Application of Artificial Neural Network in Wildfire Early Prediction Systems

HRISTINA NIKOVA, SNEJANA YORDANOVA, RADOSLAV DELIYSKI
Department of Automation,
Technical University of Sofia,
8 Kl. Ohridski blvd., 1000 Sofia,
BULGARIA

Abstract: - The preservation of forest ecosystems is of vital importance to life on our planet. The increased losses of forests due to fires make the task of forest fire prevention of crucial significance. The present paper describes the development of an artificial neural network (ANN) for forest fire early prediction. The ANN predictor consists of two layers with 5 neurons in the hidden layer. It is trained through backpropagation of an error learning algorithm and is validated to provide prediction with a high degree of accuracy. An additional advantage of the designed predictor is the use of a limited number of input data based on weather and moisture conditions and of an output of a prior computed probability for fire. The training and validation datasets consist of 82 records of real measurement data. The developed and validated ANN can contribute to improvement of the current forest fire prediction systems.

Key-Words: - wildfire, data processing, neural network, training, validation, prediction.

Received: March 25, 2023. Revised: October 21, 2023. Accepted: December 15, 2023. Published: December 31, 2023.

1 Introduction

Forest ecosystems play a very important role for life on Earth in general, covering nearly one-third of all land on the planet. They have a considerable contribution to many vital processes in various aspects of health, life, and economics. Due to forest diversity and its complicated structure, such processes as regulation of the climate and the carbon cycle, are possible. Furthermore, these ecosystems are home to up to 80% of terrestrial biodiversity on Earth. Forests also help in the prevention of erosion, and enrich and conserve soil.

The loss of forest can cause a significant ecosystem disaster. Since the twentieth century, the increasing numbers of wildfires have turned them into the main danger for forest destruction. The latest research predicts increasing in wildfire occurrence by 14% by the end of 2030 and even 30% by the end of 2050, [1]. That makes the need for forest fire prevention and detection crucial. Effective fire management is an approach for the better preservation of forest ecosystems.

Looking forward to the next few decades, an escalating challenge in managing wildfires will be faced. A large research in both improving knowledge of fire physics on one hand (theoretical part) and fire modeling techniques on the other (practical part), is expected. Despite the notable advances in technology development in the last years, the problems with modeling and prediction of forest fire occurrence persist.

The fire prediction is based on classical statistical techniques, [2], [3], and intelligent approaches known also as soft computing, [4], [5]. The basic intelligent methods include Artificial Neural Networks (ANN), Fuzzy Logic (FL) Models, [6], and Adaptive Neuro- Fuzzy Inference System (ANFIS), [6], which integrates both ANN and FL principles. They all are trained to discover valuable relationships between a great number of measured or estimated variables for the factors related to the forest fires occurrence and the predicted variable for fire. Some of the most popular training algorithms are the Support Vector Machine (SVM) – a supervised algorithm for linear and non-linear classification and regression problems, and the backpropagation of error algorithm (BP), [5], [6], [7], [8], [9], [10].

The ANN has become one of the leading methods in the prediction of different processes in a great number of areas such as power energy consumption, environmental science, water resource planning and management, transport, agriculture, medicine, etc., [11], [12], [13], [14], [15]. In the field of forest fire detection and prevention, the ANN can find a considerable application due to their abilities to learn from examples and to generalize the knowledge into the learning process to new and unseen examples. Another significant advantage is their potential to find solutions to difficult problems, which are rich in data but poor in models.

The basic input data used in most predictors are the measured temperature, relative humidity, wind speed, precipitation (rain), [6], [16], [17], [18], vegetation, [19], [20], as well as satellite images, [21]. They are extended in some predictors with indices computed from the measured variables, time, topographic, and spatial variables, [18], [20]. The predicted variables are the burned forest fire area, [6], [16], [17], [18], the fire danger index defined by the daily number of forest fires, [19]; fire spatial and temporal probability expressed as dates and locations of fire events, etc., [19], [20], [21]. The ANN-based predictors are most often multi-layer networks trained by BP. All developed predictors are duly validated.

Fuzzy logic predictors of the total burned area are suggested in [6]. They are based on Fuzzy Inductive Reasoning and ANFIS tuned by gradient descent and least-square algorithms. The data cover 17 years. A meteorological station records 12 input variables - the first five of the basic and also spatial location, month, day, and 4 indexes characterizing the fuel moisture according to the Canadian Fire Weather Index System. Due to the limitations of the fuzzy model, the most significant five variables are selected - the first four of the basic and Fine Fuel Moisture codes.

All 12 variables are used in the development of an ANN predictor of the total burned area [16]. The ANN has one hidden layer of heuristically determined 36 neurons.

Two ANN predictors for the risk of forest fire occurrences, defined by a fire danger index on a scale of 1–4 (1 for the lowest and 4 for fire the highest danger) depending on the daily number of forest fires, are suggested in [19]. The input data are from fixed weather stations across the country covering 8 years and consist of two variables - relative humidity and cumulative precipitation, selected from six weather variables – the minimal and the maximal temperatures, the average humidity of the day, the solar radiation, the average wind speed, and the cumulative precipitation. The ANN has three layers with 4 neurons in each hidden layer and 1 neuron in the output layer. All neuron activation functions are hyperbolic tangent sigmoid. The first ANN predictor is trained by Levenberg–Marquardt BP while the second ANN predictor - by SVM with a Gaussian kernel function.

In [20] an ANN for identifying areas of forest fires (ignition) by predicting their spatial probability, i.e. the dates and locations of fire events is developed. The input data cover 10 years and contain 12 variables including topographic, anthropogenic, hydrologic, vegetation, and land (identified features include elevation, aspect, slope, tree cover density, forest type, settlement proximity, settlement density, water proximity, power line proximity, normalized vegetation density index, modified normalized water density index, and land use and cover. The activation function in the two ANN hidden layers is rectified linear (ReLU) and in the output layer - logarithmic sigmoidal (logsig).

In [17], a two-layer ANN is trained to predict the forest fire spread by evaluating the historical forest fire disturbance data– time and location from 18 years. The 16 input variables characterize climatic, topographic, combustible factors, and land cover (the 4 basic, wind direction, slope and slope direction, elevation, vegetation, surface water content, roads, railways, settlements, lakes, ditches, wells). The activation function of the neurons in the hidden layer is hyperbolic tangent and in the output layer – “logsig”.

