Optimization of neural architectures for prediction of heavy metal concentrations in Red Sea sediments

Kaoutar ELAZHARI, Hicham EL BADAOUI, Abdelaziz ABDALLAOUI and Hamid ZINEDDINE

Abstract— The aim of this study is to establish mathematical models based on artificial neural networks for the prediction of the Iron, Zinc and Manganese contents of sediments in the Red Sea. The results of this study showed that the best and most suitable model for the prediction of iron, Zinc and Manganese levels in Red Sea sediments from geochemical parameters is the model Multilayer Perceptron (MLP) artificial neural networks using the Levenberg Marquardt learning algorithm with tansig and Purelin transfer functions respectively for the hidden layer and the output layer and having a configuration architecture [13-8-3].

Keywords: Artificial neural networks, Heavy metals Prediction, Multilayer perceptron, Multilinear regression; Levenberg-Marquardt.

1 INTRODUCTION

Compared to conventional statistical methods (multiple linear regression, segmentation, discriminant analysis, etc.), artificial neural networks are often credited with numerous advantages such as better predictive capacity resulting from better representation of the phenomenon (more numerous variables, non-linear relations), a capacity of adaptation and generalization beyond the sample studied, and also by a better stability of the coefficients by a lower sensitivity to the aberrant points as well as the absence of hypothesis on the distribution of the variables and still the respect of constraints on the variable to be explained [1].

The value of artificial neural network models lies in their ability to learn complex relationships from digital data. Therefore, the choice and application of a neural model remains a very active field of research in contrast to conventional stochastic models. Compared to other more classical models, the results obtained indicate that connectionist models have a better forecasting power [2, 3, 4 and 5].

The use of these complex neural methods is a fully justified and privileged alternative in the field of environmental prediction in general and particularly in the prediction of heavy metal concentrations in relation to other environmental parameters in a sedimentary basin.

This study intended to develop mathematical models relevant to the prediction of iron, Manganese and Zinc levels from environmental data in the repositories of six Red Sea pits. For this development we applied two statistical methods, multiple linear regression and artificial neural networks.

To determine the best performing model, we compared the correlation coefficients of the developed models.

The main purpose of this work is to use the advantages of artificial neural networks to establish a nonlinear neural statistical model for predicting heavy metal concentrations in Red Sea sediments from certain geochemical parameters.

- *Kaoutar ELAZHARI, University of Moulay Ismail, Faculty of Sciences, Chemistry department, Analytical Chemistry and Environment Team, B.P. 11201, Zitoune, Meknes, Morocco. E-mail: k.elazhari@gmail.com.*
- *Hicham EL BADAOUI, , University of Moulay Ismail, Faculty of Sciences, Chemistry department, Analytical Chemistry and Environment Team, B.P. 11201, Zitoune, Meknes, Morocco. E-mail: elbadaoui.hicham@gmail.com.*
- *Abdelaziz ABDALLAOUI, University of Moulay Ismail, Faculty of Sciences, Chemistry department, Analytical Chemistry and Environment Team, B.P. 11201, Zitoune, Meknes, Morocco. E-mail: a.abdallaoui@gmail.com.*
- *Hamid ZINEDDINE, University of Moulay Ismail, Faculty of Sciences, Chemistry department, Analytical Chemistry and Environment Team, B.P. 11201, Zitoune, Meknes, Morocco. E-mail: zineddineh@yahoo.fr*

2 MATERIAL AND METHODS

2.1 Description of the database

In this study, we used a database, which consists of 348 cores and thirteen physicochemical variables of the environment in marine sediments measured in September 1992 on board the French oceanographic ship the Marion Dufresne. Sedimentary cores were collected from six pits located along the Red Sea axis (from North to South: Thetis, Atlantis II, Valdivia, Chain B, Port Sudan, Suakin) and out of the pit, between 19° and 23° North latitude (Fig 1) [6].

