Comparing Artificial Neural Networks with Multiple Linear Regression for Forecasting Heavy Metal Content

Rachid El Chaal[*] , Moulay Othman Aboutafail

ENSA of Kenitra, Engineering Sciences Laboratory, Data Analysis, Mathematical Modeling and Optimization Team, Ibn Tofail University, 14000 Kenitra, Morocco

[*] Correspondence: Rachid El Chaal (rachid.elchaal@uit.ac.ma)

Abstract: This paper adopts two modeling tools, namely, multiple linear regression (MLR) and artificial neural networks (ANNs), to predict the concentrations of heavy metals (zinc, boron, and manganese) in surface waters of the Oued Inaouen watershed flowing towards Inaouen, using a set of physical-chemical parameters. XLStat was employed to perform multiple linear and nonlinear regressions, and Statista 10 was chosen to construct neural networks for modeling and prediction. The effectiveness of the ANN- and MLR-based stochastic models was assessed by the determination coefficient (R^2), the sum squared error (SSE) and a review of fit graphs. The results demonstrate the value of ANNs for prediction modeling. Drawing on supervised learning and back propagation, the ANN-based prediction models adopt an architecture of [18-15-1] for zinc, [18-11-1] for manganese, and [18-8-1] for boron, and perform effectively with a single cached layer. It was found that the MLR-based prediction models are substantially less accurate than those based on the ANNs. In addition, the physical-chemical parameters being investigated are nonlinearly correlated with the levels of heavy metals in the surface waters of the Oued Inaouen watershed flowing towards Inaouen.

Keywords: Artificial neural networks (ANN); Mathematical modelling; Multiple linear regression (MLR); Heavy metal; Residual; Prediction

1. Introduction

Water can be polluted by many things, namely, wastewater, industrial products, and domestic products. The discharge of pollutants, particularly those containing harmful nanoparticles or metallic heavy metals, makes the water toxic and poses risks to human health, biological life, and environmental security. Metals have a special status in water pollution, because they are employed frequently in many applications. The heavy metal elements are brought up the food chain by water. These elements are often only present in tiny amounts, but when they bioaccumulate in living organisms, they pose an increasingly dangerous problem [1].

Public agencies have set limitations to manage emissions, examine threats, and protect our ecosystem against the toxicity of heavy metals. The scientific community and Moroccan government officials are working hard on environmental issues [2]. Morocco's environmental plans introduce the management and management of natural resources, and consider the preservation of our ecosystem and the preservation of natural resources [3].

This paper aims to create mathematical models that can predict the presence of heavy metals like manganese, boron, and zinc based on various environmental factors. The objects are the surface waters in the Oued Inaouen watershed. Two prediction models, namely, artificial neural network (ANN) and multiple linear regression (MLR) were compared on Statista 10 and XLSTAT 2017.

2. Methodology

2.1 Database

Our database consists of 100 surface water samples (observations) from the province of Taza. The collection,

transportation, and storage of water samples in 2014 and 2015 were accomplished in accordance with the policies and practices of the National Drinking Water Office. The lab run by the University Sidi Mohamed Ben Abdellah (USMBA) of Fez, known as the Regional University Interface Center (CURI), executed part analysis, part fabrication, and part manufacturing [4-6].

2.2 Input Selection

A total of 16 physical-chemical features are identified in these samples, and taken as independent (explanatory) factors, namely, total dissolved solids (TDS), conductivity (Cond), temperature (T/°C), dissolved oxygen (DO), pH, potassium (K), magnesium (Mg), calcium (Ca), nitrate (NO_3), chloride (Cl), phosphorus (P), ammoniacal nitrogen (NH_4), total alkalinity ($CaCO_3$), sodium (Na), and sulfates (SO_4) [5]. Manganese (Mn), zinc (Zn), and boron(B) levels were identified as the three dependent variables to be predicted.

From the overall database, 70% of the data were selected randomly to train the dependent variable forecast model [7]. The remaining 30% of data were used to prevent over fitting, and assess the prediction effect of the trained model.

2.3 Data Preprocessing

Data normalization makes models less difficult. The input data are made up of the 16 independent variables above, which are of various magnitudes. The data are transformed into a standardized variable to equalize the measuring scales. The I values of each independent variable are adjusted to match their means and standard deviations, using the connection $X(v_i)$:

$$X_s(v_i) = \left(\frac{X(v_i) - \bar{X}(v_i)}{\sigma(v_i)}\right), i \in \{1\ldots\ldots16\}$$ (1)

where, $X_s(v_i)$, $X(v_i)$ and $\bar{X}(v_i)$ are the standardized, measured, and mean values, respectively. The mean value can be calculated by:

$$\bar{X}(v_i) = \frac{1}{100} \sum_{k=1}^{100} X_k(v_i)$$ (2)

The standard deviation $\sigma(v_i)$ can be expressed as:

$$\sigma(v_i) = \sqrt{\frac{1}{100} \sum_{k=1}^{100} (X_k(v_i) - \bar{X}(v_i))^2}$$ (3)

The values of all variables are standardized to prevent computing excessively large or small exponents, and thus reduce the variability of the mean [8]. The data of the explanatory variable was also normalized to the range of [0,1] to meet the constraints of the transfer function used by the neural networks. The normalization can be realized by the following relationship [9]:

$$Y_n = \left(\frac{Y - Y_{min}}{Y_{max} - Y_{min}}\right)$$ (4)

where, Y_n, Y, Y_{min} and Y_{max} are the standardized, measured, minimum, and maximum values, respectively.

