

Evaluation of Several Empirical Models in Estimating the Monthly Runoff in Urban Region. Case Study: Kafr Kela Al Bab Region, Egypt

Youssef Kassem^{*1}, Hüseyin Gökçekuş², Sarah Ahmed Helmy Salem³

^{1,2,3} Near East University, Nicosia, North Cyprus

yousseuf.kassem@neu.edu.tr

huseyin.gokcekus@neu.edu.tr

20233447@std.neu.edu.tr

Abstract. For flood risk management in a particular location, precise runoff prediction is crucial, although it's still a challenging process because it's unexpected and imprecise. Accordingly, the main goal of this study is to predict urban runoff in the Kafr Kela Al Bab region of Egypt using the quadratic model (QM), Elman neural network (ENN), cascade forward neural network (CFNN), and Poisson regression (PR). To accomplish this purpose, global meteorological data are collected for analysis in this study between 1985 and 2023. The results showed that the QM model outperforms the other models, with $R^2 = 0.9906$ and 0.9902 during the validation phase. Furthermore, the RSME and MAE demonstrate that the observed values and the QM models' derived values strongly agree with each other in terms of statistical parameters.

Keywords: Global Meteorological Data, Artificial Neural Network, PR, Weather Parameter, Runoff.

1 Introduction

Climate change has led to changes in Earth's weather patterns. One major consequence of climate change is the disruption of rainfall patterns [1]. Shifting patterns in rainfall have a direct impact on urban runoff and flooding, according to Kassem et al. [2] and Sefry et al. [3]. Additionally, due to the increasing volume of runoff, urban drainage systems may flood in areas that receive frequent and heavy rains [4]. In urban environments, impermeable surfaces typically impede the infiltration of natural water. Urban runoff is the result of rainwater flowing over surfaces, accumulating contaminants and finding its way into stormwater drains [5]. Moreover, the risk of flooding in urban areas is increased by the combination of increased rainfall and urban runoff [6].

Therefore, it is crucial to forecast and simulate urban runoff in order to make well-informed decisions and to prepare for floods, manage pollutants, and promote sustainable urban growth. Recently, empirical models like artificial neural networks (ANN) have been applied as an alternative tool for better understanding how rainfall-runoff behaves in urban area. Numerous studies have utilized ANN to predict rainfall-runoff at different time scales worldwide. For instance, Kassem et al. [7] utilized different ANN models to forecast monthly runoff in the Bared River, Lebanon. Tiwari et al. [8] developed a rainfall-runoff prediction model using a radial basis function neural network. Mandal and Biswas [9] employed ANN and the Soil Conservation Service Curve Number method to predict monthly runoff in the Mayurakshi River Catchment, India.

Therefore, to predict runoff in Kafr Kela Al Bab, Egypt, the study utilized a variety of empirical models, such as quadratic model (QM), Poisson regression (PR), Elman neural network (ENN), and cascade forward neural network (CFNN). Global meteorological dataset were utilized for predicting the runoff in the chosen area to accomplish this goal. The variables are as follows: wind speed (WS), solar radiation (SR), vapor pressure (VP), actual evapotranspiration (AE), rainfall (R), runoff (Q), maximum temperature (Tmax), minimum temperature (Tmin), and soil moisture (SM). For this study, average monthly data from 1958 to 2023 were gathered.

2 Material and Method

2.1 Study area

The Kafr Kila al-Bab is situated at 30.69 North and 31.15 East. This region is distinguished by its level terrain and closeness to the Nile River. The region is distinguished by its level terrain and its proximity to the Nile River. Due to its low-lying terrain and the annual flooding that the Nile causes during the flood season, the delta is susceptible to flooding. Generally speaking, the region may see significant rainfall from October to March, or the rainy season. Additionally, the combination of flat topography and impermeable terrain can lead to rapid runoff that overwhelms drain-age systems and results in widespread flooding.

2.2 Dataset

The global meteorological reanalysis dataset has been used in the current study due to measurement limitations. TerraClimate was used to collect data for the chosen site. In addition to global monthly meteorological data with high-spatial resolution ($1/24^\circ$, ~ 4 km) produced from the WorldClim dataset, the TerraClimate dataset includes monthly station data from 1958 to the present. The monthly worldwide meteorological data for this study were gathered between 1985 and 2023 and included the maximum temperature (Tmax), minimum temperature (Tmin), rainfall (R), runoff (Q), wind speed (WS), solar radiation (SR), vapor pressure (VP), actual evapotranspiration (AE), and soil moisture (SM). The descriptive statistics for the data are listed in Table 1.