A deep learning ANN for early warning of forest fire occurrence based on Long- and Short-Term Memory network (LSTM) was developed in [18]. The data used are from 536 historical records for a set of 12-dimensional meteorological measured influencing variables – special coordinates, month of the year, day of the week, temperature, relative humidity, wind speed, rain, 4 fuel moisture indexes derived from the Canadian Fire Weather Index System - Fine Fuel Moisture Code, Duff Moisture Code, Drought Code (DC) and the Initial Spread Index as well as the burned area. The designed ANN consists of a 6-layered deep architecture (one input layer, one LSTM layer, two fully connected layers, one dropout layer, and one regression layer). The number of neurons in the LSTM layer is set to 100. The dropout probability is set to 0.5. The sigmoid function is used to scale the signals in the interval [0, 1]. Similarly, the hyperbolic tangent function scales the output of a particular memory cell. Four other machine learning methods are applied for the prediction of forest fire to the same dataset - Decision Trees with fine tree architecture, Linear Regression, SVM with a linear kernel function, and Narrow Neural Network with ReLU activation function. The result shows that LSTM outperforms the rest of the methods.

In [21], a Region-Based Convolutional Neural Network (R-CNN) is trained to predict forest fire occurrence based on satellite images. R-CNN object detection model has full image convolutional features. The raw images and the training dataset images are fed respectively into the historical records and the data from the previous n-weeks. SVM is used for classifying the candidate region as either real fire or non-fire.

A common characteristic of the developed ANN predictors is the complex structure of several layers and many neurons. Besides, great resources are required to collect the data needed. The predicted output is often not directly related to the fire occurrence. As a whole, each developed ANN predictor serves a specific purpose related to a specified area and data stations. However, a new comparative analysis of the advantages and drawbacks of various methods for wildfire prediction including ANN, Binary regression as well as normalized Burning index, Energy Release Component, and Severe Fire Danger Index is presented in [22], based on data from the same period and geographical area.

The present investigation aims to develop a general approach for predicting forest wildfire occurrence with a high degree of accuracy under a limited input data sample using simple ANN. The novelties conclude in the algorithm for the design of a multi-layer ANN predictor and the trained ANN to predict the probability of forest fire occurrence based only on meteorological data and fuel moisture conditions. The training data is collected from the region of Whitmore, North California, USA.

The geographical location of Whitmore is 40°37'45"N latitude and 121°54'59"W longitude. The area is located within the Cascade maintenance range. Figure 1 presents the satellite image and average data of some of the climate characteristics of the region: average maximum temperature, wind speed, and rainfall. The climate is the Mediterranean type classified as "Csb" based on the Köppen Climate Classification system, [23], with hot and dry summers. The United States system for estimating forest fire danger - the National Fire Danger Rating System (NFDRS), [24], [25], [26], based on its fuel models, [27], classifies the region as "Fuel Model X". The main vegetation in the area is 30 years old mixed chaparral and dense brush fields with a height of more than 1,85 m. The amount of dead fuel available in the region increases the risk of intensive wildfires, especially in the summer season and the necessity of significant efforts for forest protection.

The next sections of this article are organized as follows. The development of an ANN fire predictor is presented in Section 2. The validation results are described in Section 3. Section 4 contains the conclusion and a vision for future research.

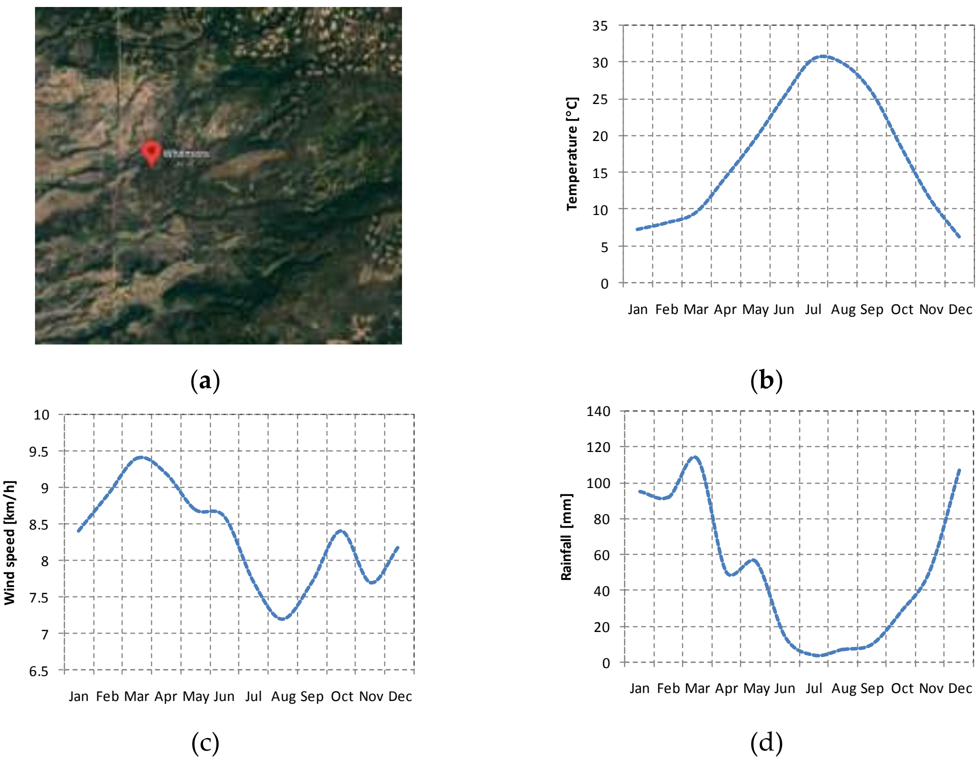

Fig. 1: (a) Topographic location of the area; (b) Average high temperature per month; (c) Average wind speed per month; (d) Average rainfall per month

2 Development of ANN Fire Predictor

The prediction of fire is based on a great number of measured and expert-assessed variables that describe various weather and moisture conditions. A two-layer ANN with nonlinear activation functions of the hidden neurons and a linear activation function of the output neuron trained by backpropagation of error learning algorithm is accepted as the most proper for a fire predictor. Such ANNs are commonly used to approximate any complex nonlinear relationship of many variables.