2.4 Data Analysis

Two methods are adopted to realize the modelling process, namely, MLR and MLP-type ANN, aiming to optimize the link between manganese (Mn), zinc (Zn), and boron (B), and physicochemical elements [10].

2.4.1 MLR
Regression analysis is a statistical technique that enables the examination of the potential relationships between a dependent variable and one or more independent variables [11].

Linear regression uses a straight line to summarize, interpret, and forecast the fluctuations of a dependent variable (Y) according to another independent variable (X). Regardless of the objective, the success of regression

analysis is substantially influenced by the relationships between the explanatory variables [12].

A statistical technique known as MLR describes the fluctuations of an endogenous variable connected to numerous exogenous variables [13]. It addresses the same problem as the simple linear regression (SLR). The difference is that MLR seeks to explain the values of Y not by a single variable X, but by several variables $\{X_j\}$ known as explanatory variables. After slightly modifying the previous notations, it is assumed that Y and $\{X_j\}$ has a linear relationship [13]:

$$Y = a_0 + a_1 X_1 + a_2 X_2 + \cdots + a_p X_p + \varepsilon \tag{5}$$

where, Y is the dependent variable; $X_1, X_2, \dots, X_p$ are the independent variables; p is the number of explanatory variables; $a_0, a_1, a_2, \dots, a_p$ are the parameters to be estimated; a_0 is the estimated intercept; $a_1, a_2, \dots, a_p$ are the slopes (partial regression coefficients); ε is the error of the model that expresses or summarizes the missing information in the linear explanation of Y values from $X_1, X_2, \dots, X_p$ [14].

The MLR coefficients can be interpreted the same as the SLR coefficients: The intercept represents the expected value of Y when the values of all the X_j are fixed at 0. Each slope a_j indicates the expected change in Y for a change of one unit in X_j, while all other things are equal. The last rule is important, for it requires all other X_j to have fixed values [15].

2.4.2 MLP-type ANN

The ANN is a computational model mimicking biological behaviors of the human nervous system [16]. The human brain is composed of a vast number of neurons, which correspond to the nodes, the fundamental elements of an ANN. The ANN is an excellent tool for data analysis [17].

Each ANN contains many nodes that communicate with each other by sending signals through links called synaptic connections. In general, there are three kinds of nodes in the neural system [18]: Input nodes that receive data, output nodes that send data as the system output; hidden nodes whose input and output signals remain in the system.

(a) Formal node

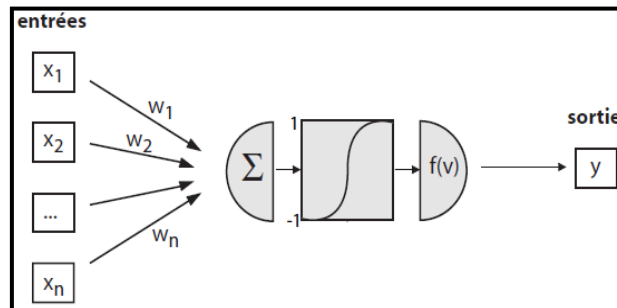In statistics, a formal node is a parameterized algebraic function with nonlinear parameters [19] and bounded values:

$$Y = f(X_1, X_2, \dots, X_n; W_1, W_2, \dots, W_n) \tag{6}$$

where, X_i is the model input; w_i is the model parameter; f is the activation function of the node [20].
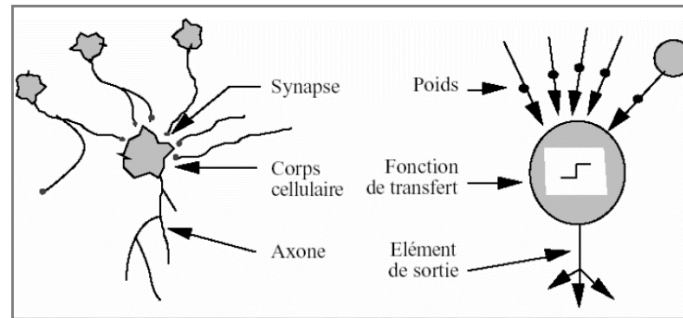
Figure 1. Mapping biological node to artificial node

Figure 2. Schematic of a formal node

Figure 1 compares the structure of an artificial node with that of a human neuron. The artificial node (an elementary processor) receives some inputs from upstream nodes. Each input is associated with a weight W, which represents the strength of the connection, and an output function, whose output is imported to other similar nodes [21].