Table 1. Data's descriptive statistics

Variable	Unit	Mean	Standard deviation	Minimum	Maximum
Tmax		27.64	6.05	16.17	37.50
Tmin		12.92	5.50	2.06	22.70
R	mm	2.94	4.85	0.00	33.90
WS	m/s	2.58	0.43	1.31	4.27
SR	W/m ²	231.54	60.85	120.70	327.10
VP	KPa	1.50	0.45	0.78	2.48
SM	mm	0.19	0.20	0.10	1.40
AE	mm	2.79	4.61	0.00	32.20
Q	mm	0.15	0.25	0.00	1.70

2.3 Empirical Model

Accurate estimation of a PV system's output performance is crucial. Recently, empirical models such as artificial neural networks (ANN) and mathematical models have utilized as effective tools for predicting PV system outputs. Therefore, Elman neural network (ENN), and cascade forward neural network (CFNN), Poisson regression (PR), and quadratic model (QM) is utilized to predict the runoff. The description of these models can be found in Ref. [10–14].

a) ENN

It is an extraordinarily computationally powerful feedback neural network. The four levels are the context layer, hidden layer, output layer, and input layer. The signal transmission function is accomplished by the input layer. However, there is simply a linear weight impact in the output layer. The context layer distinguishes Elman neural networks (ENN) from backpropagation neural networks.

b) CFNN

In general, feed-forward neural networks and CFNNs are comparable. An input layer, one or more hidden layers, and an output layer make up its three layers. The input-derived weights are present in the first layer. Weights from the input and all earlier layers are also included in each subsequent layer. Every layer contains biases. The output layer comes in last. The weights and biases of each layer need to be set up. The model's performance throughout the training phase was assessed using the mean squared error.

c) PR

One of the most popular models for modeling uncommon occurrences and count data is PR. PR has been used in a number of scientific research to simulate unusual events and determine the relationship between factors. When PR is applied, the following two presumptions are typically used: the output variable follows a Poisson distribution, and (2) the mean of the response variable is equal to the variance.

d) QM

QM is a statistical approach that examines the connection between input variables and output responses to optimize processes and enhance product quality. Creating experiments, fitting mathematical models to the data from the trials, and figuring out the best input variable settings are all part of it. RSM is useful for deriving the ideal operating conditions and for examining intricate systems with numerous input variables. It facilitates effective process optimization and the comprehension of the relationships between variables.

2.4 Statistical Indices (SI)

Various statistical indicators have to be employed to assess the performance of the constructed models. Three indicators are used in this study to assess the model's efficacy. Eqs. (1)-(3) provide specific mathematical formulas for the metrics utilized in this study.

$$R^2 = 1 - \frac{\sum_{i=1}^n (a_{a,i} - a_{p,i})^2}{\sum_{i=1}^n (a_{p,i} - a_{a,ave})^2} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (a_{a,i} - a_{p,i})^2} \quad (2)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |a_{a,i} - a_{p,i}| \quad (3)$$

where n is the number of data, $a_{p,i}$ is the predicted value, $a_{a,i}$ is the actual value, $a_{a,ave}$ is the average actual value, and i is the number of input variables.

3 Results

3.1 Results of the ENN and CFNN

The proposed study aims to improve the performance of the artificial neural network (ANN) model by designing different network configurations. The optimal model for neural networks is found by an iterative method that employs hidden layer configurations ranging from 1 to 15 and experimental iterations ranging from 10,000 to 1,000,000. The endurance and quickness of the Levenberg-Marquardt training approach contributed to its adoption. To limit the probability of local minima, many training cycles use randomized variable initialization. Model selection is based on the minimal mean square error (MSE). Furthermore, the dataset was separated into two categories: testing data (2000–2019) and training data (all data except 2000–2019), for use in the development and evaluation of ANN models. According to the study, the hyperbolic tangent function was the most effective activation function for the CFNN and ENN model's input and output layers. Furthermore, it was demonstrated that the optimum network topology comprised of one hidden layer with 15 neurons for CFNN and 8 neurons for ENN.

