Predicting Modified Fournier Index by Using Artificial Neural Network in Central Europe

Endre Harsányi 1,2, Bashar Bashir 3, Firas Alsilibe 4, Muhammad Farhan Ul Moazzam 5, Tamás Ratonyi 1, Abdullah Alsalman 3, Adrienn Széles 1, Aniko Nyeki 6, István Takács 7 and Safwan Mohammed 1,2,*

1 Institute of Land Use, Technical and Precision Technology, Faculty of Agricultural and Food Sciences and Environmental Management, University of Debrecen, 4032 Debrecen, Hungary
2 Institutes for Agricultural Research and Educational Farm, University of Debrecen, Bősörményi 138, 4032 Debrecen, Hungary
3 Department of Civil Engineering, College of Engineering, King Saud University, P.O. Box 800, Riyadh 11421, Saudi Arabia
4 Department of Transport Infrastructure and Water Resources Engineering, Széchenyi István University, Egyetem tér 1, 9026 Győr, Hungary
5 Department of Civil Engineering, College of Ocean Science, Jeju National University, 102 Jejudaehakro, Jeju 63243, Korea
6 Department of Biosystems and Food Engineering, Faculty of Agricultural and Food Sciences, Széchenyi István University, Vár Square 2, 9200 Mesombányovár, Hungary
7 Doctoral School of Humanities, University of Debrecen, Egyetem Tér 1, 4032 Debrecen, Hungary
* Correspondence: safwan@agr.unideb.hu

Abstract: The Modified Fournier Index (MFI) is one of the indices that can assess the erosivity of rainfall. However, the implementation of the artificial neural network (ANN) for the prediction of the MFI is still rare. In this research, climate data (monthly and yearly precipitation (p, P_total) (mm), daily maximum precipitation (P_d-max) (mm), monthly mean temperature (T_avg) (°C), daily maximum mean temperature (T_d-max) (°C), and daily minimum mean temperature (T_d-min) (°C)) were collected from three stations in Hungary (Budapest, Debrecen, and Pécs) between 1901 and 2020. The MFI was calculated, and then, the performance of two ANNs (multilayer perceptron (MLP) and radial basis function (RBF)) in predicting the MFI was evaluated under four scenarios. The average MFI values were between 66.30 ± 15.40 (low erosivity) in Debrecen and 75.39 ± 15.39 (low erosivity) in Pécs. The prediction of the MFI by using MLP was good (NSE_Budapest(SC3) = 0.71, NSE_Pécs(SC2) = 0.69). Additionally, the performance of RBF was accurate (NSE_Budapest(SC4) = 0.68, NSE_Pécs(SC3) = 0.73). However, the correlation coefficient between the observed MFI and the predicted one ranged between 0.83 (Budapest (SC2-MLP)) and 0.86 (Pécs (SC3-RBF)). Interestingly, the statistical analyses promoted SC2 (P_d-max + p + P_total) and SC4 (P_total + T_avg + T_d-max + T_d-min) as the best scenarios for predicting MFI by using the ANN–MLP and ANN–RBF, respectively. However, the sensitivity analysis highlighted that P_total, P_d-max and T_d-min had the highest relative importance in the prediction process. The output of this research promoted the ANN (MLP and RBF) as an effective tool for predicting rainfall erosivity in Central Europe.

Keywords: land degradation; machine learning; climate change; Hungary

1. Introduction

In many regions across the world, the most predominant type of land degradation is soil erosion, which has adverse environmental and socioeconomic consequences [1–3]. Soil erosion is the process of moving soil particles by external forces, such as mass movement, wind, and water [4,5]. In Europe, where a humid climate dominates, water-induced soil erosion is the main form of erosion, which poses a serious environmental concern in many European countries [6]. Furthermore, soil erosion by wind and dust storms is one of the challenges in European countries [7–9].
Soil erosion by water has numerous environmental impacts. For instance, detaching soil particles from the upper layer of the soil causes a deterioration in agriculture productivity through the loss of organic matter, nutrients, and soil depth [10]. Moreover, moving soil particles over vast distances affects the ecosystem service quality in downstream rivers by increasing the sedimentation and the contamination of aquatic life [11,12]. Since measuring soil erosion at a large scale is difficult, expensive, and time consuming, several models have been developed in recent decades to estimate soil erosion [13–15].

In Europe, the Universal Soil Loss Equation (USLE) [15], and its modified version, the Revisited Universal Soil Loss Equation (RUSLE) [14], is the most widely used in quantifying soil erosion at multiple scales across Europe. At large spatial scales, RUSLE is typically the most frequently used model to estimate soil erosion [16]. In the RUSLE model, the average annual soil erosion is calculated by multiplying six factors, including the rainfall erosivity factor (R factor). These factors are slope length (L-factor), soil erodibility (K-factor), slope steepness (S-factor), supporting conservation practices (P-factor), and crop type and management (C-factor). In this sense, rainfall erosivity is considered the most important, as rainfall has a direct impact on detaching and moving the soil particles [15].

Rainfall erosivity is the potential force of raindrops to detach and erode soil particles [17]. As it is one of the main causes of floods and landslides, researchers have highlighted rainfall erosivity as an important indicator to be investigated [18]. The rainfall erosivity factor is calculated using rainfall records with 1–5 min precipitation intervals [19]; however, these records are rarely accessible for long enough in most of the world. As a result, the kinetic energy concept has been widely employed to estimate the rainfall erosivity factor from half-hourly or hourly datasets [20].