2.1 ANN Training Data Collection and Pre-processing

The ANN training data is collected from open-access databases of wildfire occurrences of the NFDRS. The real-time data is registered once per hour – 24 hours per day. The sample used for training and validation of the ANN fire predictor is extracted from several 12 of the biggest forest fires based on the affected area for 10 years between 2007 and 2017 in the region of Whitmore, North California, USA. Ten /variables that characterize the weather and the fuel moisture condition are selected as the most contributing factors for the forest fire occurrence, [28] and are included in the input data sample. They are the current temperature (Temp), relative humidity (RH), total solar radiation for the day (SolR), rain in terms of total precipitation amount (Rain), the maximum temperature during the day (maxT), the average wind speed within 10 min period before measurement (Wind), fuel moisture - FM1, FM100, FM1000 as well as Keetch Byram Drought Index (KBDI), [28]. The values of the variables are depicted in Figure 2. No precipitation is registered except for 6 records (from 12 to 17) when the precipitation amount of 0.508 mm remains constant. Therefore, “rain” is not depicted in Figure 2. The record numbers 11, 17, 23, 29, 36, 49 53, 59, 66, and 72 are registered during the fire occurrence and presented with solid black vertical lines.

FM1 is fuel moisture for the dead 1-hour time lag fuels with the size of less than 0,65 cm in diameter, FM100 is fuel moisture for the dead 100-hour time lag fuels with a size from 2,50 to 7,60 cm in diameter and FM1000 is fuel moisture for the dead 1000-hour time lag fuels with size between 7,60 and 20,30 cm in diameter.

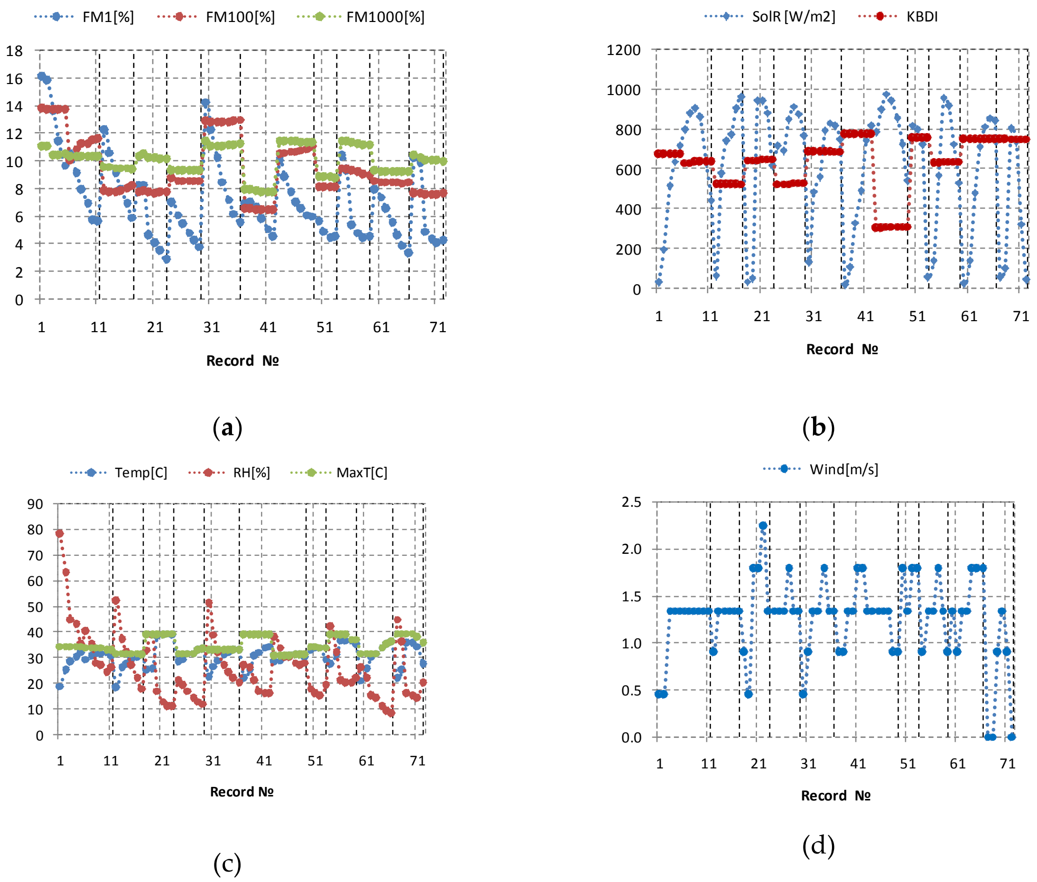

Fig. 2: The 72 records of (a) FM1, FM100, FM1000; (b) SolR and KDBI; (c) Temp, RH and MaxT; (d) Wind speed

Dead Fuels are naturally occurring fuels whose moisture content depends on environmental conditions and vegetation characteristics. Fuel moisture measures the amount of water contained in vegetation.

When the fuel moisture content is less than 30 percent, the fuel is considered as dead. The lower the fuel moisture content is, the more easily the fires ignite and the more rapidly they spread. If the fuel moisture content is high, the probability of fire ignition is very low and even if a fire starts it will not spread rapidly.

The small fuel moisture (FM1) class includes grass, leaves, and small plants or roundwood. Also the upper layer of litter on the forest floor. In the Larger fuels (FM100, FM1000) classes are mainly included roundwood with large sizes.

KBDI is a dimensionless index with a lower value equivalent to 0 and upper – to 800, estimating the amount of precipitation that is needed to bring the soil back to saturation (a value of 0 means complete saturation of the soil). It is a good indicator of the drought level and the availability of drought fuel.

Each n^{th} (n=1÷N) record is a combination of measured or assessed values for the variables that make the vector $\mathbf{p}_{nx10}$:

$\mathbf{p}_{nx10}$= [Temp, RH, SolR, Rain, maxT, Wind, FM1, FM100, FM1000, KBDI].

All N combinations build the input data matrix $\mathbf{P}_{Nx10}$= [$\mathbf{p}_{nx10}$].

The input data contains *Ntotal* =82 data vectors $\mathbf{p}_{nx10}$ – 12 vectors with real fire occurrence and 70 vectors with no presence of fire, respectively.

The data are selected to cover variable conditions and to ensure that there is no linear dependence between the different vectors.

The collected data are split into N=72 training vectors (87%) in $\mathbf{P}_{72x10}$ used for predictor modeling and N^V=10 validation vectors (13%) in $\mathbf{P}^{\mathbf{V}}_{10x10}$ used for predictor validation.

The corresponding target vectors $\mathbf{T}_{72x1}$ and $\mathbf{T}^{\mathbf{V}}_{10x1}$ consist of the computed for each data vector fire probability p_f, [28].