As mentioned previously, a variety of statistical indices were employed in the performance assessment of the proposed models. The observed value and the anticipated value calculated for the testing phase using CFNN and ENN are compared in Fig. 1.

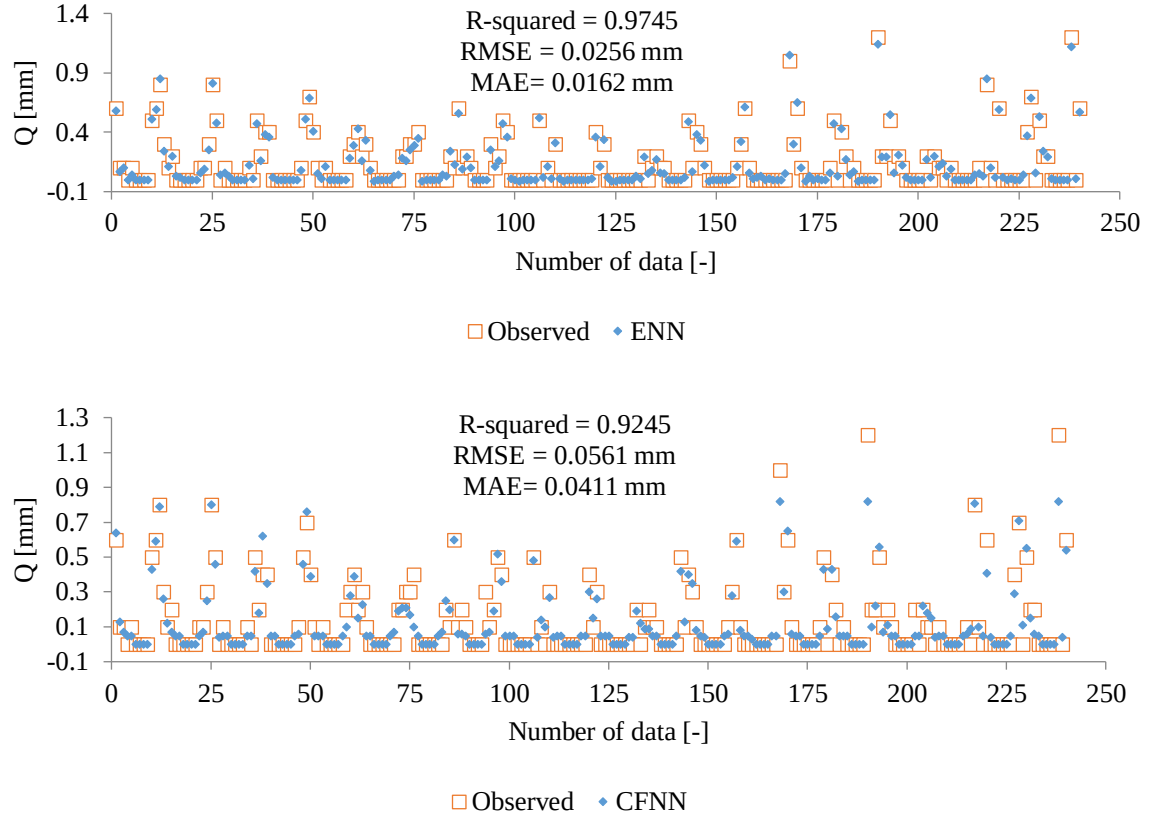


Fig. 1. Comparison between the observed and with the estimated values by MLPNN and RBF

3.2 Results of the QM and PR

QM and PR models were developed for this study to estimate monthly runoff at the selected region. The dataset's data from 2000 to 2019 was used in the testing phase, and the remaining data was utilized to generate the mathematical equation that is displayed in Eq. (4) and Eq. (5) for QM and PR model, respectively.