To accurately estimate the R factor using the kinetic energy concept, it is necessary to measure both the intensity and the kinetic energy of the rain, but it is highly challenging to achieve this directly since the equipment needed is expensive and measuring the distribution of the rainstorm’s drop sizes is a tedious process [21]. To overcome this, researchers have developed numerous empirical equations that describe the relationship between rainfall intensity and its kinetic energy [22]. To provide a comprehensive review of these equations, Dash et al. [21] compared six of the most universal equations in more detail and provided a deep evaluation of their applicability in calculating the R factor. Alternative methods for calculating the R factor include index techniques, such as the Modified Fournier Index (MFI), especially when high-resolution rainfall records (half-hourly or hourly) are not available. The MFI is one of the methods suggested by Arnoldus [23] for calculating the R factor based on the monthly rainfall data. However, some adjustment is required for calculating the R factor based on the MFI result [24]. The MFI was used to estimate catastrophic erosion by evaluating rainfall erosivity and its association with other meteorological factors [25].

| Country               | Period              | Comment                                      | Reference                       |
|-----------------------|--------------------|----------------------------------------------|---------------------------------|
| NE Spain              | 1997–2006          | *MFI* has poor performance against observed R-value | Angulo-Martínez and Beguería [24] |
| India                 | 1981–2019 (CHIRPS-2.0 Global_daily) | High correlation between *MFI* and R | Raj et al. [26] |
| Uruguay               | 1931–2000          | *MFI* increased in the 1980s and 1990s | Munka et al. [27] |
| Iran                  | 1970–1992          | Roose’s index more efficient than *MFI* | Sadeghi [28] |
| Pannonian basin       | 1961–2014          | Low erosive class of *MFI* was recorded | Lukić et al. [11] |

Recently, the artificial neural network (ANN) and machine learning algorithms have been widely used to predict environmental processes (erosion, contamination, and drought)
in many parts of the world [29]. For instance, the multilayer perceptron neural network (MLPNN) model is one of the most widely used models for predicting hydrological data [30]. Mishra and Desai [31] used the ANN, RBF, and adaptive neural network-based fuzzy inference system (ANFIS) to forecast drought (SPI) at various timescales and found that ANN has better performance than RBF and ANFIS. In Iran, MLP, ANFIS, and multiple linear regression models were used for forecasting precipitation; the output showed that MLP produced better results [32]. Jalalkamali et al. [33] compared stochastic models with the ANN to forecast SPI-9 in Iran, and their results revealed that stochastic models performed better. The different model results depend on the drought index and its scale [34]. More examples of the implementation of the ANN for predicting certain environmental variables are presented in Table 2.

Table 2. Implementation of ANN for predicting some environmental variables in many parts of the world.

| Country | Environmental Phenomenon | Method | Comment | Reference |
|---------|--------------------------|--------|---------|-----------|
| Two watersheds in Australia and France | Rainfall-runoff modeling | EANN and FFNN | EANN is better than FFNN | Nourani [35] |
| Kasilian watershed (Iran) | Soil erosion | ANN and GIS | ANN with GIS perfectly predicts soil erosion | Gholami et al. [36] |
| Langat River basin (Malaysia) | Meteorological drought | ANN-ARIMA | Engagement of Wavelet-based ARIMA–ANN was recommended | Khan et al. [37] |
| Bojnourd (Iran) | Drought forecasting | ANN, ANFIS, support vector machine (SVM) | SVM was the most accurate one | Mokhtarzad et al. [38] |
| South Korea | Groundwater level | ANN and SVM | SVM was better than ANN | Yoon et al. [39] |

Based on the literature, few studies used the ANN to predict rainfall erosivity. However, limited information is available on MFI changes in Central Europe. Thus, the main goals of this research were to: (1) assess the Modified Fournier Index (MFI) as a representative for the erosivity index in three stations in Hungary between 1901 and 2020; (2) evaluate the ability of ANNs (multilayer perceptron (MLP) and the radial basis function (RBF)) to predict the MFI; and (3) rate the importance of input variables in predicting the MFI based on sensitivity analysis (∂). Overall, the implementation of ANN to predict MFI is still less common, which give this work novelty in its field, where the output will serve researchers, planners, and decision makers.

2. Materials and Methods

2.1. Data Collection

Data were collected from the Hungarian Metrological Center (https://www.met.hu/en/eghajlat/magyarorszag_eghajlata/eghajlati_adatsorok/Pecs/adasok/havi_adatosok/, accessed on 1 June 2022). The data included the monthly rainfall (mm), daily maximum precipitation in the month (mm), monthly mean temperature (°C), daily maximum mean temperature in the month (°C), and daily minimum mean temperature in the month (°C) and were collected from three meteorological stations: Budapest (47°30′40″ N, 19°01′41″ E), Debrecen (47°29′44″ N; 21°37′48″ E), and Pécs (46°04′37″ N, 18°13′29″ E). Interestingly, the data cover 120 years from 1 January 1901 to 1 December 2020.