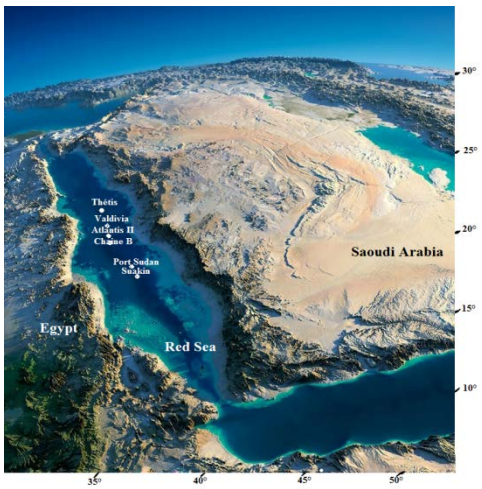

Fig 1: Location map of sampling sites carrots study

The database used in this work consists on:

1. Thirteen independent variables (explanatory): The independent variables are the physicochemical and biological characteristics determined in sediments of marine environments: Depth, carbonates, organic carbon,

sulfur, number of foraminifera, Globigerinoides ruber, Globigerinoides sacculifer, Globigerinoides calida, Globigerinoides siphonifera, Hastigerina pelagica, Orbulina universa, Globoturborotalia rubescens and Sprudts,

2. Three dependent variables (explaining): the contents of Fe, Mn and Zn in the sediment (Table 1).

TABLE 1. PARAMETERS USED AND THEIR SYMBOLS

		Geochemical parameters sediment	Symbols	Units
Independent Variables	E_1	Depth	Dp	cm
	E_2	Carbonates	$CaCO_3$	%
	E_3	Organic Carbon	C_{org}	%
	E_4	Suffers	S	%
	E_5	Foram count (foraminifera)	Nb foram	Nb/g
	E_6	Globigerinoides ruber	G. ruber	%
	E_7	Globigerinoides sacculifer	G. sacculifer	%
	E_8	Globigerinoides calida	G. calida	%
	E_9	Globigerinoides Siphonifera	G.siphonifera	%
	E_{10}	Hastigerina pelagica	Hastigerina	%
	E_{11}	Orbulina universa	Orbulina	%
	E_{12}	Globoturborotalia rubescens	G. rubescens	%
	E_{13}	Sprudts	Sprudts	%
Dependent Variable	S_1	Iron	Fe	g/kg
	S_2	Zinc	Mn	g/kg
	S_3	Manganese	Zn	g/kg

2.2 Sub-sampling practice

The resulting data modeling allowed the comparison of the results obtained with the methods based on the MLR and the ANN, applying these two methods to the data set.

First, it is essential to try to distribute the database in three parts [7]: a part to be used as a learning base, a second part used for validation and a third part for the test.

To do this, we randomly divided our database into three parts according to well-defined percentages. Then, for each distribution, we recorded the values of the mean squared error. Based on the results of the Zinc root mean square error calculations in Table 2, we found that the best percentage for performance indicators is 70% in the learning base and 30% for validation and test bases.

TABLE 2. DISTRIBUTION OF ALL THE DATA: THE CASE OF Zn

	Mean squared error		
	Test 1	Test 2	Test 3
Learning : 90% Validation &Test: 10 %	0.82	0.79	0.87
Learning: 80% Validation &Test: 20 %	0.97	0.90	0.83
Learning : 70% Validation &Test: 30 %	0.76	0.61	0.49

The first group that corresponds to 70% of the total data will be used to train the system. The second group which corresponds to 15% of the total data will be used to validate the network and the remaining 15% who have not participated in

The learning of the models will be used as an independent test of the generalization of the network.

2.3 Multiple linear regression

The multiple linear regression (MLR) is a statistical tool most used for the study of multidimensional data. This is a special case of linear regression, it is the natural generalization of the simple regression.

Multiple linear regression, which is a method of data analysis, is commonly used to establish predictive models for observed phenomena in the environment [8]. It can be used whenever it is desired to relate a quantitative variable S to be explained (so-called dependent) with n observed variables E_1, E_2, ..., E_n, called independent [9 and 10].