The formal node is the essential element of a neural network. It is a straightforward mathematical operator whose numerical value can easily be calculated. A typical node has N inputs, each with a synaptic weight (Figure 2), an output function, an output which in turn serves as an input to other similar nodes [22]. The result is a nonlinear function f of a linear combination of the inputs (X_i). The most frequently used potential v is the weighted sum of the inputs X_i weighted by the coefficients (w_i) also called connection weights [23].

Then, a formal node can be characterized as follows [24]: The input signals are accepted by a set of synaptic connections, which are defined by synaptic weights determining the effect of the signal by node i on node j. In addition, a combination function or adder performs the weighted activation sum that converges to node j. This weighted sum can be expressed as:

$$A_j = \sum_i W_{ij} X_i \tag{7}$$

where, W_{ij} is the synaptic weight; X_i is the input value. An activation (or transfer) function drives a node by determining its activation as the node output [25]:

$$\sum W_{ij} X_i + \theta_j \tag{8}$$

On this basis, a transfer function φ computes the value of the node state, called activation X_j. It is then transmitted to the forward nodes:

$$Y_j = \varphi \left(\sum W_{ij} X_i + \theta_j \right) \tag{9}$$

where, θ_j is the bias of node j based on a kind of local weight, i.e., an inhibitory input employed in many types of activation functions. The input makes the network more flexible by varying the node's trigger threshold through weight adjustment via training. This is adopted in several types of activation functions [26].

(b) Activation function

The activation function non-linearizes the functioning of the node. It generally has three intervals [27]: The node is inactive, when the value of the function is below the threshold (the output is often 0 or -1); the node is in the transitional state, when the value of the function is around the threshold; the node is active, when the value of the function is above the threshold (the output is often 1).

There are several activation functions, including hyperbolic tangent, Gaussian, sigmoid, affine function, and arc tangent function. The most popular one is the sigmoid function (Figure 3) [28]:

$$f(x) = \frac{1}{1 + e^{-x}} \tag{10}$$
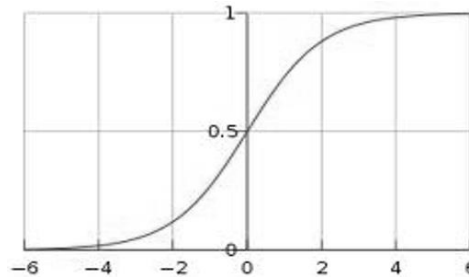
Figure 3. Graphical representation of the sigmoid function

3. Results and Discussion

The Results section may be divided into subsections. It should describe the results concisely and precisely, provide their interpretation, and draw possible conclusions from the results. The results from the STATISTICA neural network method are displayed in Table 1. Using the Broyden-Fletcher-Goldfarb-Shanno algorithm [29],

According to the topology, they provide the two layers' determination factors, cycles, and training algorithm (BFGS).

By changing the total number of hidden nodes and the number of training cycles, we changed the network's design (number of iterations). We gradually changed the number of hidden nodes to achieve this (1, ..., 15). The findings demonstrated that the optimum prediction model for heavy metal concentrations is obtained when the number of hidden nodes equals eleven for Mn, eight for B, and fifteen for Zn. At this moment, the lowest mean square error, the most iterations, and the highest determination coefficient have been achieved.

The network's training in the Boron scenario is shown in Figure 4. After 102 steps, the outcome is reached. The learning error and test gradients converge properly with eight nodes in the caching layer.

Table 1. The balance of the best architectures obtained by Statistica for the three metals

	Architecture	Training algorithm	Error function	Activation function of the hidden layer	Output layer activation function	Number of iterations
Zn	16-15-1	BFGS	SOS	Tanh	Logistic	217
B	16-8-1	BFGS	SOS	Tanh	Exponential	102
Mn	16-11-1	BFGS	SOS	Logistic	Exponential	228
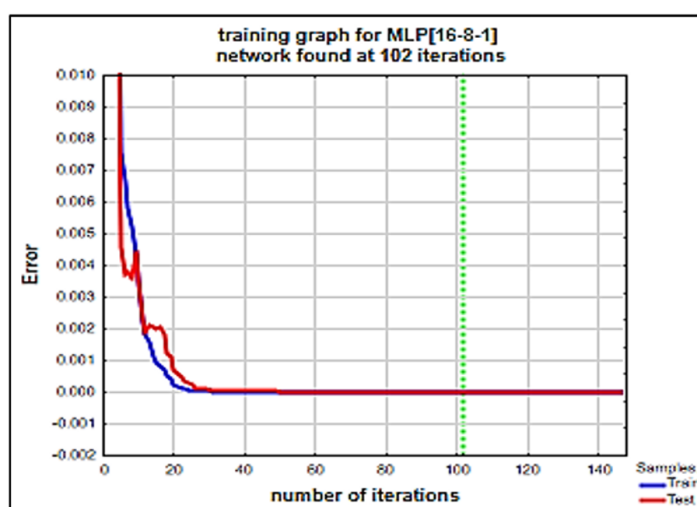
Figure 4. Eight nodes in the cache layer and the development of the RMSE in the case of Boron

