there was some interesting behavior. First, it is noticeable that almost all the models show rapid convergence, although the PSO, SCA, and RUN show better results after iteration number 400. According to the mean descriptive statistics (Table 2) calculated based on 30 runs of each algorithm, the Runge Kutta Optimization Algorithm performs better, followed by the Crow Search Algorithm, the Fuzzy Logic, and the Grey Wolf Optimizer.

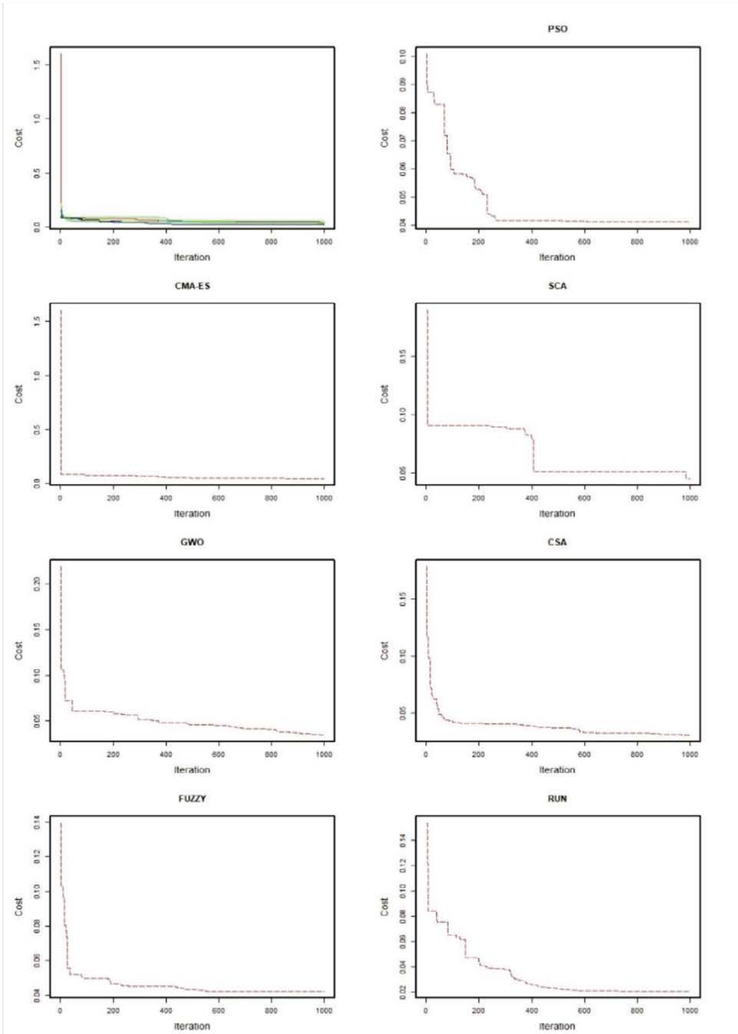

Fig. 2. Convergence of the metaheuristic algorithms.

This behavior can also be observed in Figure 3. The asymptotic behavior of the Ruge Kutta approaching the value of 0.02 is noticeable.

During the algorithms' running, the PSO algorithm's results fluctuated wildly, although their execution time can be considered low. On the other hand, the CMA-ES algorithm, which was expected to perform well, showed the highest variability between runs, which can be confirmed through its standard deviation shown in Table 2. Although not designed with a focus on climatological data, the SCA algorithm shows a behavior similar to that observed with CMA-ES. Regarding the execution time, the Fuzzy Logic was the most time-consuming because it uses approximately six minutes per run compared to the two minutes of execution of the other algorithms. Nevertheless, this time of execution is considered as low because, in all the runs, the algorithms stopped until they reached 1000 iterations.