2.2. Modified Fournier Index (MFI)

Rainfall erosivity (R) represents the ability of rain drops to initiate erosion. To calculate the R factor, the measurement of rainfall intensity and rainfall duration is required [40]. As such data are not available in many places in the world, many indices were developed to determine the R factor. The Modified Fournier Index (MFI), which was proposed by
Arnoldus [23], is a widely used index for estimating the R factor. The MFI is based on monthly precipitation ($p_i$) and total yearly precipitation ($P_{total}$):

$$MFI = \sum_{i=1}^{12} \frac{p_i^2}{P_{total}}$$

(1)

The output of Equation (1) can be categorized as presented in Table 3. Based on that, the MFI will have high values where the rainfall values are high. In this sense, regions with high amounts of total annual rainfall and rainfall precipitation concentration will have a high MFI value [41]. However, a strong correlation between MFI and R-factor was recorded in the literature [41,42]. Overall, the calculation of the MFI could provide a realistic estimation of the potential rainfall erosivity factor [43].

Table 3. Description of the MFI erosivity range [42].

| MFI Value | Category |
|-----------|----------|
| <60       | Very low |
| 60–90     | Low      |
| 90–120    | Moderate |
| 120–160   | High     |
| >160      | Very high|

2.3. Predicting Rainfall Erosivity Based on Artificial Neural Network (ANN)

2.3.1. Thermotical Background of ANN Algorithms (MLP and RBF)

The artificial neural network consists of various interconnected neurons, nodes, or perceptrons that are called artificial neurons. Each node transmits a signal to another node; therefore, it can keep the information between various connections and distinguish the patterns [44]. The interconnected node obtains signals, processes them, and transforms them further. The transferring signal between nodes is a real number, and its output can be estimated using a nonlinear function by summing up all the inputs. The output of any network architecture works as an input for the preceding neuron [45,46].

There are several neural networks, but the multilayer perceptron (MLP) is widely used in environmental studies. The MLP connects nodes in a feedforward ANN. The MLP connections between nodes cannot form a cycle. The MLP is sometimes used as any feedforward ANN, and sometimes it refers to a network with various layers [47].

The MLP is a supervised learning technique used in backpropagation for training the dataset and has the ability to split the data that are not linearly separable. These attributes differentiate it from the linear MLP [48]. Ali et al. [49] used the MLP and found that it has the potential to predict drought as one of the ecosystem components in different performance measures. Therefore, in this study, we also used the MLP model. To estimate the $y$ using a three-layer network with $n$ number of neurons in the hidden layers and $m$ number of inputs, we can use Equation (2):

$$y = f \left[ \sum_{j=0}^{n} w_{ij} g \left( \sum_{i=1}^{m} w_{ji} x_i + w_{j0} + w_{j} \right) + w_{0} \right]$$

(2)

Here, weight $w_{ij}$ joined with the $j$th neuron in the hidden layer and the output layer $w_{ji}$ weight make a connection between the $i$th input variable and the $j$th neuron in the hidden layer, where $x_i$ is the $i$th independent variable, $w_{j0}$ is the bias of the $j$th neuron, $g$ is the activation function for the neuron of the hidden layer, and $f$ is the activation function for the output layer [50].

The radial basis function (RBF) is another form of ANN, which was used in this research. The RBF was first proposed in 1988 by Broomhead and Lowe to solve the ill-conditioned problems in interpolation [51]. The RBF is a base of radial networks comprising neural network groups, i.e., a statistical neural network. Euclidean distance is the net input
for the activation function of a neuron between its weight ($w$) and vector ($i$) multiplied by the bias $b$. The equation below (Equation (3)) presents the radial basis function network [50,52]:

$$a = (lw - ib)$$

Despite the differences between these two algorithms (i.e., MLP and RBF), both were used for predicting the MFI values in Central Europe.

2.3.2. Modeling Framework

Input Variable

Based on Equation (1), the only necessary data for calculating the MFI is rainfall data ($p_i$ and $P_{total})$. However, for the modeling approach, we engaged other climatic factors, including the daily maximum precipitation (mm) ($P_{d-max}$), monthly mean temperature (°C) ($T_{avg}$), daily maximum mean temperature (°C) ($T_{d-max}$), and daily minimum mean temperature (°C) ($T_{d-min}$). An overview of the input variable for each station is presented in Figures 1 and 2 and Table 4.

![Figure 1](image_url)

**Figure 1.** Evolution of monthly mean temperature ($T_{avg}$) (A–C), daily maximum mean temperature ($T_{d-max}$) (D–F), and daily minimum mean temperature ($T_{d-min}$) (G–I) in Central Europe (Budapest, Debrecen, Pécs) between 1901 and 2020.

For the modeling approach, five scenarios were adopted, as can be seen in Table 5. The main purpose of adopting different scenarios is to assess the function of the ANN (MLP and RBF) in predicting the MFI based on different input variables. For instance, the first scenario includes all input variables (rainfall (daily + monthly + total) + temperature (monthly)), while the last scenario includes only two rainfall parameters.
Table 4. Descriptive statistical analysis of input variables in three locations, Budapest, Debrecen, and Pécs, between 1901 and 2020.