The vectors p_f are obtained using Binary Logistic Regression Analysis for forest wildfire occurrence through the IBM® statistical software platform - SPSS®, [29], with the weather and fuel moisture conditions as input data. This approach estimates a nonlinear relationship between a set of dependent variables (weather and fuel moisture conditions) and a single independent variable of binomial (categorical) type representing “Fire” or “No fire” occurrence. The weight coefficients in the regression equation measure the unique strength of the relationship between dependent and independent variables and are obtained by Newton-Raphson root-finding method for non-linear functions. The binary regression analysis provides additionally valuable information about the significance of the dependent variables and, accuracy of the model and shows high level of the total success rate, [28]. The used as a target statistical model for computation of p_f derived in [28], is:

$$p_f = \frac{e^F}{1+e^F}, \tag{1}$$

where:
$F = C_0 + C_1Rain + C_2MaxT + C_3Wind + C_4FM1 + C_5FM100 + C_6FM1000 + C_7KBDI$;
C0= -6.43; C1= 13.211; C2= 1.159; C3= -2.005 C4= -2.767; C5= 2.107; C6= -2.039; C7= -0.015.

2.2 ANN Model Training

The general requirement for determination of the number of the free parameters (w) of an ANN is as follows:

$$w \approx S1 * (R + S2) \ll \frac{q}{5}, \tag{2}$$

where $S1$ and $S2$ are the number of neurons in the hidden and the output layer respectively, R is the input vector size and q is the size of the training set. Thereby, the size of the hidden layer should be at least (rule of thumb):

$$S1 >= \frac{q}{5*(R+S2)} \tag{3}$$

Accounting for the size N=72 of the training sample and the number of the input variables R=10, the simplest possible ANN is a two-layered feed-forward. The number of the neurons in the hidden layer is $S1$=5 which ensures that the ANN with such an architecture satisfies also the criterion for preventing of over-fitting and under-fitting in training by the backpropagation method $R*S1+S1+S1*S2+S2 \leq N$ or 61≈72 (here $S2$=1).

The two-layer feed-forward ANN accepted is trained to predict forest fire by the backpropagation of error learning algorithm, [30], for minimization of a selected cost function - the Mean Square Error (MSE) at the output layer.

The algorithm is illustrated in Figure 3. The principle is to compare the desired output (target) **T** and the computed ANN output **Y2** and the obtained in this way error e=T-Y2 to be used to update the

ANN weights and biases by gradient based Levenberg-Marquardt BP.

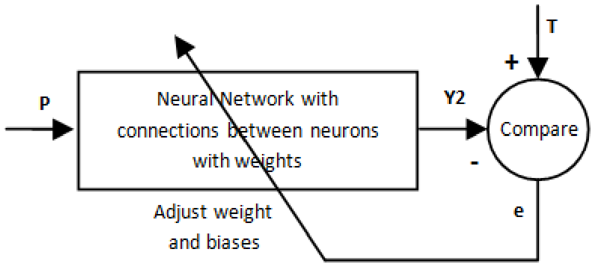

Fig. 3: Neural network backpropagation learning algorithm

The update **ΔW**=[ΔWjl] of the weights **W** and the update **ΔB**=[Δbj] of the biases **B** starts from the second layer end moves to the first layer:

$W2_{jl}(k+1) = W2_{jl}(k) + \Delta W2_{jl}(k), \Delta W2_{jl}(k) = \alpha\delta_j Y1_l$ (4)

$b2_j(k+1) = b2_j(k) + \Delta b2_j(k), \Delta b2_j(k) = \delta_j = = e_j(k)$ (5)

$W1_{li}(k+1) = W1_{li}(k) + \Delta W1_{li}(k), \Delta W1_{li}(k) = \alpha.\delta_l.p_i$ (6)

$b1_l(k+1) = b1_l(k) + \Delta b1_l(k), \Delta b1_l(k) = \delta_l = = e_l(k)$ (7)

where:

- $\delta_j = (T_j - Y2_j).f2'(n2_j) = e_j.f2'(n2_j)$,
$\delta_l = f1'(n1_l).\sum_{m=1}^{S2}\delta_m.W2_{lm}$ with
$\delta_m = (T_m - Y2_m).f2'(n2_l) = e_m.f2'(n2_m)$;

-**Y1**= [$Y1_1 \ldots Y1_{S1}$] and **Y2**= [$Y2_1 \ldots Y2_{S2}$] are the output vectors of the first and the second layer respectively with $S1$ and $S2$ the number of the neurons in the first and the second layer;
- p_l is the i^{th} input variable (i=1÷R);
- k is the number of epochs;
-$W2_{jl}$ is the weight between the output of the l^{th} neuron (l=1÷$S1$) of the first layer and the input to the j^{th} neuron (j=1÷$S2$) from the second layer;
- $b2j$ is the bias of the j^{th} neuron of the second layer;
- f2(n2) are the activation functions of all neurons in the second layer, f2′ are the derivative of f2 concerning n2,
$n2_j = (W2_{j1}Y1_1 + W2_{j2}Y1_2 + \cdots W2_{j\,S1}Y1_{S1}) + b2_j$;
- the weights $W1_{li}$ and the biases $b1$ are determined analogically, f1(n1) with
$n1_l = (W1_{l1}p1_1 + W1_{l2}p1_2 + \cdots W1_{lR}p1_R) + b1_l$;
- α is a learning rate.

Levenberg-Marquardt BP improves the convergence of the tuning procedure for the weights and the biases (2)-(5) by introducing second order derivatives of the cost function with respect to the unknown weights and biases.

The tuning process is repeated until the end condition is satisfied – a reached desired minimum value of MSE=$\frac{1}{N}\sum_{1}^{N}(\mathbf{T}-\mathbf{P})^2$ or an elapsed number of epochs (iterations). The monotonously reduced MSE with the epochs in the ANN fire predictor training is a proof of lack of over-fitting or under-fitting. The initial weights and biases are generated as random numbers.