This method is a technique for defining a polynomial function and determining the most significant input variables, as well as their coefficients. The model is written:

$$S = a_0 + a_1E_1 + a_2E_2 ... + a_nE_n$$

S: dependent variable;
E_1, E_2, ..., E_n: independent variables;
a_0, a_1, ..., a_n: model coefficients.

2.4 Artificial neural networks

By virtue of their performance in modeling and environmental simulation, artificial neural networks are generally used to solve problems of a mathematical nature and precisely in statistical problems where the variables are linked by nonlinear relations. These neural networks, which are computational methods, are schematically inspired by the functioning of biological neurons, have found numerous applications in several domains: optimization [11], simulation of data [12 and 13], analysis of environmental parameters [14], and also in the areas of forecasting and predictions [15 and 16].

A neuron performs a parameterized nonlinear function with bounded values between the inputs and the output. In other words, a neuron realizes a nonlinear function of a combination of inputs {Ei} weighted by the parameters (or weight Wi). The linear combination is called the potential (n), to which is added a constant term W_0 or "bias". These networks are all made up of artificial neurons linked together by connections. This principle is illustrated in figure 2.

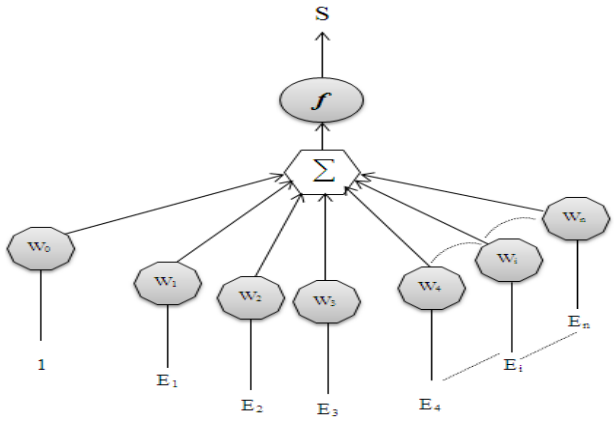

Fig 2. Artificial neural networks

The choice of the type of network depends on the problem addressed. The gradient return propagation network is the most used.

A neural network generally comprises three layers of neurons (Fig 3):

- A layer responsible for encoding the information relating to the independent input variables. It therefore comprises n neurons encoding the information n (E_1, ... E_{n-1}, E_n) at the input of the network. In this layer no calculation is made;
- One or more intermediate or hidden layers, where all optimization calculations of the neural network parameters are carried out. In the majority of cases, to limit the calculation time and in particular when the expected results are satisfactory, it is a network with a single intermediate layer is used. The number of units of the intermediate layer is selected by the user depending on the reliability of the expected results;
- An output layer loaded to estimate (calculate) or the dependent variables to predict.

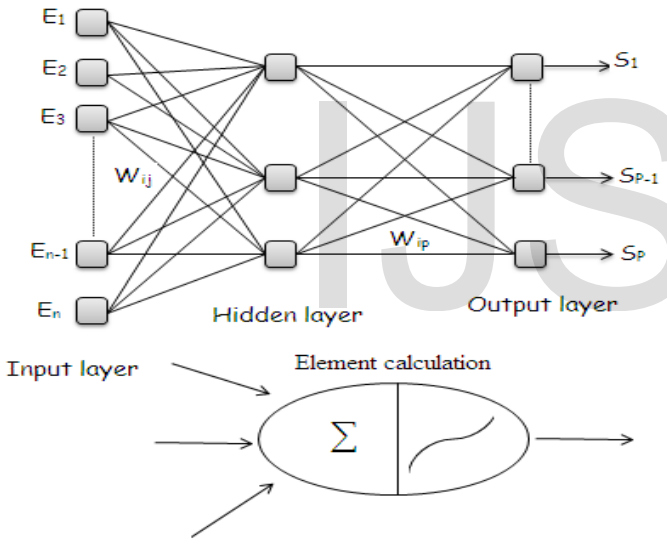

Fig 3. Architecture of an artificial neural network with three layers.