Table 2. Changes in the root mean square sum of errors (RMSE), the iteration number (IN), and the determination coefficient R^2 for the elements boron, manganese, and zinc as a function of the number of hidden layer nodes (NHLN)

| Metals	NHLN=8			NHLN=11			NHLN=15		
	R^2	IN	RMSE	R^2	IN	RMSE	R^2	IN	RSME
B	0.99	102	0.0001	0.98	199	2.07	0.95	81	5.96
Mn	0.93	75	4.25	0.99	228	0.003	0.96	123	1.13
Zn	0.97	58	0.98	0.94	12	2,98	0.99	217	0,001

Table 3. Between the heavy metal concentrations observed and those anticipated by the MRL, a coefficient of determination ANN

|		Boron	Manganese	Zinc
MLR		$R^2 = 0.17$	$R^2 = 0,22$	$R^2 = 0,4$
ANN		Architectures of the selected models		
		[16-8 -1]	[16-11 -1]	[16-15 -1]
		Learning		
		$R^2 = 0,997$ SSE = 0,0001	$R^2 = 0,998$ SSE = 0,003	$R^2 = 0,9998$ SSE=0,001
		Test		
		$R^2 = 0,996$ SSE = 0,008	$R^2 = 0,997$ SSE = 4.291	$R^2 = 0,997$ SSE = 3.902

Overlearning has occurred due to the network's extensive training; this occurrence has occurred 102 times.

Comparing neural network models to those discovered using MLR, increases in the explanation of variance of up to 82% are possible. For manganese, boron, and zinc, it is 77%, 82%, and 59% respectively. The coefficients of determination rise from 0.224 to 0.99, 0.172 to 0.99, and 0.40 to 0.99, respectively. The outcomes shown in Table 2 demonstrate that the models created by the ANN are superior to those discovered by the MLR approach. In fact, the models created by the ANN have estimated coefficients of determination that are substantially higher (above 0.9), while the models created by the MLR have calculated coefficients that are lower (between 0.17 and 0.40). Referring to in subgraphs (a) and (b) of Figure 5 and the subsequent Table 2 and Table 3 makes this clear.

It may be deduced from the determination coefficients found when analyzing the accuracy of the ANN models that they closely match learning-related models. On the other hand, there are some differences between the determination coefficients found during training and those used to assess the models' applicability to MLR. This reveals that there are nonlinear correlations between the Physico-chemical characteristics of the environment and the concentrations of heavy metals in the waters of the Oued Inaouen catchment area. What typically manifests itself in an aquatic environment [30, 31].