The ANN is accepted to have 5 neurons in the hidden layer and one neuron in the output layer. The activation function of the neurons in the hidden layer is nonlinear differentiable logsig with output in the range [0, 1]:

$$\mathrm{logsig}(n) = \frac{1}{1+\exp(-n)} \quad (8)$$

The input to the lth hidden neuron (l=1÷5) is:

$$n_l = W1_{l1}.p_1 + W1_{l2}.p_2 + \ldots + W1_{lR}.p_R + b1_l \quad (9)$$

where $W1_{li}$ are the weights between the i^{th} variable (i=1÷R, R=10) of the n^{th} input data vector $\mathbf{p}_{nx10}$ and the l^{th} hidden neuron, p_i are the corresponding values of the i^{th} variable in the input data vector $\mathbf{p}_{nx10}$ and $b1_l$ is the hidden neuron bias.

The activation function of the single output neuron for pf is linear. The ANN is designed, initialized, trained and validated using MATLAB™, [31]. Its block diagram is presented in Figure 4.

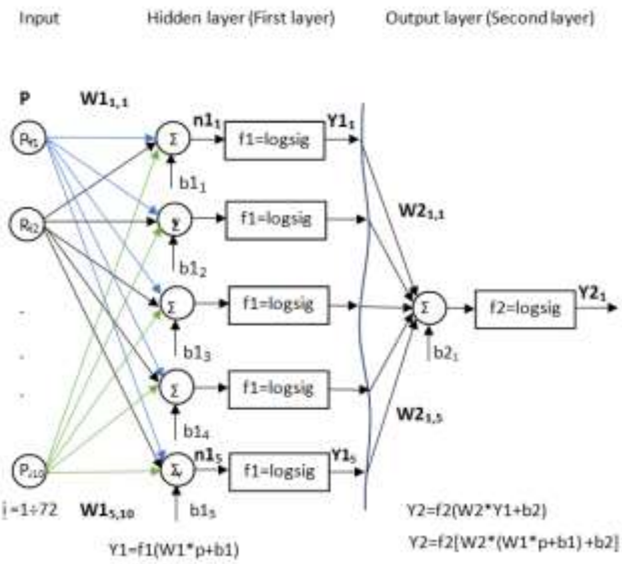

Fig. 4: Two-layer neural network with 5 neurons in the hidden layer

A general algorithm for development and validation of the ANN predictor is depicted in Figure 5.

The input data are the training and validation couples ($\mathbf{P}_{72x10}$, $\mathbf{T}_{72x1}$), ($\mathbf{P}^V_{10x10}$, $\mathbf{T}^V_{10x1}$), the number of layers and neurons in each layer, the activation functions, the training method, the cost function and the end condition.

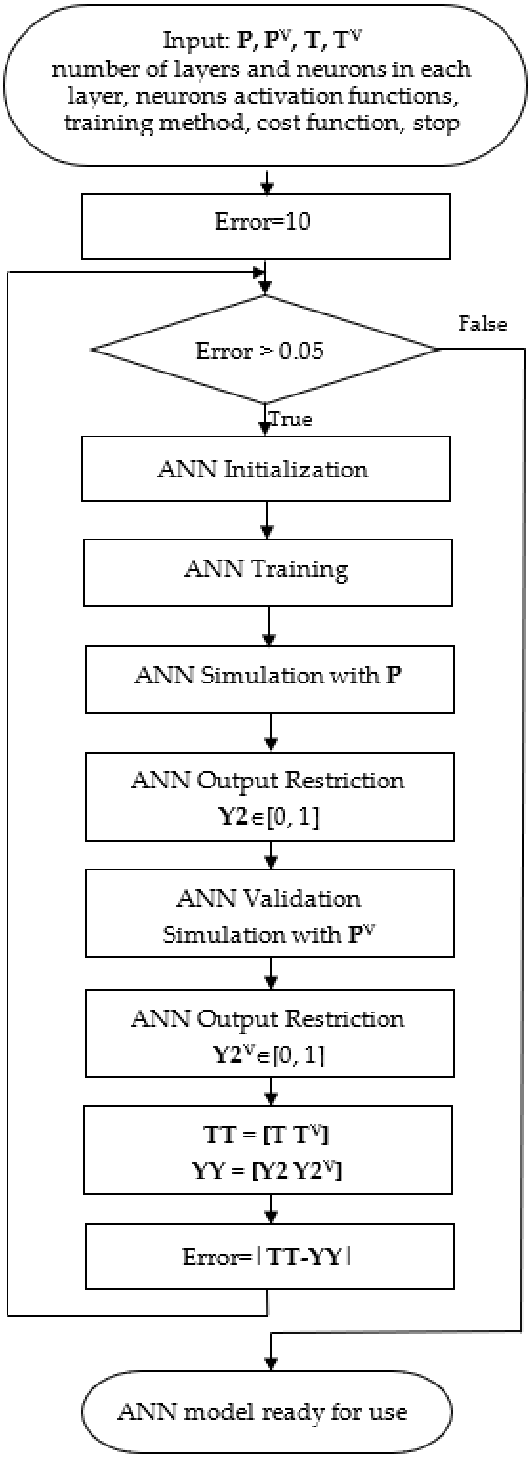

Fig. 5: General algorithm for development and validation of multi-layer ANN predictor

Based on this input data a two-layer artificial neural network with 5 neurons in the hidden layer has been initialized assigning random initial values to biases and weights. Then it is trained by BP to minimize MSE. After reaching the accepted end condition of 2000 epochs, the ANN is simulated with the input for training $\mathbf{P}_{72x10}$ to compute the output $\mathbf{Y2}$=p_f.

The linear activation function of the output neuron may cause **Y2** to exceed the probability range which requires restriction of the values **Y2** in the range [0, 1]. The predictor is validated using ANN simulation with independent input data from the matrix $\mathbf{P}^V_{10x10}$. The output of the validation process $\mathbf{Y2}^V$ is restricted to take values in the range [0, 1]. Then the two target vectors and the corresponding output vectors are united TT= [T TV], YY = [Y2 Y2V] and the absolute error computed Error=|TT-YY|.

The training and the validation are successful if Error < 0.05 and then the trained ANN predictor is ready for use. If this condition is not satisfied the process is repeated from different random initial values for the biases and the weights. If the Error persists to be high the ANN predictor is trained for new input data.