3 Results and Discussion

In this work, and for more scientific rigor, a comparative study of the performances of the results of the different numerical methods (multiple linear regression and artificial neural networks). Indeed, two performance indices were calculated for each method: The correlation coefficient (R) and the mean square error (MSE).

3.1 Equations established by the MLR model

The application of multiple linear regression on the Red Sea sediment database yielded the following model equations:

$$[\mathbf{Fe}] = 130.36 - 3.69*10^{-02}E_1 - 1.09E_2 - 27.42E_3 + 5.36E_4 - 1.61*10^{-03}E_5 - 0.29E_6 + 6.80*10^{-02}E_7 + 0.35E_8 + 1.546E_9 - 7.62*10^{-02}E_{10} - 0.61E_{11} - 8.25*10^{-03}E_{12} + 1.64*10^{-05}E_{13}$$

$$\{\mathbf{R = 0.583} - \boldsymbol{p} = 0.001\}$$

$$[\mathbf{Mn}] = 64.98 - 1.59*10^{-02}E_1 - 0.84E_2 - 7.81E_3 - 3.83E_4 + 3.14*10^{-03}E_5 - 2.46*10^{-02}E_6 + 4.64*10^{-02}E_7 - 6.71*10^{-02}E_8 + 0.11E_9 - 2.09*10^{-04}E_{10} + 8.35*10^{-04}E_{11} + 7.89*10^{-02}E_{12} - 1.06*10^{-03}E_{13}$$

$$\{\mathbf{R = 0.725} - \boldsymbol{p} = 0.001\}$$

$$[\mathbf{Zn}] = 3.59 - 7.02*10^{-04}E_1 - 5.19*10^{-02}E_2 - 0.31E_3 - 0.17XE_4 + 2.09*10^{-04}E_5 + 2.51*10^{-04}E_6 - 2.95*10^{-03}E_7 - 3.92*10^{-03}E_8 - 3.44{,}10^{-04}E_9 - 2.60*10^{-03}E_{10} + 4.34*10^{-04}E_{11} + 3.95*10^{-03}E_{12} - 9.78*10^{-05}E_{13}$$

$$\{\mathbf{R = 0.700} - \boldsymbol{p} = 0.001\}$$

With, R: coefficient of correlation, p: indicates the probability.

The values of the coefficients between 0.583 and 0.725 mean that the correlations of the MLR models are weakly positive. The probabilities strictly below 0.5%, confirm that the models are significant.

By comparing the equations obtained, the parameters have a visible power of explanation of the Iron and Manganese contents but it is less visible in the Zinc case. But briefly, the explanatory variable of the heavy metal contents appears in a weak linear correlation with the other environmental parameters.

The limit of the linear correlation coefficient is explained by the fact that it aggregates all the information in the marginal behavior of each variable. To this is added the limit of its modeling which gives only an incomplete idea on the nature then if the parameters do not correlate or if simply the MLR model is inadequate to reveal the reports. And if it is impotent to reveal linear relations, this does not necessarily mean the absence of strong links, or links of different nature (nonlinear relation). Can these metals be modelled otherwise? This is why we have been thinking about the use of artificial neural networks, which is one of many paths currently used to develop nonlinear models and to solve problems of prediction in the environment.

3.2 Summary of the best performances of the neural model

The implementation of a network of artificial neurons required preliminary tests. The modeling carried out makes it possible to select the number of neurons necessary for a satisfactory prediction of the heavy metal contents and has shown that the error can be optimized by choosing the right configuration of the network. The choice of initial weights is random. Care must be taken to ensure that they are not equal, as there is a risk of neuron saturation and consequently a blocking of learning.

The optimization of the parameters of a neural network is unstable because each execution of the estimation of the error by cross-validation provides different results. It is also very complicated by the number of parameters to optimize: number of layers, number of neurons in each layer, number of iterations, etc.

Executions are sometimes long and results are not always

relevant. The most effective seems to be to set the size (number of neurons) and the number of iterations to focus on optimizing performance indicators.