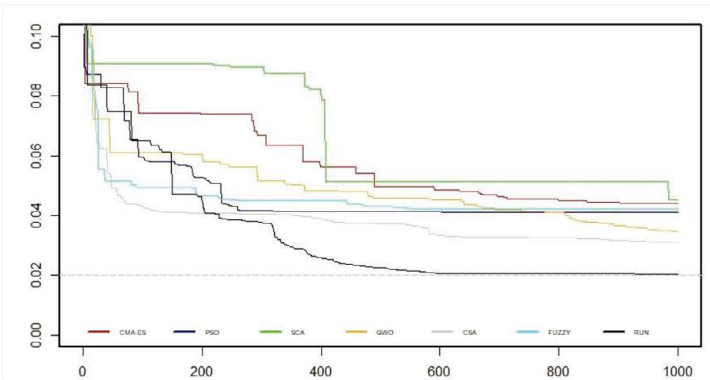

Fig. 3. Comparison of the convergence of the metaheuristic algorithms.

Table 2. Mean descriptive statistics of the running tests.

Statistic	PSO	SCA	GWO	CSA	CMA-ES	FUZZY	RUN
Minimum	0.061	0.058	0.044	0.041	0.059	0.45	0.024
Maximum	0.138	0.168	0.165	0.158	0.888	0.161	0.174
Mean	0.067	0.072	0.051	0.047	0.068	0.050	0.041
STD	0.012	0.015	0.010	0.011	0.029	0.011	0.021

4 Conclusion

In this research, machine learning techniques and metaheuristic algorithms were implemented to find models with better performance for the prediction of evapotranspiration. The algorithms selected for the study were the Particle Swarm Optimization (PSO), Crow Search Algorithm(CSA), Sine Cosine Algorithm (SCA), Grey Wolf Optimizer (GWO), Covariance Matrix Adaptation Evolution Strategy (CMA-ES), FUZZY LOGIC (FUZZY), and the Runge Kutta Optimizer (RUN) in conjunction with a multilayer perceptron. The results show that hybridization can improve the performance of the models focused on the prediction of hourly evapotranspiration, mainly when it is used in the hybridization of the Runge Kutta Optimization Algorithm, the Crow Search Algorithm, the Fuzzy Logic, and the Grey Wolf Optimizer Algorithm. Nevertheless, it is necessary to carry out more experiments to confirm and improve the performance of the models shown in this research.

References

1. Wu, L., Huang, G., Fan, J., Ma, X., Zhou, H., Zeng, W.: Hybrid extreme learning machine with meta-heuristic algorithms for monthly pan evaporation prediction. Computers and Electronics in Agriculture 168, 105115 (2020).
2. Núñez-González, G., Velázquez Pérez, D., Pelayo-Cortés F.J., Barboza-Jiménez, P.: Analysis of reference evapotranspiration behavior during the rainy season at five weather stations in the Lerma-Chapala basin. Ingeniería Agrícola y Biosistemas 11(2), 147–159 (2019).
3. Allen, R.G., Pereira, L.S., Raes, D., Smith, M.: Crop evapotranspiration: guidelines for computing crop water requirements. FAO, Rome (2006).
4. Priestley, C.H.B., Taylor, R.J.: On the assessment of surface heat flux and evaporation using large-scale parameters. Monthly weather review 100(2), 81–92 (1972).
5. Hargreaves, G.H., Samani, Z.A.: Reference crop evapotranspiration from ambient air temperature. In: Winter Meeting American Society of Agricultural Engineers, pp. 1–12, Chicago, (1985).
6. Allen, R.G., Pruitt, W.O.: FAO-24 Reference evapotranspiration factors. Journal of Irrigation and Drainage Engineering 117(5), 778–773 (1972).
7. Chia, M.Y., Huang, Y.F., Koo, C.H., Fung, K.F.: Recent Advances in Evapotranspiration Estimation Using Artificial Intelligence Approaches with a Focus on Hybridization Techniques—A Review, Agronomy 10(1), 101 (2022).
8. Kumar, M., Raghuwanshi, N. S., Singh, R., Wallender, W. W., Pruitt, W. O.: Estimating Evapotranspiration using Artificial Neural Network, Journal of Irrigation and Drainage Engineering, 128(4), 224–233 (2002).
9. Rahimikhoob, A.: Estimation of evapotranspiration based on only air temperature data using artificial neural networks for a subtropical climate in Iran, Theor Appl Climatol, 101(1), 83–91 (2010).
10. Jain S. K., Gupta, A. K.: Investigation of Multilayer Perceptron Regression-based Models to Forecast Reference Evapotranspiration (ETo), Research Reports on Computer Science, 2(3): 122–134 (2023).