The final computed weights **W** and biases **B** of the trained ANN to serve as a forest fire predictor are presented as the following matrices and vectors respectively:

- for Layer 1 (Hidden Layer with 5 neurons) **W1** and **B1**

$$\mathbf{W1}_{10x5}=\begin{bmatrix} -0{,}221 & 0{,}959 & -1{,}005 & 2{,}583 & 3{,}441 \\ 0{,}274 & -0{,}460 & -3{,}781 & 1{,}473 & 2{,}360 \\ -3{,}565 & -0{,}519 & 1{,}556 & -1{,}154 & -2{,}971 \\ -2{,}513 & 1{,}292 & -3{,}193 & 0{,}774 & -0{,}654 \\ 3{,}274 & 4{,}029 & -2{,}799 & 3{,}560 & -0{,}374 \\ -0{,}334 & -0{,}958 & 1{,}549 & -1{,}833 & 1{,}110 \\ -3{,}347 & 0{,}070 & 10{,}693 & -2{,}189 & 0{,}927 \\ 0{,}928 & 0{,}162 & -6{,}780 & 1{,}036 & 4{,}525 \\ -1{,}704 & -5{,}613 & 3{,}726 & -4{,}480 & -6{,}975 \\ 0{,}980 & 3{,}067 & 2{,}407 & 1{,}673 & 0{,}578 \end{bmatrix} \quad (10)$$

$$\mathbf{B1}_{1x5}=[5{,}244 \quad -4{,}345 \quad -1{,}516 \quad -2{,}975 \quad 4{,}880] \quad (11)$$

- for Layer 2 **W2** (the weights between the 5 outputs from the hidden neurons and the input of the single neuron from the output layer) and **B2**

$$\mathbf{W2}=[-0{,}195 \quad -1{,}783 \quad -2{,}076 \quad 1{,}699 \quad 0{,}135] \quad (12)$$

$$\mathbf{B2}=[1{,}14]. \quad (13)$$

The target $\mathbf{T}_{72x1}$ and the obtained ANN output as a function of the record number are presented in Figure 6. The target is depicted in black dots whereas the ANN predictor output – is in yellow dots.

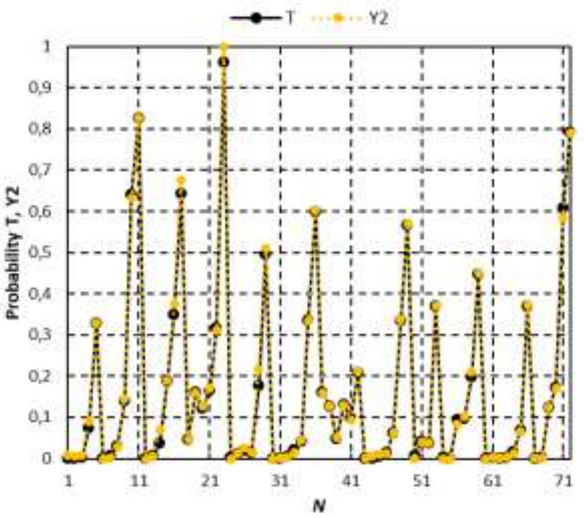

Fig. 6: Target and ANN predictor output

It can be seen that the ANN output keeps close to the target with a high degree of accuracy of less than 5%. The biggest difference is observed in records 12 and 23, where the errors remain still far below the set value of 5 %.

3 ANN Fire Predictor Validation

The validation of the designed ANN fire predictor is tested by the accuracy of prediction for input data not used in the training of the ANN. It is included as a second part of the algorithm in Figure 5 where the input validation couple is ($\mathbf{P}^V$10x10, $\mathbf{T}^V$10x1) which makes 13% of the total number of 82 records. The validation is based on 2 records with real fire occurrence (values for pf close to 1) and 8 records with no presence of fire (values for pf close to 0). The input data, the target, and the predicted output probability are shown in Table 1.

Table 1. Validation dataset with ANN output

Fire	Temp [C]	RH [%]	SolR [Wm^2]	Rain [mm]	MaxT [C]	Wind [m/s]	FM1 [%]	F100 [%]	FM1000 [%]	KBDI	T	Y2
NO	18,3	70	7	0	34,4	0,45	14,5	13,8	11	674	0,015	0,04
YES	33,9	30	894	0	34,4	1,34	8,2	13,7	10,5	674	0,666	0,685
NO	32,8	24	769	0	33,3	0,89	6,2	11,4	10,3	636	0,57	0,527
NO	15	56	21	0,02	31,7	0,89	11,6	7,8	9,6	521	0	0
NO	38,9	9	763	0	39,4	2,24	3,1	7,7	10,1	651	0,534	0,513
NO	27,2	25	557	0	31,7	0,89	8	8,7	9,3	526	0,002	0,003
YES	35,6	14	838	0	38,9	1,34	4,1	6,3	7,7	773	0,445	0,418
NO	30,6	17	370	0	33,9	1,34	4,2	8,1	8,8	756	0,34	0,035
NO	37,2	21	856	0	38,3	1,34	4,5	9,1	11,2	638	0,242	0,276
NO	35	10	433	0	37,8	0,89	3,8	7,6	10	751	0,48	0,494

In Figure 7 the black dots connected by a solid line represent the target values T^V for pf whereas the yellow dots connected by a dashed line indicate the computed output $Y2^V$ of the ANN fire predictor.

The result shows that an accuracy of 95% has been achieved in validation which has been set up as a requirement in the algorithm in Figure 5. The error is less than 5% even for the biggest difference between the corresponding dots from record 3. The accuracy illustrates how well the ANN has been trained to predict fire forest events for any new records.

The ANN training and the validation that follows after it are inseparable processes as seen from the algorithm in Figure 5. This sequence is repeated in a cycle till the absolute difference (error) between the total ANN predictor output from training and validation and the corresponding target becomes less than 0.05 (5%).

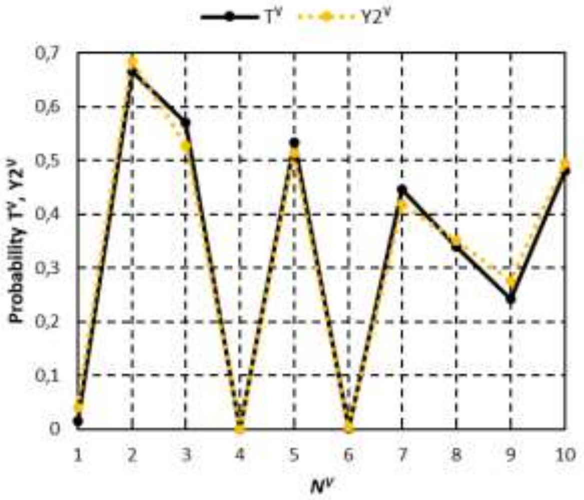

Fig. 7: Target and ANN predictor output from validation

The ANN training in each new cycle starts from different random initial ANN weights and biases, generated during ANN initialization. Therefore, each new design for the same input data in the algorithm leads to a new ANN predictor, which differs in weights and biases.