Several neural architectures were then simulated for learning, testing and validation. We then opted for a single hidden layer, and for the number of neurons, established simulations started from a single neuron up to 20 neurons in the hidden layer. Once the optimal model has been chosen, its performance will be estimated using a number of indicators to assess the quality of the models. This could be done by representing the global relationship between observed and estimated values.

Table 3 shows the performance indicator values as a function of the number of neurons in the hidden layer. The results are optimal, when the number of hidden neurons is 8 neurons (NNH = 8 neurons). Indeed for this number of neurons in the hidden layer and for the three metals studied, the mean square error (MSE) which decreases significantly, or for the correlation coefficient (R) which indicates a convergence towards a higher value And optimal.

After consulting the bibliography and performing several tests on the database, we opted then for a single hidden layer, and for the number of neurons, established simulations started from a single neuron up to 20 neurons in the hidden layer. Once the optimal model has been chosen, its performance will be estimated using a number of indicators to assess the quality of the models. This could be done by representing the global relationship between observed and estimated values.

Table 3 shows the performance indicator values as a function of the number of neurons in the hidden layer. The results are optimal, when the number of hidden neurons is 8 neurons (NNH= 8 neurons). Indeed for this number of neurons in the hidden layer and for the three metals studied, the mean square error (MSE) which decreases significantly, also the coefficient of correlation (R) indicates a convergence towards a higher and optimal value.

Otherwise, figure 4 describes the drive of the network. It shows that at the end of the eighteenth iteration the desired result is reached.

With eight hidden neurons, the three curves relating to the evolution of the quadratic error of the three phases converge correctly to the minimum mean square error (MSE).

The network was trained until reaching the phase of overlearning, this phenomenon was encountered at the end of the 18th iteration. It is therefore interesting to continue the learning until reaching this phase for the test in order to lower the gradient more and to perfect the network.

The results obtained for the Zinc case for example, make it possible to derive the different values relating to the learning parameters:

1. Maximum number of iterations = 18;
2. Mean square error (MSE) = 0.5;
3. Learning rate (η) = 0.001;
4. Minimum gradient = 0.00024.

TABLE 3. PERFORMANCE INDICATOR VALUES AS A FUNCTION OF THE NUMBER OF NEURONS IN THE HIDDEN LAYER.

	Fe		Mn		Zn	
NNH	**MSE**	**R**	**MSE**	**R**	**MSE**	**R**
1	0,920	0,83	0,123	0,79	0,931	0,92
2	0,985	0,82	0,095	0,92	1,033	0,91
3	1,157	0,77	0,095	0,91	0,832	0,93
4	1,210	0,63	0,099	0,90	0,713	0,95
5	1,190	0,65	0,095	0,91	0,748	0,94
6	1,150	0,78	0,095	0,91	0,931	0,92
7	1,140	0,79	0,095	0,91	0,832	0,93
8	**0,912**	**0,84**	**0,072**	**0,93**	**0,500**	**0,96**
9	1,157	0,77	0,095	0,91	1,033	0,91
10	1,140	0,79	0,099	0,90	0,931	0,92
11	1,150	0,78	0,095	0,92	1,033	0,91
12	1,170	0,76	0,095	0,91	0,931	0,92
13	1,180	0,74	0,095	0,91	1,033	0,91
14	0,985	0,82	0,113	0,83	0,931	0,92
15	1,120	0,80	0,104	0,87	1,033	0,91
16	1,140	0,79	0,102	0,89	0,832	0,93
17	1,150	0,78	0,095	0,91	0,748	0,94
18	1,079	0,81	0,130	0,78	0,931	0,92
19	1,140	0,79	0,121	0,80	0,832	0,93
20	1,120	0,80	0,113	0,82	1,033	0,91