(a)   (b)
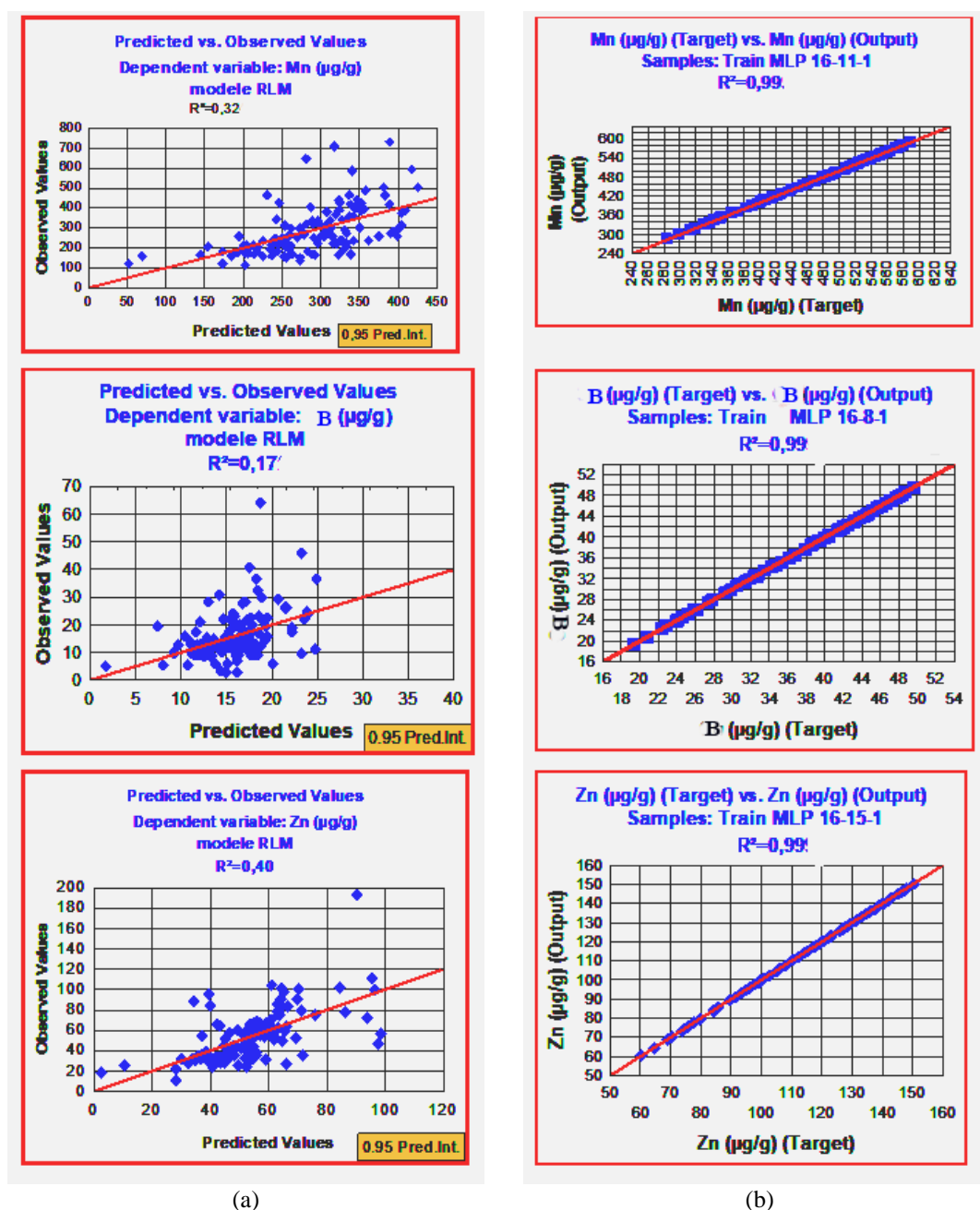
Figure 5. (a) MLR (MLR) models' estimates of boron, manganese, and zinc concentrations and the values actually measured. (b) ANN study relating the predicted amounts of boron, manganese, and zinc (ANN)
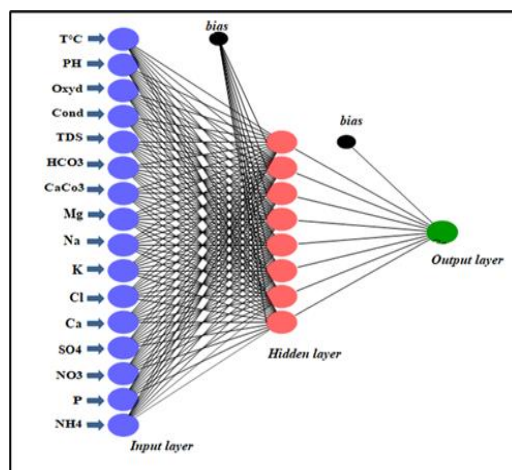
Figure 6. Design of the three-layer configuration neural network [16-8-1] employed in this research

(a)   (b)
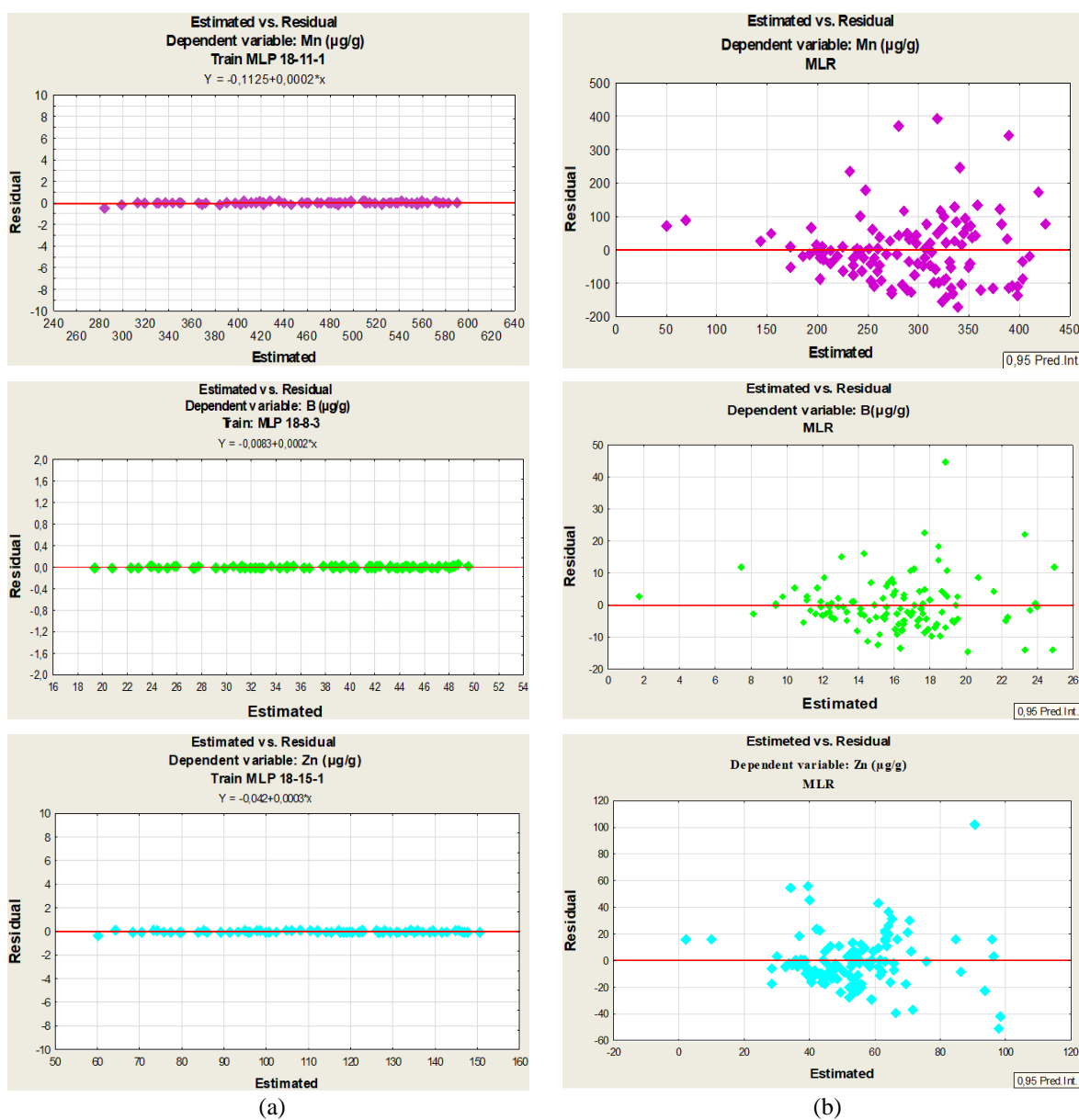
Figure 7. Relationship between the estimated values and their residuals from the NR and relative MLR model