The computational time for the derivation of a validated ANN predictor is determined by the number of cycles necessary to obtain an ANN predictor with the desired accuracy and the computer and MATLAB^{TM} performance. The time for the execution of a cycle depends on the ANN architecture (number of layers and neurons), the training method, the training data - size and pre-processing (normalization, noise, and co-linearity elimination, etc.), the generated random initial values of the ANN tunable parameters (weights and the biases) which determine how far from the optimal values where MSE is minimum the training starts, the time for simulation after training and for validation.

4 Discussion and Analysis

In the present research, the complexity of the ANN predictor is reduced by using a simple architecture, a fast-converging training algorithm, and a small amount of training data which proper selection and pre-processing that ensures diversity and representativeness assists the fast convergence. This contributes to a decrease in the absolute time for the execution of the algorithm. It is in the range [5÷20] minutes for the design of several ANN fire predictors using a medium powerful computer and MATLAB™ Release 13.

However, the time for designing an ANN predictor is of small importance since ANN predictors are derived offline and once for a great period. The representative training sample used is selected and properly processed from a variety of collected data over the course of many years on end. This ensures that the ANN can accurately predict future events described by similar data used in the ANN design. A derivation of a new ANN predictor may be required when thoroughly new conditions arise.

5 Conclusions and Future Research

The novelty and the main contributions of the presented research are concluded in the following.

- A general algorithm for the development and validation of a multi-layer ANN predictor is suggested.
- Based on that algorithm a backpropagation two-layer ANN for prediction of the probability of fire occurrence is designed using MATLAB™. The ANN predictor accounts for the most contributing factors for the wildfire occurrence, i.e., the weather and the fuel moisture condition.
- The trained ANN predicts wildfire events with higher accuracy (error <5%) than the regression prediction model from [28]. Besides, it uses only a selected limited input data sample of 82 records from 10 years in the region of Whitmore, North California, USA. The comprehensive and low-complexity algorithm ensures fast processing. To this, contribute the ANN simple architecture, the fast-converging training method, and the small amount of properly pre-processed training data.
- An advantage of the ANN predictor is its simple structure with 5 hidden neurons. The ANN predictor enables the analytical description of the nonlinear relationship between a great number of significant variables and the fire probability. The results from this research may help in improving the available forest fire early prediction and prevention systems worldwide. Furthermore, the general methodology suggested can be applied to different regions and input data.
- To the authors' knowledge the present investigation is the first approach to develop an ANN for predicting the probability of forest fire occurrence from meteorological data and fuel moisture conditions for the region of Whitmore, California.

The future research will focus on the integration of the developed ANN model in the existing fire alarm system in the region of Whitmore, Northern California, and comparison with the real data. Besides, the developed algorithm can be also applied to other hazard areas.

Abbreviations:

SVM - support vector machine
BP - backpropagation of error algorithm
ANN - artificial neural network
FL - fuzzy logic
ANFIS - adaptive neuro-fuzzy inference system
KBDI - Keetch Byram drought index
MSE - mean square error
Logsig - logarithmic sigmoidal
DC - drought code
LSTM - long- and short-term memory network
ReLU - rectified linear activation function
NFDRS - national fire danger rating system
R- CNN -region-based convolutional neural network

Designations

FM - fuel moisture
Temp - current temperature
RH - relative humidity
SolR - solar radiation for the day
Rain - precipitation amount
maxT - maximum temperature during the day
Wind - average wind speed within 10 min period
N - number of records
Pf - fire probability
P, T - training input-target matrices
"V" - superscript for validation data

Acknowledgments:
The authors would like to thank the Research and Development Sector at the Technical University of Sofia for the financial support.

References:

[1] Report of the United Nations Environment Assembly of the United Nations Environment Programme, *United Nations*, 2022, ISSN: 0252-2055, [Online]. https://www.unep.org/ (Accessed Date: February 17, 2024).

[2] Chettouh Samia, Hamzi Rachida, Statistical Fire Models : Review. *International journal of Engineering research & Technology (IJERT)*, December 2013, Vol. 2, Issue 12.

[3] Ma S, Liu Q, Zhang Y. A prediction method of fire frequency: Based on the optimization of SARIMA model. *PLoS One*. 2021 Aug e0255857 https://doi.org/10.1371/journal.pone.0255857.

[4] Al-Janabi, Samaher & AlShourbaji, Ibrahim & Salman, Mahdi Abed. Assessing the Suitability of Soft Computing Approaches for Forest Fires Prediction. *Applied Computing and Informatics* 2018, *14*. 214-224. https://doi.org/10.1016/j.aci.2017.09.006.

[5] M. Zahari, R. R. Karri, M, Isa, El-Said Zahran, S. M. Nagendra, "Soft computing techniques for prediction of forest fire occurrence in Brunei Darussalam". *AIP Conference Proceedings*, 2643(1): 030023, 2023, https://doi.org/10.1063/5.0110349.

[6] Nebot, À.; Mugica, F. Forest Fire Forecasting Using Fuzzy Logic Models. *Forests* 2021, *12*, 1005. https://doi.org/10.3390/f12081005.

[7] M. Bisquert, E. Caselles, J. Sanchez, "Application of artificial neural networks and logistic regression to the prediction of forest fire danger in Galicia using MODIS data". *International Journal of Wildland Fire*, 2012, Vol. 21, pp. 1025–1029, http://dx.doi.org/10.1071/WF11105.

[8] K. R. Singh, K. P. Neethu, K Madhurekaa, A Harita, P. Mohan, "Parallel SVM model for forest fire prediction". *Soft Computing Letters*, *Vol. 3*, 2021, https://doi.org/10.1016/j.socl.2021.100014.

[9] J. O. Otieno, "*A machine learning algorithm for predicting wild fire occurrence*". Thesis, Strathmore University, 2020, [Online]. http://hdl.handle.net/11071/12085 (Accessed Date: February 17, 2024).

[10] G. Zhang, M. Wang, K. Liu, "Forest Fire Susceptibility Modeling Using a Convolutional Neural Network for Yunnan Province of China". *International Journal of Disaster Risk Science*, , 2019, *Vol. 10*, pp. 386–403, https://doi.org/10.1007/s13753-019-00233-1.