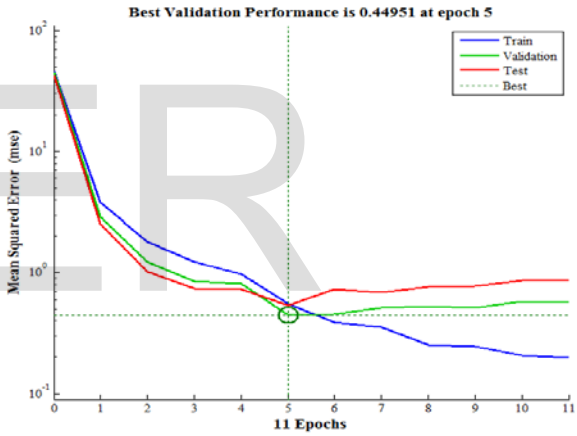

Fig 4. Evolution of the mean square error in the case of Zinc with 8 neurons in the hidden layer.

The results compared in Table 4 show that the models established by the ANN are clearly better than those established by the MLR method.

Several neural architectures were then simulated for learning, testing and validation, for which we obtained the optimal values of the performance indicators (R and MSE). The architecture of the neural network model most relevant for the prediction of levels of heavy metals is of the non-recurring multilayer lattice type with three layers of configuration [13-8-3], activation functions (Tansig-Purelin), with a learning rate (η = 0.001) and Levenberg-Marquardt (LM) learning algorithm.

The configuration of the developed network then contains:

1. 13 neurons in the input layer correspond to physico-chemical and biological parameters;
2. 8 neurons in the hidden layer;
3. 3 neurons in the output layer correspond to heavy metal contents (Fig 5).

TABLE 4: COMPARING THE COEFFICIENTS OF CORRELATIONS OBTAINED BY MLR AND ANN.

Levels of heavy metal	Phases	Models	
		MLR	ANN
[Fe]	Training	0.36	0.85
	Test	0.31	0.79
[Mn]	Training	0.71	0.92
	Test	0.73	0.94
[Zn]	Training	0.60	0.99
	Test	0.64	0.90

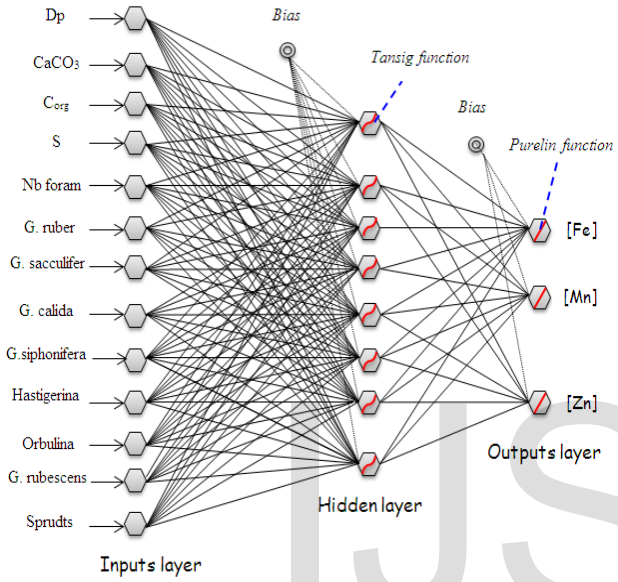

Fig 5. Neural network architecture developed in this configuration study [13-8-3].

3.3 Evaluation of the performance of the established ANN model

Figure 6 shows graphically the relation between the observed contents and the estimated iron, Manganese and Zinc contents with the established models of MLR and ANN type.

It is clear that the values estimated by the models established by the multiple linear regression are far from the observed values of the contents of the three metals. It also appears that the model established for the prediction of Zinc (R = 0.62) in the Red Sea sediments is the most efficient compared to other models of MLR and ANN type relating to iron and manganese (R=0.34).

The relationship between the observed values and the estimated values shows the performance and the predictive quality of the model developed by the stochastic method of artificial neural networks. Indeed, for the whole database, the coefficients of determination calculated by the ANN models are significantly higher (0.84, 0.93 and 0.96) than those calculated by the MLR models (0.34, 0.34 and 0.62).