$$\begin{aligned}
 Q = & 0.259 - 0.0511 \cdot T_{max} + 0.043 \cdot T_{min} + 0.514 \cdot R + 0.0569 \cdot WS + 0.00035S \cdot R + \\
 & 0.043VP + 0.082 \cdot SM - 0.487 \cdot AE + 0.002 \cdot T_{max}^2 + 0.00228 \cdot T_{min}^2 + 0.0052 \cdot R^2 + 0.00267 \cdot \\
 & WS^2 - 0.0148 \cdot VP^2 + 0.00228 \cdot SM^2 - 0.00583 \cdot AE^2 - 0.00448 \cdot T_{max} \cdot T_{min} - 0.031 \cdot T_{max} \cdot \\
 & R - 0.00479 \cdot T_{max} \cdot WS + 0.0067 \cdot T_{max} \cdot VP - 0.00728 \cdot T_{max} \cdot SM + 0.0331 \cdot T_{max} \cdot AE + \\
 & 0.0538 \cdot T_{min} \cdot R - 0.0036T_{min}VP + 0.00751T_{min}SM - 0.0567T_{min}AE + 0.0781RWS - \\
 & 0.00116 \cdot R \cdot SR - 0.182 \cdot R \cdot VP + 0.084 \cdot R \cdot SM - 0.0176 \cdot WS \cdot VP + 0.0095 \cdot WS \cdot SM - \\
 & 0.0833 \cdot WS \cdot AE - 0.000323 \cdot SR \cdot VP + 0.000103 \cdot SR \cdot SM + 0.00122 \cdot SR \cdot AE - 0.089 \cdot SM \cdot \\
 & AE
 \end{aligned}
 \quad (4)$$

$$\begin{aligned}
 Q = & \exp(-0.98 - 0.765 \cdot T_{min} - 0.0116 \cdot SR - 1.39 \cdot VP + 0.3067 \cdot AE - 0.00685 \cdot AE^2 + \\
 & 0.00373 \cdot T_{min} \cdot SR + 0.581 \cdot T_{min} \cdot VP + 0.0097 \cdot SR \cdot VP - 0.00301 \cdot T_{min} \cdot SR \cdot VP)
 \end{aligned}
 \quad (1)$$

Fig. 2 shows the observed value during the testing phase with the estimated value calculated using QM. Also, Fig. 3 compares the measured value with the predicted value estimated using ARIMA for testing phase.