The ANN architecture we employed for Boron prediction is shown in Figure 6.

A topology network [16-8-1] is what it is.

The connections between measured and estimated amounts of heavy metals for models created using the ANN and MLR approaches are shown in subgraphs (a) and (b) of Figure 5. The effectiveness of the ANN technique is established by the other two metals and manganese, which both demonstrate a high coefficient of determination of 0.99. (Boron and Zinc). We can show that the models discovered by the ANN are more effective by contrasting them with those produced by the MLR. We are looking at in subgraphs (a) and (b) of Figure 5, it is simple to observe this.

4. Residual Studies

The relative error is the error made by the models created using each method on a specific member of the sample used to build the model [32].

For starters, it enables us to spot any outliers or observations crucial in deciding the regression. This examines the residuals (Yi-Yi') essential [33]. Second, the only way to empirically test the accuracy of the model's assumptions is frequently reviewed these residuals.

The relationships between the estimated heavy metal values from the neural network and MLR models and their residuals are depicted in the following figures. These graphs demonstrate that for the four metals under investigation, the residuals generated by neural network approach models are less dispersed [34] (near zero) than those obtained by MLR models. This demonstrates once more that the current method, which is based on the principle of ANNs, is more effective than the method based on MLR, which is typically used in the development of linear prediction models for predicting the heavy metal contents of the surface water from Physico-chemical parameters.

The relationships between the predicted heavy metal levels using the neural network (NN) and MLR (MLR) models and their residuals are shown in Figure 7.

The analysis of the residuals, which represent the differences between the estimated and observed values, shows that the models developed using the method based on the neural network principle are more efficient than those generated using MLR for the three metals studied. This proof shows a nonlinear relationship between the surface water heavy metal levels and the examined Physico-chemical parameters.

5. Conclusions

This study used MLR and ANNs to estimate the concentrations of heavy metals (manganese, boron, and zinc) in the surface waters of the Inaouen watershed based on their Physico-chemical parameters.

There are benefits and drawbacks to these statistical methods. Although more effective and efficient, neural networks do not clearly explain the relationship between the input and output data. Because there is no linear relationship between the independent and dependent variables, MLR allows for the determination of equations that describe the relationship between the explanatory input variables and the output variables.

The calculations of the mean square errors performed in this study revealed that a three-layer architecture with an input layer, a hidden layer, and an output layer is preferable for the establishment of prediction models using ANNs related to the metals studied in the surface waters of the Inaouen catchment area. They also demonstrated that, when employing a backpropagation algorithm and supervised learning, a hidden layer with 15 nodes for zinc, 11 for manganese, and 8 for boron produces mathematical models that are more accurate at predicting the future than those that use fewer nodes.

The three heavy metals found in the surface waters of the Inaouen watershed have shown significant learning and predictive capabilities thanks to ANNs.

To put things in perspective, the following ideas can help us develop this topic further:
• Work with different network types, like recurrent networks;
• Utilize other activation methods;
• To predict the contents of these metals in the surface waters of the Moroccan watershed, try extrapolating the effective models developed by ANNs related to the prediction of metal contents studied in the surface waters of the Inaouen watershed.
• The outcomes encourage further consideration of the approach that enhances the work done thus far. To test another RNA architecture to see which one yields the best results, or to investigate the possibility of simultaneously predicting the three parameters, for instance, would be very interesting.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] S. S. Sonone, S. Jadhav, M. S. Sankhla, and R. Kumar, "Water contamination by heavy metals and their toxic effect on aquaculture and human health through food Chain," *Lett. Appl. NanoBioScience*, vol. 10, no. 2, pp. 2148-2166, 2020. https://doi.org/10.33263/LIANBS102.21482166.