| Location | Statistic | n * | Min. | Max. | Range | Median | X | SD | SK | KU |
|----------|-----------|-----|------|------|-------|--------|---|----|----|----|
| Budapest | $P_{d-max}$ | 1440 | 0.0  | 115.4| 115.4 | 14.2   | 16.6| 11.5| 1.8 | 7.2 |
|          | $p_i$ | 1440 | 0.0  | 263.1| 263.1 | 42.8   | 48.5| 33.4| 1.3 | 2.9 |
|          | $T_{avg}$ | 1440 | -8.7 | 26.6 | 35.3  | 11.5   | 11.3| 8.0 | -0.1| -1.2|
|          | $T_{d-max}$ | 1440 | -2.9 | 33.1 | 36.0  | 18.1   | 17.4| 7.7 | -0.2| -1.2|
|          | $T_{d-min}$ | 1440 | -20.6| 21.7 | 42.3  | 5.5    | 5.1 | 8.5 | -0.2| -1.0|
| Debrecen | $P_{d-max}$ | 1440 | 0.0  | 80.3 | 80.3  | 13.0   | 15.4| 10.7| 1.7 | 4.2 |
|          | $p_i$ | 1440 | 0.0  | 232.3| 232.3 | 40.3   | 47.2| 32.0| 1.2 | 1.8 |
|          | $T_{avg}$ | 1440 | -10.2| 24.9 | 35.1  | 10.4   | 10.0| 8.3 | -0.2| -1.2|
|          | $T_{d-max}$ | 1440 | -1.6 | 30.7 | 32.3  | 17.4   | 16.4| 7.8 | -0.2| -1.1|
|          | $T_{d-min}$ | 1440 | -23.9| 20.8 | 44.7  | 3.9    | 3.5 | 9.3 | -0.3| -0.9|
| Pécs    | $P_{d-max}$ | 1440 | 0.0  | 97.4 | 97.4  | 15.4   | 18.0| 11.8| 1.7 | 4.6 |
|          | $p_i$ | 1440 | 0.0  | 227.0| 227.0 | 47.7   | 55.0| 35.5| 1.1 | 1.3 |
|          | $T_{avg}$ | 1440 | -8.6 | 25.9 | 34.5  | 11.5   | 11.5| 8.0 | -0.2| -1.2|
|          | $T_{d-max}$ | 1440 | -2.3 | 32.8 | 35.1  | 17.9   | 17.3| 7.4 | -0.2| -1.1|
|          | $T_{d-min}$ | 1440 | -21.8| 21.7 | 43.5  | 4.9    | 4.4 | 8.6 | -0.3| -0.9|

* n: number of observations; Min.: minimum; Max.: maximum; X: mean; SD: standard deviation (σ); SK: skewness (Pearson); KU: kurtosis (Pearson).

Table 5. Developed scenarios for MFI prediction based on ANN (MLP and RBF) in Central Europe.

| Number | Scenarios | Input | Climate Element |
|--------|-----------|-------|-----------------|
| 1      | SC1       | $P_{d-max} + p_i + P_{total} + T_{avg} + T_{d-max} + T_{d-min}$ | Rainfall (daily + monthly + total) + temperature (monthly) |
| 2      | SC2       | $P_{d-max} + p_i + P_{total}$ | Rainfall (daily + monthly + total) |
| 3      | SC3       | $P_{total} + T_{avg} + T_{d-max} + T_{d-min}$ | Rainfall (total) + temperature (monthly) |
| 4      | SC4       | $p_i + P_{total}$ | Rainfall (monthly + total) |

Training, Testing, and Sensitivity Analysis for Different ANN (MLP and RBF) Algorithms

For the five implemented scenarios and ANN (MLP and RBF) algorithms, data were divided randomly into 70% for training and 30% for testing. As this work was conducted in an SPSS environment, the initial conduction for each algorithm was adopted. For instance, the number of layers of hidden units was from 1 to 50, and the training type was Batch (initial Lambda = 0.000005) for the MLP algorithm, while the architecture of the RBF
algorithm was based on automatically finding the number of units in the hidden layer with the normalized RBF as an activation function.

Finally, sensitivity analysis ($\partial$) was used to highlight the relationship between the input variable for each scenario and the predicted MFI, as shown in Figure 3.

Assessing the ANN Performance

To assess the performance of the ANN algorithms (MLP and RBF) in predicting the MFI, four indices were used. The indices are model efficiency (NSE) [53], index of agreement correlation ($d$) [54], root mean square error (RMSE) [55], and Pearson correlation coefficient ($r$) [56], as shown in Table 6.

### Table 6. Indices for evaluation of ANN performance for predicting MFI erosivity.

| Index | Equation * | Range | Note |
|-------|------------|-------|------|
| NSE   | $NSE = 1 - \frac{\sum_{i=1}^{n} (MFI_{cal} - MFI_{prd})^2}{\sum_{i=1}^{n} (MFI_{cal} - MFI_{cal})^2}$ | $-\infty$ and 1 | When the NSE reaches 1, it is a perfect match between $MFI_{cal}$ and $MFI_{prd}$ |
| $d$   | $d = 1 - \frac{\sum_{i=1}^{n} (MFI_{cal} - MFI_{cal})^2}{\sum_{i=1}^{n} (MFI_{cal} - MFI_{cal})^2}$ | 0 to 1 | When $d$ approaches +1, this indicates an ideal agreement between $MFI_{cal}$ and $MFI_{prd}$ |
| RMSE  | $RMSE = \sqrt{\frac{\sum_{i=1}^{n} (MFI_{cal} - MFI_{cal})^2}{n}}$ | 0 to $+\infty$ | A lower RMSE value denotes a perfect match between $MFI_{cal}$ and $MFI_{prd}$ |
| $r$   | $r = \frac{\sum_{i=1}^{n} [(MFI_{cal} - MFI_{cal})(MFI_{cal} - MFI_{cal})]}{\sqrt{\sum_{i=1}^{n} (MFI_{cal} - MFI_{cal})^2} \sqrt{\sum_{i=1}^{n} (MFI_{cal} - MFI_{cal})^2}}$ | $-1$ to 1 | When $r = +1$, this shows an ideal positive linear relationship between $MFI_{cal}$ and $MFI_{prd}$ |