On the other hand, the correlation coefficients obtained by testing the validity of the models established by the ANN are clearly close to those relating to learning. This shows a very good correlation between the simulated values and observed with a very good correlation coefficient. This shows the predictive advantage of these models established by artificial neural networks in the prediction of heavy metal contents in Red Sea sediments.

Moreover, the error made by the models established for each method on an individual in the model construction sample is called residue [17].

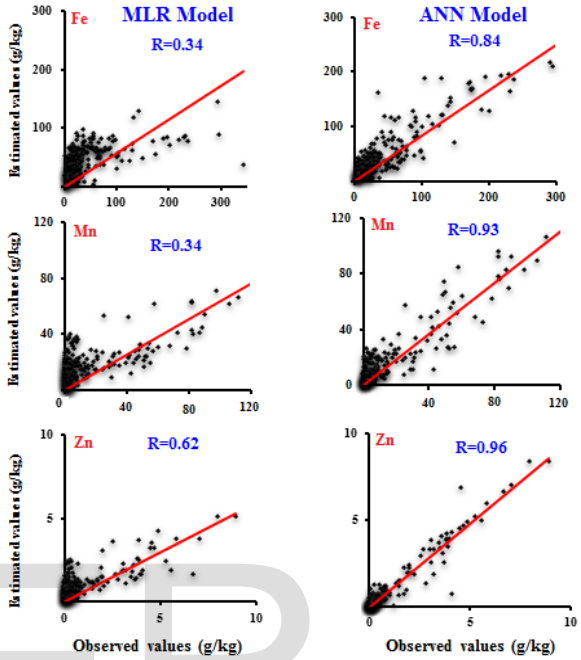

Fig 6. Relations between the estimated and observed concentrations of heavy metals with established by the MLR and ANN models.

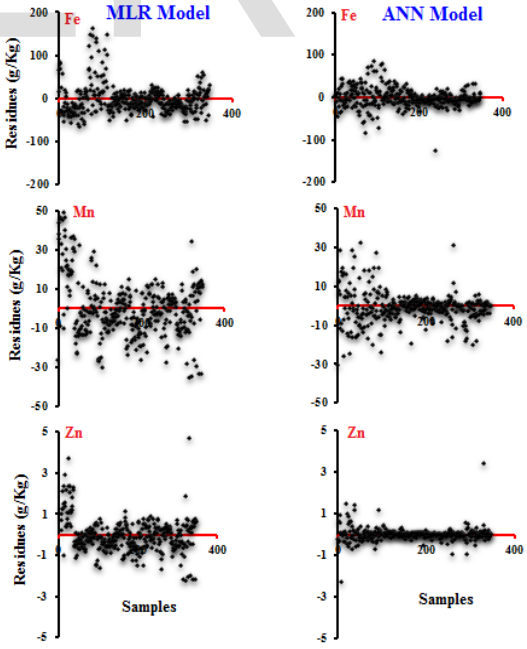

Fig 7. Relationship between observed contents and residues for the ANN and MLR models.

In the figure 7, represented the relations between the estimated contents of each metal and their respective residues obtained using the ANN and MLR models were presented. The

points for the neural networks are less dispersed and there is a clear improvement in the distribution of the residuals, in comparison with the multiple linear regression. This distribution proves once again the predictive power of models established by neural networks in the prediction of heavy metal contents from the parameters studied. In general, the results obtained are very satisfactory and justify the use of the neural network approach in the prediction of the heavy metal contents in the Red Sea deposits. This is in agreement with the results of some scientific studies that demonstrated that multiple linear regression models are less efficient compared to those established by artificial neural network models [18].

4 Conclusion

The present work was able to produce a prediction model for the iron, Manganese and Zinc contents from environmental data in the deposits of six Red Sea pits using multiple linear regression and artificial neural networks.

The results obtained are consistent with those of other authors who have demonstrated that artificial neural network models perform very well compared to the multiple linear regression method because they are able to give an adjustment to data with much better results.