[2] Q. H. Vuong, "From children's literature to sustainability science, and young scientists for a more sustainable Earth," *J. Sustain. Educ.*, vol. 24, 2020.

[3] H. A. Kacem, Y. Bouroubi, Y. Khomalli, S. Elyaagoubi, and M. Maanan, "The economic benefit of coastal blue carbon stocks in a Moroccan Lagoon ecosystem: A case study at Moulay Bousselham Lagoon," *Wetlands*, vol. 42, no. 2, pp. 1-15, 2022. https://doi.org/10.1007/s13157-022-01533-x.

[4] R. El Chaal and M. O. Aboutafail, "Statistical modelling by topological maps of kohonen for classification of the physicochemical quality of surface waters of the inaouen watershed under matlab," *J. Niger. Soc. Phys. Sci.*, vol. 4, no. 2, pp. 223-230, 2022. https://doi.org/10.46481/jnsps.2022.608.

[5] R. El Chaal and M. O. Aboutafail, "Development of stochastic mathematical models for the prediction of heavy metal content in surface waters using artificial neural network and multiple linear regression," E3S Web of Conferences, vol. 314, Article ID: 02001, 2021. https://doi.org/10.1051/e3sconf/202131402001.

[6] R. El Chaal and M. O. Aboutafail, "A comparative study of back-propagation algorithms: Levenberg-Marquart and BFGS for the formation of multilayer neural networks for estimation of fluoride," *Commun. Math. Biol. Neurosci.*, vol. 2022, pp. 1-17, 2022. https://doi.org/10.28919/cmbn/7355.

[7] A. Gheziel, S. Hanini, and B. Mohamedi, "Artificial neural network (ANN) for prediction indoor airborne particle concentration," *Int. J. Vent.*, vol. 21, no. 1, pp. 74-87, 2022. https://doi.org/10.1080/14733315.2021.1876408.

[8] R. Magallanes-Quintanar, C. E. Galván-Tejada, J. I. Galván-Tejada, S. J. Méndez-Gallegos, A. García-Domínguez, and H. Gamboa-Rosales, "Narx neural networks models for prediction of standardized precipitation index in central Mexico," Atmosphere-Basel, vol. 13, no. 8, pp. 1254-1254, 2022. https://doi.org/10.3390/atmos13081254.

[9] J. Lause, P. Berens, and D. Kobak, "Analytic Pearson residuals for normalization of single-cell RNA-seq UMI data," Genome Biol., vol. 22, no. 1, pp. 1-20, 2021. https://doi.org/10.1186/s13059-021-02451-7.

[10] A. E. H. H. Alayat, H. El Badaoui, A. Abdallaoui, and D. Abrid, "Development of mathematical models for predicting the iron concentrations of lake oubeira waters (ne Algerian)," J. Fundam. Appl. Sci., vol. 10, no. 1, pp. 83-96, 2018. http://dx.doi.org/10.4314/jfas.v10i1.6.

[11] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Comput. Sci.*, vol. 7, Article ID: e623, 2021. https://doi.org/10.7717/peerj-cs.623.

[12] P. Roback and J. Legler, Beyond multiple linear regression: applied generalized linear models and multilevel models in R, New York, NY, USA: Chapman and Hall/CRC, 2021. https://doi.org/10.1201/9780429066665.

[13] G. Ciulla and A. D'Amico, "Building energy performance forecasting: A multiple linear regression approach," *Appl. Energy*, vol. 253, Article ID: 113500, 2019. https://doi.org/10.1016/j.apenergy.2019.113500.

[14] N. A. F. Sulaiman, S. M. Shaharudin, S. Ismail, N. H. Zainuddin, M. L. Tan, and Y. Abd Jalil, "Predictive modelling of statistical downscaling based on hybrid machine learning model for daily rainfall in east-coast peninsular malaysia," *Symmetry-Basel*, vol. 14, no. 5, Article ID: 927, 2022. https://doi.org/10.3390/sym14050927.

[15] L. Plonsky and H. Ghanbar, "Multiple regression in L2 research: A methodological synthesis and guide to interpreting R2 values," *Mod. Lang. J.*, vol. 102, no. 4, pp. 713-731, 2018. https://doi.org/10.1111/modl.12509.

[16] T. Dalgaty, J. P. Miller, E. Vianello, and J. Casas, "Bio-inspired architectures substantially reduce the memory requirements of neural network models," *Front. Neurosci.*, vol. 15, Article ID: 612359, 2021. https://doi.org/10.3389/fnins.2021.612359.

[17] K. Ostad-Ali-Askari and M. Shayan, "Subsurface drain spacing in the unsteady conditions by HYDRUS-3D and artificial neural networks," *Arab. J. Geosci.*, vol. 14, no. 18, pp. 1-14, 2021. https://doi.org/10.1007/s12517-021-08336-0.