*$MFI_{cal}$: calculated MFI value based on Equation (1); $MFI_{prd}$: predicted value based on the ANN algorithms (MLP and RBF); $MFI_{cal}$: average of calculated values; $MFI_{prd}$: average of predicted values.
Additionally, the Taylor diagram [57] was used to plot the $MFI_{Cal}$ against $MFI_{Prd}$. In this sense, the Taylor diagram provides a full overview of the best model/scenarios (Table 5) based on the correlation and standard deviation.

Finally, it is important to mention that all input and output along with the modeling approach was conducted in IBM SPSS Statistics (V. 24). The SPSS was chosen as it provides a user-friendly platform along with a variety of options that could optimize the output and ANN algorithm. However, we used the initial recommended sets (i.e., batch is the type of training, initial Lambda is 0.0000005, initial Sigma is 0.00005) by SPSS for conducting the modeling.

3. Results

3.1. MFI Variability in Hungary

In the three studied stations, the $MFI$ follows a normal distribution (Figure 4). The average $MFI$ values were between 66.30 ± 15.40 (low erosivity) in Debrecen and 75.39 ± 15.39 (low erosivity) in Pecs (Table 7). In Budapest (central Hungary), the highest $MFI$ (129.17, high) value was recorded in 1955, which corresponds to 898.8 mm of rainfall, while the lowest $MFI$ value (37.21) was recorded in 1997. In Debrecen (eastern Hungary), the highest $MFI$ value was 126.23 (high) in 1977. However, the maximum $MFI$ value in Pecs (southern Hungary) was 121.09 (high) in 1972. Interestingly, the highest frequency of the $MFI$ values was 31 in Budapest, 29 in Debrecen, and 24 in Pecs for the values (65.04, 74.32), (54.28, 63.37), and (81.41, 89.53), respectively (Figure 4).

Table 7. An overview of $MFI$ values in the three studied stations (1901–2020).

| Station   | $n^*$ | Min  | Max  | Range | MD   | Mean  | SD   | SK   | K    |
|-----------|-------|------|------|-------|------|-------|------|------|------|
| Budapest  | 120   | 37.21| 129.17| 91.97 | 68.66| 68.72 | 16.24| 0.90 | 1.43 |
| Debrecen  | 120   | 36.10| 126.23| 90.13 | 64.91| 66.30 | 15.40| 0.59 | 0.76 |
| Pecs      | 120   | 40.83| 121.09| 80.25 | 76.48| 75.39 | 15.39| 0.14 | 0.09 |

* $n$: number of observations; Min: minimum; Max: maximum; MD: median; SD: standard deviation ($n$); SK: skewness (Pearson); K: kurtosis (Pearson).

3.2. MFI Prediction by ANN–MLP and ANN–RBF

In the three stations, a combination of different climate variables (four scenarios) was used by ANN–MLP and ANN–RBF for predicting the $MFI$ values. The predicted $MFI$ values are presented in Figures 5 and 6.

Figure 4. Normal distribution of $MFI$ values for studied stations between 1901 and 2020.

Table 7. An overview of $MFI$ values in the three studied stations (1901–2020).
3.2. MFI Prediction by ANN–MLP and ANN–RBF

In the three stations, a combination of different climate variables (four scenarios) was used by ANN–MLP and ANN–RBF for predicting the MFI values. The predicted MFI values are presented in Figures 5 and 6.

For the ANN–MLP, each scenario exhibited a different performance in predicting MFI values (Figure 5). In Budapest, the Pearson correlation coefficient ($r$) ranged between 0.82 (SC1-MLP) and $r_{MFI \text{observed} \text{vs} \text{MFI predicted}} = 0.83$ for the rest of the scenarios. The $d$ index ranged between 0.88 (SC1-MLP) and 0.9 (SC3-MLP). The efficiency of the ANN–MLP was assessed using the NSE. However, the NSE value was above 0.6, which indicates a good model performance for all scenarios. However, the highest value was NSE = 0.7 in SC3. Interestingly, the highest NSE value and lowest RMSE were recorded in SC3. Based on the statistical indicator, the efficiency of the scenarios in predicting the MFI can be highlighted as follows: SC3 > SC2 > SC4 > SC1. For Debrecen, the ANN–MLP exhibited a good performance (Figure 7). The $r$ values and those of other statistical indicators were lower than those recorded in Budapest. For instance, the $r$ ranged between 0.79 and 0.81, and the NSE between 0.62 and 0.66, while the RMSE was higher than Budapest. Based on the four suggested scenarios, the ANN–MLP performance could be ranked as follows: SC1 > SC2 > SC4 > SC3. Similar to Budapest, the ANN–MLP performance in Pecs was better than that in Debrecen (Figure 7). The $d$ index was higher than 0.88, and the NSE was good (NSE > 0.66). Based on this, we can draw the following rank: SC2 > SC3 > SC1 > SC4.
The output of ANN–RBF algorithm (predicted MFI) and observed one in the studied station under four scenarios (SC1, SC2, SC3, and SC4): (C) evolution of observed and predicted MFI in Budapest (1901–2020); (D) scatter plot $MFI_{\text{observed}}$ vs. $MFI_{\text{predicted}}$ in Budapest; (G) evolution of observed and predicted MFI in Debrecen (1901–2020); (H) scatter plot $MFI_{\text{calculated}}$ vs. $MFI_{\text{predicted}}$ in Debrecen; (K) evolution of observed and predicted MFI in Pecs (1901–2020); (L) scatter plot $MFI_{\text{calculated}}$ vs. $MFI_{\text{predicted}}$ in Pecs.