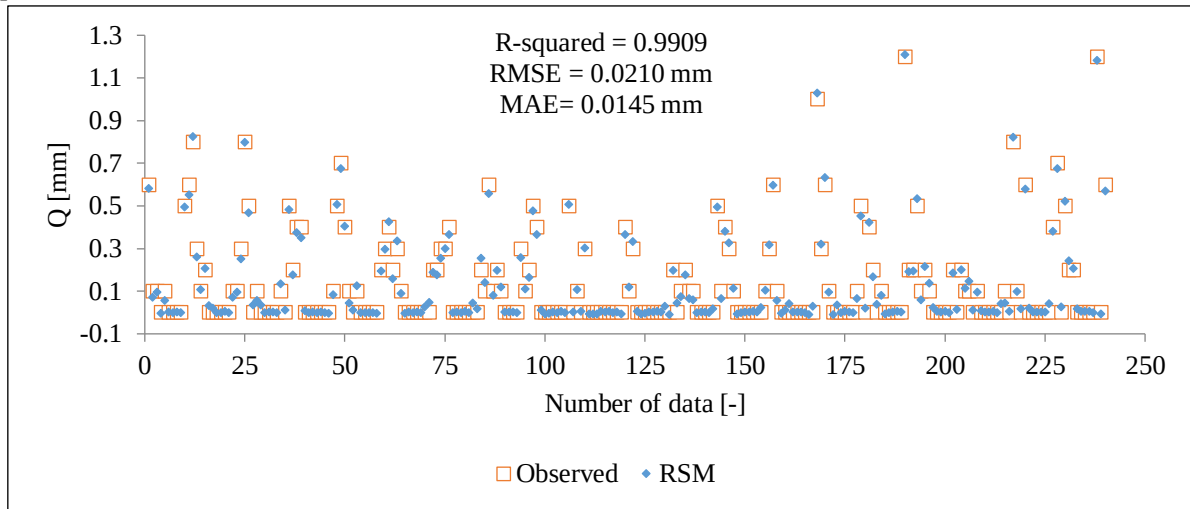


Fig. 2. Comparison between the measured with the estimated value by RSM

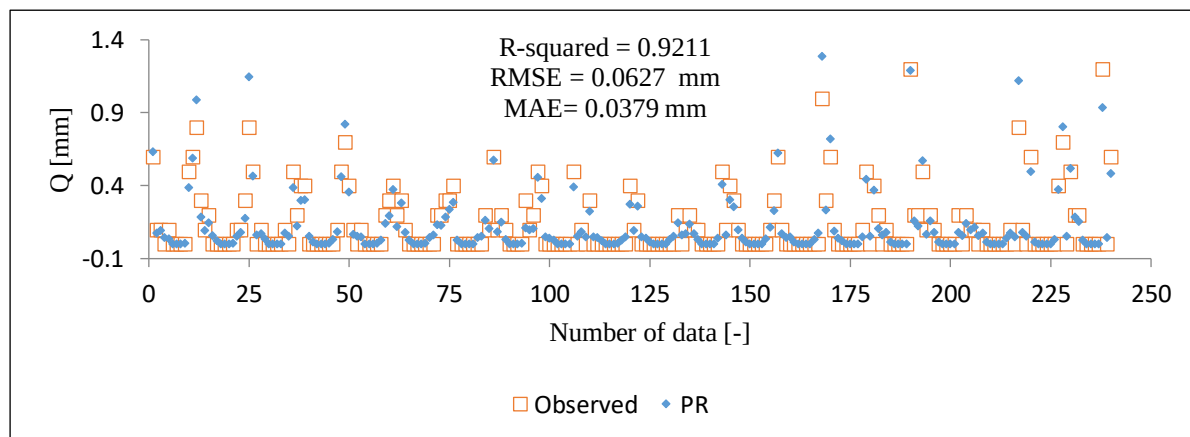


Fig. 3. Comparison between the measured with the estimated value by ARIMA

4 Discussion and Conclusions

The results of this study hold great significance for water quality and flood risk management that rely on runoff. Urban runoff impacts water quality as it collects pollutants like oil, chemicals, heavy metals, and litter from impervious surfaces. This contaminated runoff then enters stormwater drainage systems, eventually discharging into local water bodies. The presence of pollutants can degrade water quality, harming aquatic ecosystems, and posing risks to public health if the water is used for drinking or recreational purposes. Urban runoff plays a major role in flood risk management. In densely populated areas with extensive impervious surfaces, rainwater cannot infiltrate the ground effectively. This leads to increased surface runoff during heavy rainfall events, overwhelming stormwater systems, and causing flooding. Accurate prediction of runoff can aid in designing and managing stormwater infrastructure, ensuring that it can handle peak flows and reduce flood risk. Urban runoff and its prediction play a crucial role in shaping the sustainability and resilience of urban environments. As cities continue to grow and urbanization expands, understanding the significance of urban runoff and accurately predicting its behavior become essential for several reasons. According to Ref. [15], the amount, intensity, and duration of rainfall events determine the rate at which water accumulates on the land surface and the potential for

runoff. Heavy and prolonged rainfall can lead to higher runoff rates as the soil becomes saturated and unable to absorb more water [16].

Therefore, precise urban runoff prediction is essential for managing flood risk in a given area, but it remains a difficult procedure due to its unpredictability and imprecision. According to the findings, the results showed that the QM model outperforms the other models, with $R^2 = 0.9906$ during the validation phase. Furthermore, the RSME and MAE demonstrate that the observed values and the QM models' derived values strongly agree with each other in terms of statistical parameters.