[18] B. Zhang and Y. Yao, "Precipitable water vapor fusion based on a generalized regression neural network," *J. Geod.*, vol. 95, no. 3, pp. 1-14, 2021. https://doi.org/10.1007/s00190-021-01482-z.

[19] S. Scardapane and D. Wang, "Randomness in neural networks: An overview," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 7, no. 2, Article ID: e1200, 2017. https://doi.org/10.1002/widm.1200.

[20] S. Eger, P. Youssef, and I. Gurevych, "Is it time to swish? Comparing deep learning activation functions across NLP tasks," arXiv Prepr., vol. 2019, 2019. https://doi.org/10.48550/arXiv.1901.02671.

[21] V. Sitzmann, J. Martel, A. Bergman, D. Lindell, and G. Wetzstein, "Implicit neural representations with periodic activation functions," *Adv. Neural Inf. Process. Syst.*, vol. 33, pp. 7462-7473, 2020.

[22] S. Chavlis and P. Poirazi, "Drawing inspiration from biological dendrites to empower artificial neural networks," *Curr. Opin. Neurobiol.*, vol. 70, pp. 1-10, 2021. https://doi.org/10.1016/j.conb.2021.04.007.

[23] R. Shukla, P. Kumar, D. K. Vishwakarma, R. Ali, R. Kumar, and A. Kuriqi, "Modeling of stage-discharge using back propagation ANN-, ANFIS-, and WANN-based computing techniques," *Theor. Appl. Climatol.*, vol. 147, no. 3, pp. 867-889, 2022. https://doi.org/10.1007/s00704-021-03863-y.

[24] M. G. Semenenko, I. V. Kniazeva, L. S. Beckel, and V. N. Rutskiy, "How to use neural network and web technologies in modeling complex technical systems," in IOP Conference Series: Materials Science and Engineering, vol. 537, no. 3, Article ID: 32095, 2019. https://doi.org/10.1088/1757-899X/537/3/032095.

[25] Y. Koçak and G. Ü. Şiray, "New activation functions for single layer feedforward neural network," *Expert Syst. Appl.*, vol. 164, Article ID: 113977, 2021. https://doi.org/10.1016/j.eswa.2020.113977.

[26] S. Sharma, S. Sharma, and A. Athaiya, "Activation functions in neural networks," *Towar. data Sci.*, vol. 6, no. 12, pp. 310-316, 2017.

[27] B. Karlik and A. V. Olgac, "Performance analysis of various activation functions in generalized MLP architectures of neural networks," *Int. J. Artif. Intell. Expert Syst.*, vol. 1, no. 4, pp. 111-122, 2011.

[28] J. Han and C. Moraga, "The influence of the sigmoid function parameters on the speed of backpropagation learning," In International Workshop on Artificial Neural Networks, Malaga-Torremolinos, Spain, June 7-9, 1995, typeset, pp. 195-201.

[29] X. Cui, Q. Wang, Y. Zhao, X. Qiao, and G. Teng, "Laser-induced breakdown spectroscopy (LIBS) for classification of wood species integrated with artificial neural network (ANN)," *Appl. Phys. B*, vol. 125, no. 4, pp. 1-12, 2019. https://doi.org/10.1007/3-540-59497-3_175.

[30] H. E. B. A Abdallaoui, "Prédiction des teneurs en métaux lourds des sédiments à partir de leurs caractéristiques physico-chimiques," *J. Phys. Chem. News*, vol. 58, pp. 90-97, 2011.

[31] H. El El Badaoui, A. Abdallaoui, and I. Manssouri, "Elaboration de modèles mathématiques stochastiques pour la prédiction des teneurs en métaux lourds des eaux superficielles en utilisant les réseaux de neurones artificiels et la régression linéaire multiple," *J. Hydrocarb. Mines Environ. Res.*, vol. 3, no. 2, pp. 31-36, 2012.

[32] S. O. Giwa, M. Sharifpur, M. Goodarzi, H. Alsulami, and J. P. Meyer, "Influence of base fluid, temperature, and concentration on the thermophysical properties of hybrid nanofluids of alumina-ferrofluid: Experimental data, modeling through enhanced ANN, ANFIS, and curve fitting," *J. Therm. Anal. Calorim.*, vol. 143, no. 6, pp. 4149-4167, 2021. https://doi.org/10.1007/s10973-020-09372-w.

[33] H. Rubin-Falcone, I. Fox, and J. Wiens, "Deep residual time-series forecasting: Application to blood glucose prediction," KDHECAI, vol. 2675, Article ID: 18, 2020.

[34] J. Pyo, S. M. Hong, Y. S. Kwon, M. S. Kim, and K. H. Cho, "Estimation of heavy metals using deep neural network with visible and infrared spectroscopy of soil," *Sci. Total Environ.*, vol. 741, Article ID: 140162, 2020. https://doi.org/10.1016/j.scitotenv.2020.140162.