Similar to ANN–MLP, the ANN–RBF showed a good ability to predict the MFI under different scenarios (Figure 6). In Budapest and Debrecen, SC4 had the highest correlation $r_{\text{MFI, pred}}$ (0.85, 0.82), with the highest and lowest NSE, respectively, which indicates that SC4 (ANN–RBF) ($p_l + P_{\text{total}}$) is the best scenario for Budapest and Debrecen. In this sense, the scenarios can be ranked for both stations as follows: SC4 > SC2 > SC3 > SC1 (Figure 7). In Pecs, SC3 (ANN–RBF) ($P_{\text{total}} + T_{\text{avg}} + T_{d_{-\text{max}}} + T_{d_{-\text{min}}}$) outperformed the rest of the scenarios ($r_{\text{MFI, pred}} = 0.86, d = 0.92, \text{NSE} = 0.73, \text{RMSE} = 7.8$). However, the performance of the four scenarios can be ranked as SC3 > SC4 > SC2 > SC1.
than those recorded in Budapest. For instance, the \( r \) ranged between 0.79 and 0.81, and the NSE between 0.62 and 0.66, while the RMSE was higher than Budapest. Based on the four suggested scenarios, the ANN–MLP performance could be ranked as follows: SC1 > SC2 > SC4 > SC3. Similar to Budapest, the ANN–MLP performance in Pecs was better than that in Debrecen (Figure 7). The \( d \) index was higher than 0.88, and the NSE was good (NSE > 0.66). Based on this, we can draw the following rank: SC2 > SC3 > SC1 > SC4.

Figure 7. Performance analysis of ANN (MLP and RBF) algorithms in predicting MFI values under four scenarios (SC1, SC2, SC3, and SC4) at three stations in Budapest (BUD), Debrecen (DB), and Pecs (PE): (A) Pearson correlation coefficient (\( r \)), (B) index of agreement correlation (\( d \)), (C) model efficiency (NSE), and (D) root mean square error (RMSE).

The Taylor diagram (Figure 8) reveals that SC2 and SC3 for Budapest and SC1 and SC2 for Debrecen and Pecs are the best scenarios in terms of the ANN–MLP (Figure 8). However, for the ANN–RBF, SC4 was the most appropriate scenario for Budapest and Debrecen, while SC3 was the best one for Pecs. Overall, these analyses promoted SC2 (\( p_{d_{\text{max}}} + p_{i} + P_{\text{total}} \)) and SC4 (\( P_{\text{total}} + T_{\text{avg}} + T_{d_{\text{max}}} + T_{d_{\text{min}}} \)) as the best scenarios for predicting MFI using the ANN–MLP and ANN–RBF, respectively.
However, for the ANN–RBF, SC4 was the most appropriate scenario for Budapest and Debrecen, while SC3 was the best one for Pecs. Overall, these analyses promoted SC2 ($P_{\text{d-max}} + p_i + P_{\text{total}}$) and SC4 ($P_{\text{total}} + T_{\text{avg}} + T_{d\text{-max}} + T_{d\text{-min}}$) as the best scenarios for predicting MFI using the ANN–MLP and ANN–RBF, respectively.

Figure 8. Taylor diagram for assessing the performance of ANN (MLP and RBF) algorithms in predicting MFI. (a,b) Budapest. (c,d) Debrecen. (e,f) Pecs.

3.3. Comparing between ANN–MLP and ANN–RBF in MFI Prediction

To compare the outputs of each algorithm in each station, the outputs were plotted in a Taylor diagram (Figure 9). The main point of this step is to test all the scenarios for both algorithms against the calculated MF. For Budapest and Debrecen stations, the RBF-SC4 followed by the MLP-SC2 was the best predictor. In Pecs, the RBF-SC3 followed by the MLP-SC1 was superior compared to the others. Notably, in the three stations, the RBF-SC1 had the worst performance (Figure 9). Interestingly, the RBF outperformed the MLP.
the MLP-SC1 was superior compared to the others. Notably, in the three stations, the RBF-SC1 had the worst performance (Figure 9). Interestingly, the RBF outperformed the MLP.

Figure 9. Analysis of ANN–MLP and ANN–MLP RBF performance in predicting MFI using a Taylor diagram. (A) Budapest. (B) Debrecen. (C) Pecs.