References

1. Shahzad, A., Ullah, S., Dar, A. A., Sardar, M. F., Mehmood, T., Tufail, M. A. ... & Haris, M.: Nexus on climate change: Agriculture and possible solution to cope future climate change stresses. *Environ Sci Pollut Res*, 28, 14211-14232 (2021). <https://doi.org/10.1007/s11356-021-12649-8>
2. Kassem, Y., Gökçekuş, H., & Alijl, N.: Flash flood risk assessment for Girne region, Northern Cyprus. *Engineering, Technology & Applied Science Research*, 12(3), 8728-8737 (2022). <https://doi.org/10.48084/etasr.4939>
3. Sefry, S. A., Youssef, A. M., AbuAlfadael, E. Y., & AbuAbdullah, M. M.: Flood peak discharge and volume runoff envelope curves for 100-year return period for the arid region: Saudi Arabia. *Natural Hazards*, 115(2), 1439-1460 (2023). <https://doi.org/10.1007/s11069-022-05600-8>
4. Wang, X., Kinsland, G., Poudel, D., & Fenech, A.: Urban flood prediction under heavy precipitation. *J Hydrol*, 577, 123984 (2019). <https://doi.org/10.1016/j.jhydrol.2019.123984>
5. Gorgoglione, A., Gioia, A., & Iacobellis, V.: A framework for assessing modeling performance and effects of rainfall-catchment-drainage characteristics on nutrient urban runoff in poorly gauged watersheds. *Sustainability*, 11(18), 4933 (2019). <https://doi.org/10.3390/su11184933>
6. Jemberie, M. A., & Melesse, A. M.: Urban flood management through urban land use optimization using LID techniques, city of Addis Ababa, Ethiopia. *Water*, 13(13), 1721 (2021). <https://doi.org/10.3390/w13131721>
7. Kassem, Y., Gökçekuş, H., Diah, F. S., Quoigoah, M. P., & Godwin, M. H.: Prediction of Runoff Using Artificial Neural Networks, MLR Regression, and ARIMA Model (A Case Study: Bared River, Lebanon). In *International Conference on Theory and Applications of Fuzzy Systems and Soft Computing* (pp. 247-255). Cham: Springer Nature Switzerland (2023). https://doi.org/10.1007/978-3-031-25252-5_35
8. Tiwari, D. K., Tiwari, H. L., & Nateriya, R.: Runoff modeling in Kolar river basin using hybrid approach of wavelet with artificial neural network. *J Water Clim Chang*, 13(2), 963-974 (2022). <https://doi.org/10.2166/wcc.2021.246>
9. Mandal, S., & Biswas, S.: Runoff Prediction Using Artificial Neural Network and SCS-CN Method: A Case Study of Mayurakshi River Catchment, India. In *Water Security and Sustainability: Proceedings of Down To Earth 2019* (pp. 27-42). Springer Singapore (2021). https://doi.org/10.1007/978-981-15-9805-0_4
10. Kassem, Y., Gökçekuş, H., Babangida, A., Larmouth, E. J., & Mafela, L. G.: Time Series Forecasting of Solar Power Generation for 5.4 kW Off-Grid PV System: A Case Study in Al Mahmra, Lebanon. In *International Conference on Intelligent Computing & Optimization* pp. 621-631 Springer International Publishing (2022). https://doi.org/10.1007/978-3-031-19958-5_58
11. Kassem, Y., Gökçekuş, H., Diah, F. S., Quoigoah, M. P., & Godwin, M. H.: Prediction of Runoff Using Artificial Neural Networks, MLR Regression, and ARIMA Model (A Case Study: Bared River, Lebanon). In *International Conference on Theory and Applications of Fuzzy Systems and Soft Computing* pp. 247-255, Springer Nature Switzerland (2022). https://doi.org/10.1007/978-3-031-25252-5_35
12. Kassem, Y., Gökçekuş, H., & Mosbah, A. A. S.: Prediction of monthly precipitation using various artificial models and comparison with mathematical models. *Environ Sci Pollut Res*, 30(14), 41209-41235 (2023). <https://doi.org/10.1007/s11356-022-24912-7>
13. Kassem, Y., & Çamur, H.: Effects of storage under different conditions on the fuel properties of biodiesel admixtures derived from waste frying and canola oils. *Biomass Conv. Bioref.*, 8, 825-845 (2018). <https://doi.org/10.1007/s13399-018-0339-1>
14. Kassem, Y., Çamur, H., & Esenel, E.: Adaptive neuro-fuzzy inference system (ANFIS) and response surface methodology (RSM) prediction of biodiesel dynamic viscosity at 313 K. *Procedia Computer Science*, 120, 521-528 (2017). <https://doi.org/10.1016/j.procs.2017.11.274>
15. Berti, M., Bernard, M., Gregoretti, C., & Simoni, A.: Physical interpretation of rainfall thresholds for runoff-generated debris flows. *Journal of Geophysical Research: Earth Surface*, 125(6), e2019JF005513 (2020). <https://doi.org/10.1029/2019JF005513>

16. Crow, W. T., Chen, F., Reichle, R. H., Xia, Y., & Liu, Q.: Exploiting soil moisture, precipitation, and streamflow observations to evaluate soil moisture/runoff coupling in land surface models. *Geophysical Research Letters*, **45**(10), 4869-4878 (2018). <https://doi.org/10.1029/2018GL077193>

Received: January 4, 2025

Accepted: March 15, 2025