[11] A. Bouteska, P. Hajek, B. Fisher, M. Z. Abedin, "Nonlinearity in forecasting energy commodity prices: Evidence from a focused time-delayed neural network". *Research in International Business and Finance*, 2023, Vol. 64, https://doi.org/10.1016/j.ribaf.2022.101863.

[12] S. Fathi, M. Mehravar, M. Rahman, "Development of FWD based hybrid back-analysis technique for railway track condition assessment". *Transportation Geotechnics*, 2023, *Vol. 38*, https://doi.org/10.1016/j.trgeo.2022.100894.

[13] D. Tsintikidis, J. L. Haferman, E. N. Anagnostou, W. F. Krajewski, T. F. Smith, "Neural network approach to estimating rainfall from spaceborne microwave data". *IEEE Transactions on Geoscience and Remote Sensing*, 1997, *vol. 35*, no. 5, pp. 1079-1093,https://doi.org/10.1109/36.628775.

[14] R. Teschl, W. Randeu, F. Teschl, "Improving weather radar estimates of rainfall using feed-forward neural networks". *Neural networks: the official journal of the International Neural Network Society*, 2007, *Vol. 20*, Issue 4, pp. 519-27, https://doi.org/10.1016/j.neunet.2007.04.005.

[15] J. Makwana, M. Tiwari, B. Deora, "Development and comparison of artificial intelligence models for estimating daily reference evapotranspiration from limited input variables". *Smart Agricultural Technology*, 2023, Vol. 3, https://doi.org/10.1016/j.atech.2022.100115.

[16] Safi, Youssef & Bouroumi, Abdelaziz. Prediction of forest fires using Artificial neural networks. *Applied Mathematical Sciences* 2013. *7*. 271-286, http://dx.doi.org/10.12988/ams.2013.13025.

[17] Zechuan Wu, Z. Wu, Bin Wang, Mingze Li, M. Li, Yuping Tian, Y. Tian, Ying Quan & J. Liu. Simulation of forest fire spread based on artificial intelligence. *Ecological indicators*, 2022, *136*, 108653, https://doi.org/10.1016/j.ecolind.2022.108653.

[18] Omar, Naaman & Al-zebari, Adel & Sengur, Abdulkadir. Deep Learning Approach to Predict Forest Fires Using Meteorological Measurements. 2021. 1-4, http://dx.doi.org/10.1109/IISEC54230.2021.9672446.

[19] Sakr, G. & E., Imad & Mitri, G. Efficient forest fire occurrence prediction for developing countri.es using two weather parameters. *Eng. Appl. of AI*. 2011, *24*. 888-

894, https://doi.org/10.1016/j.engappai.2011.02.017.

[20] Kantarcioglu, Omer & Schindler, K. & Kocaman, Sultan. Forest fire susceptibility assessment with machine learning methods in north-east turkiye. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2023, XLVIII-M-1-. 161-167, https://doi.org/10.5194/isprs-archives-XLVIII-M-1-2023-161-2023.

[21] M.Kumar , Sowmya B., Priyanka S. , Ruchita Sharma , Shivank T. Karani, Forest Fire Prediction Using Image Processing And Machine Learning, 2021, Department of Computer Science and Engineering , Ramaiah Institute of Technology, MSRIT, Karnataka, India 560054 *Nat. Volatiles & Essent*. Oils; 8(4): 13116-13134.

[22] Nikova H., Deliyski R., Tashev T., Comparison analysis of Wildfire early prediction models", *XIX International Conference on Challenges in Higher Education and research in the 21th century*, 2023, vol.19, ISSN: 2683-0337.

[23] National Oceanic and Atmospheric Administration, JetStream Max: Addition Köppen-Geiger Climate Subdivisions, [Online]. https://www.noaa.gov/jetstream/global/climate-zones/jetstream-max-addition-k-ppen-geiger-climate-subdivisions (Accessed Date: July 14, 2023).

[24] National Fire Danger Rating System, U.S. Department of Agriculture, Forest Service, [Online]. https://www.fs.usda.gov/detail/cibola/landmanagement/resourcemanagement/?cid=stelprdb5368839 (Accessed Date: February 17, 2023).

[25] P. Schlobohm, J. Brain, Gaining and Understanding of the National Fire Danger Rating System, National WildfireCoordinating Group - Fire Danger Working Team, 2002, [Online]. https://www.nwcg.gov/sites/default/files/publications/pms932.pdf (Accessed Date: February 17, 2023).

[26] J. E. Deeming, R. E. Burgan, J. D. Cohen, "The National Fire-Danger Rating System—1978 (General Technical Report INT-39)", U.S. Department of agriculture, Forest Service: Ugden, Utah, USA, 1977, [Online]. https://www.fs.usda.gov/rm/pubs_series/int/gtr/int_gtr039.pdf (Accessed Date: February 17, 2023).

[27] NWCG National wildfire coordination group. NDFDRS Fuel models, [Online]. https://www.nwcg.gov/publications/pms437/fire-danger/nfdrs-station-catalog (Accessed Date: July 14, 2023).

[28] Nikova H., Deliyski R., "Binary regression model for automated wildfire early prediction and prevention", *11th International Scientific Conference on Computer Science (COMSCI)*, Sozopol, Bulgaria, 2023, http://dx.doi.org/10.1109/COMSCI59259.2023.10315856.

[29] Brian C. Cronk, *How to Use SPSS® A Step-By-Step Guide to Analysis and Interpretation*, Copyright 2020, Published October 11, 2019 by Routledge.

[30] Ch. Sekhar, P. Meghana, "A Study on Backpropagation in Artificial Neural Networks", *Asia-Pacific Journal of Neural Networks and Its Applications*, 2020, *Vol. 4*, pp. 21-28, http://dx.doi.org/10.21742/AJNNIA.2020.4.1.03.

[31] Demuth H., M. Beale, *Neural Network Toolbox for Use with MATLAB*. Users Guide, The Mathworks Inc., 2018a, [Online]. http://cda.psych.uiuc.edu/matlab_pdf/nnet.pdf (Accessed Date: February 17, 2023).

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

Conceptualization, methodology, and supervision, S.Y and R.D.; software, H. N. and S.Y; validation, investigation resources, writing—review and editing and visualization, H.N.; writing—original draft preparation, H.N., S.Y. and R.D; funding acquisition, R.D.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

The authors would like to thank the Research and Development Sector at the Technical University of Sofia for the financial support.

Conflicts of Interest

The authors declare no conflict of interest.