3.4. Independent Variable Importance and Sensitivity Analysis

The main goal of sensitivity analysis is to highlight the importance of the input variables in the prediction process. For the MLP in the Budapest station, the $P_{total}$ had the highest importance in all the suggested scenarios ($\partial_{SC1} = 0.46; \partial_{SC2} = 0.86; \partial_{SC3} = 0.79; \partial_{SC4} = 0.95$), followed by $T_{avg}$ in SC1, and $T_{d-min}$ in SC3 (Figure 10). For the RBF in the same station, the $P_{total}$ also had the highest importance ($\partial_{SC1} = 0.45; \partial_{SC2} = 0.64; \partial_{SC3} = 0.44; \partial_{SC4} = 0.88$). However, other independent variables exhibited a good level of importance. For example, in SC1 $T_{d-min}$, $T_{d-max}$ and $p_i$ showed $\partial_{SC1}$ importance ranging between 0.13 and 0.11, while in SC2, the $p_i$ importance reached $\partial_{SC2} = 0.2$ (Figure 10).
Figure 10. Sensitivity analysis for ANN (MLP and RBF) algorithms in the studied station under four scenarios.

For the second station (Debrecen), both algorithms showed that $P_{\text{total}}$ has an important role in $MFI$ prediction. In the MLP, the importance value reached $\delta_{SC4} = 0.73$, while it was $\delta_{SC4} = 0.82$ in the RBF. Notably, the next important variable was the $p_i$, with $\delta_{SC2-MLP} = 0.24$ and $\delta_{SC2-RBF} = 0.22$. At the Pecs station, the importance of the $P_{\text{total}}$ was more pronounced for both the MLP ($\delta_{SC4} = 0.97$) and RBF ($\delta_{SC4} = 0.83$) (Figure 10).

Based on the four scenarios and both ANN (MLP and RBF) algorithms, the sensitivity analysis showed that $P_{\text{total}}$, $p_i$, and $T_{d\text{-min}}$ had the highest relative importance in the prediction process.

4. Discussion

In this research, the $MFI$ was calculated for tracking rainfall erosivity in Central Europe; then, two ANN (RBF and MLP) algorithms were tested to assess their ability in the prediction of the $MFI$. At the three studied stations, the $MFI$ values ranged from very low to high (1901–2020) (Table 7). Previously, De Luis et al. [41] analyzed the erosivity...
trend in Western Europe (Iberian Peninsula) and detected a notable decrease in rainfall erosivity based on the MFI (1951–2000). For the Netherlands, Lukić et al. [10] reported that the MFI values ranged between 77.93 and 97.27 (1957–2016). However, changes in erosivity class from low to moderate were reported in the same study. These changes in rainfall erosivity in Europe can be mainly explained by climate change (i.e., extreme events: flood and drought), which largely affects the precipitation patterns, not only in Europe but all over the world [58–61].

The output of RBF and MLP showed that the RBF outperformed the MLP. However, both algorithms were perfectly capable of predicting the MFI values, with some differences. The differences between the output could be explained by the way that each algorithm works. The necessary step for the proper functioning of the NN is to optimize the weights, known as calibration. Different types of algorithms can be used to optimize the weight, e.g., back propagation [62] and Levenberg–Marquardt [63]. These algorithms can minimize the disparity between forecasted and observed values by adjusting the network weight [46].

Generally, the ANN works on the principle of the training dataset. There are various kinds of neural network (NN) models, but usually, two models are used in prediction applications, i.e., recurrent network and feedforward network. The backpropagation algorithm is used to train both models [49–64]. When the backpropagation algorithm is used to change the weight of neurons, it works on the gradient descent method (weights change in downward direction). The signal strength between nodes is directly dependent on the weights of neurons [49]. Feedforward NN is a basic type, and it is capable of estimating constant and integral functions.

The network architecture of MLP comprises neurons put together into layers. The MLP contains three layers of nodes, i.e., input, hidden, and output layers. The MLP can have one or more hidden layers with various numbers of neurons. In addition to the input node, the hidden and output nodes are considered neurons [65]. When we used the MLP to study rainfall erosivity (MFI), the input layer contained the variables \((P_{d-max}, p_i, P_{total}, T_{avg}, T_{d-max}, \text{and } T_{d-min})\), and the output layer presented the predicted MFI (Figure 3), while the hidden layer included a nonlinear function and utilized weight for the input layer. Neurons in the hidden layer work in a trial and error approach [34].

The MLP and RBF consist of three network layers; however, the main difference between the RBF and MLP is that the RBF’s hidden and output layers are different, unlike those of the MLP [66]. The hidden layer neurons are nonlinear, while the output layer neurons are linear in the RBFs. The nonlinear hidden layer neuron plays a significant role in the nonlinear modeling task [67]. The RBF network is simpler compared to MLP. However, the MLP is more successfully implemented in various complex problems. The RBF is a local approximation network, and its output can be estimated by hidden units in a local receptive field. The MLP network works globally, and its output is determined by all the neurons [68]. Despite the similarity between both algorithms, the differences in the architecture process led to different output and accuracy (Figures 7–9).

Overall, the implementation of the ANN for predicting the MFI or other hydrological and environmental variables was proven to be a useful tool for predicting and forecasting [69]. However, the output of this research could be useful for local planners on a county scale for predicting the MFI values based only on monthly and yearly rainfall.

5. Conclusions

Land degradation is a major issue all over the world due to its negative impact on the agroecosystem and environmental components. Recently, machine learning and the artificial neural network have been implemented in environmental research for predicting natural hazards. In this research, ANN (MLP and RBF) algorithms were implemented to predict the MFI as a representative of erosivity factor (soil erosion) in Central Europe. Five scenarios with different inputs (rainfall and temperature) were suggested for exploring the accuracy of ANN (MLP and RBF) algorithms. The output of this research can be summarized as follows:
1- The MFI ranged between 91.97 (Budapest) and 80.25 (Pecs), with a notable decrease in MFI values (1901–2020).

2- The SC2 \((P_{d-max} + p_i + P_{total})\) was the best scenario for predicting the MFI using the ANN–MLP.

3- The SC4 \((P_{total} + T_{avg} + T_{d-max} + T_{d-min})\) was the most accurate scenario for predicting the MFI by using the ANN–RBF.

4- The sensitivity analysis revealed that \(p_i\) followed by \(P_{total}\) are the most important input variables for predicting MFI values.

It is good to mention that this research was only focused on MFI as one of the factors that contribute to soil erosion based on the monthly rainfall data. Some other factors such as land use (agricultural areas), soil properties (i.e., texture, structure), vegetation cover, and inclination angle of rainfall streams were not considered in this research.

Local planers, environmental organizations, and decision makers will be able to use the output of this research, where the prediction of the MFI could be performed to a satisfactory level based on the total rainfall in the target regions. In the next steps, other machine learning methods will be implemented to test their accuracy in the prediction of the MFI. However, the output of this research could serve as a good result for both scientific and industrial communities.

Author Contributions: Conceptualization: S.M. and E.H.; writing—original draft: S.M.; methodology: S.M.; writing—review and editing: F.A., M.F.U.M., A.S., I.T., B.B., A.A., T.R., A.N. and E.H. All authors have read and agreed to the published version of the manuscript.

Funding: Project no. TKP2021-NKTA-32 was implemented with the support provided by the National Research, Development, and Innovation Fund of Hungary, financed under the TKP2021-NKTA funding scheme. This research was supported by the RESEARCHERS SUPPORTING PROJECT, grant number RSP-2021/296, King Saud University, Riyadh, Saudi Arabia.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are freely available at: https://www.met.hu/en/eghajlat/magyarorszag_eghajlata/eghajlati_adatsorok/Pecs/adatok/havi_adatok/ (accessed on 1 June 2022).

Acknowledgments: The authors would like to thank Debrecen university for their unlimited support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Luetzenburg, G.; Bittner, M.J.; Calsamiglia, A.; Renschler, C.S.; Estrany, J.; Poepppl, R. Climate and land use change effects on soil erosion in two small agricultural catchment systems Fugnitz–Austria, Can Revull–Spain. Sci. Total Environ. 2020, 704, 135389. [CrossRef] [PubMed]

2. Mohammed, S.; Al-Ebraheem, A.; Holb, I.J.; Alsafadi, K.; Dikkeh, M.; Pham, Q.B.; Linh, N.T.T.; Szabo, S. Soil management effects on soil water erosion and runoff in central Syria—A comparative evaluation of general linear model and random forest regression. Water 2020, 12, 2529. [CrossRef]

3. Mohammed, S.; Alsafadi, K.; Talukdar, S.; Kiwan, S.; Hennawi, S.; Alshihabi, O.; Sharaf, M.; Harsanyie, E. Estimation of soil erosion risk in southern part of Syria by using RUSLE integrating geo informatics approach. Remote Sens. Appl. Soc. Environ. 2020, 20, 100375. [CrossRef]

4. Bullock, P. Climate Change Impact. In Encyclopedia of Soils in the Environment; Cranfield University–Silsoe: Silsoe, UK, 2005; pp. 254–262. [CrossRef]

5. Mohammed, S.; Hassan, E.; Abdo, H.G.; Szabo, S.; Mokhtar, A.; Alsafadi, K.; Alsafadi, K.; Al-Khoury, I.; Rodrigo-Comino, J. Impacts of rainstorms on soil erosion and organic matter for different cover crop systems in the western coast agricultural region of Syria. Soil Use Manag. 2021, 37, 196–213. [CrossRef]

6. Panagos, P.; Katsoyiannis, A. Soil erosion modelling: The new challenges as the result of policy developments in Europe. Environ. Res. 2019, 172, 470–474. [CrossRef]

7. Hojan, M.; Rurek, M.; Wieclaw, M.; Krupa, A. Effects of extreme dust storm in agricultural areas (Poland, the Greater Lowland). Geosciences 2019, 9, 106. [CrossRef]

8. Szczepanik, D.M.; Ortiz-Amezquita, P.; Heese, B.; D’Amico, G.; Stachlew ska, I.S. First Ever Observations of Mineral Dust in Wintertime over Warsaw, Poland. Remote Sens. 2022, 14, 3788. [CrossRef]
67. Chen, S.; Cowan, C.F.N.; Grant, P.M. Orthogonal Least Squares Learning Algorithm for Radial. *IEEE Trans. Neural Netw.* **1991**, *2*, 302–309. [CrossRef]

68. Yu, H.; Xie, T.; Paszczynski, S.; Wilamowski, B.M. Advantages of radial basis function networks for dynamic system design. *IEEE Trans. Ind. Electron.* **2011**, *58*, 5438–5450. [CrossRef]

69. Nourani, V.; Mogaddam, A.A.; Nadiri, A.O. An ANN-based model for spatiotemporal groundwater level forecasting. *Hydrol. Process.* **2008**, *22*, 5054–5066. [CrossRef]