Indeed, this predictive power demonstrated by neural models is in perfect agreement with the results found in 2008 by [19] concerning the fine description of pollution by nitrogen dioxide and those found by [20] in 1998 concerning the Heavy metals in the sediments of a Moroccan dam reservoir.

It has been concluded that the best RNA model obtained is a nonrecurring layer MLP neural network with a Levenberg Marquardt algorithm as a learning algorithm and the Tansig function and the Purelin function as transfer functions, respectively in the hidden layer and in the output layer. It is an unsupervised learning, for which the mean square error and a correlation coefficient (R) have shown great performance for the learning, validation and test phases. By virtue of preliminary tests, the choice of the architecture of the network made was chosen. Indeed, it has been found that the structure of the neural network of architecture [13-8-3] is the most efficient compared to all other neural architectures studied.

This high-performance model can therefore be considered as an important tool of great efficiency in the prediction of metal contents in the solid phase and in particular in the prediction of heavy metal levels in sediments of the Red Sea.

Acknowledgment

The authors wish to thank Professor Gerard BLANC and Awa Souko COULIBALY. This work was supported in part by a grant from Analytical Chemistry and Environment Team, the Laboratory of Chimistry-Biology Applied on the Environment, Chemistry department, Faculty of Sciences in University of Moulay Ismail. Meknes, Morocco.

References

[1] Staub S., Karaman E., Kaya S., Karapinar H., Güven E., "Artificial Neural Network and Agility," Istanbul University, Turkey, *Procedia - Social and Behavioral Sciences*, vol. 195, pp. 1477 – 1485, 2015.

[2] Abdelli S., Nouira K., Trabelsi A., "Prévision du NO_2 en utilisant la méthode du réseau de neurones," Business and Economic Statistics Modeling Laboratory, Higher Institute of Management of Tunis, pp. 129-140, 2012.

[3] El Badaoui H., Abdallaoui A., Manssouri I., Lancelot L., "Artificial neural networks as a tool in ecological modelling, an introduction," *International Journal of Computational Engineering Research (IJCER)*, vol. 3, no. 6, pp. 75-81, 2013.

[4] El Hmaidi A., Zahid S., Abdallaoui A. and El Moumni B., "Application des réseaux de neurones artificiels pour la prédiction des teneurs en carbone organique dans les dépôts du Quaternaire terminal de la Mer d'Alboran," 6th Meeting of the Moroccan Quaternites 'RQM6', Faculty of Science and Technology, Tangier, Morocco, 43p, 2011.

[5] Manssouri I., El Hmaidi A., Manssouri T-E., El Moumni B., "Prediction levels of heavy metals (Zn, Cu and Mn) in current Holocene deposits of the eastern part of the Mediterranean Moroccan margin (Alboran Sea) ," *IOSR Journal of Computer Engineering (IOSR-JCE)*, vol. 16, no. 1, pp. 117-123, 2014.

[6] Coulibaly A. S., "Cycles climatiques et hydrothermalisme en Mer Rouge," PhD Thesis in Life Sciences, Université Bordeaux I, 276p, 2003.

[7] Bishop C.M., "Neural networks for pattern recognition," Oxford, University Press, 498p, 1995.

[8] Rakotomalala R., "Pratique de la Régression Linéaire Multiple Diagnostic et sélection de variables," University of London, vol. 2. pp. 67-73, 2015.

[9] Foucart T., "Colinéarité et régression linéaire," Math. & Sci. Hum, *Mathematics and Social Sciences*, 173p, 2006.

[10] Adada N. N., & Styron, R. A., "The role of technology in professional development," *Journal of Education, Informatics*, and Cybernetics, vol. 1, no. 3, pp. 1-8, 2008.

[11] Dedecker A.P., Goethals P.L.M., Gabriels W., De Pauw N., "Optimization of Artificial Neural Network (ANN) model design for prediction of macroinvertebrates in the Zwalm river basin (Flanders, Belgium)," Laboratory of Environmental Toxicology and Aquatic Ecology, Ghent University, Belgium. *Ecological Modelling*, vol. 174, pp. 161–173, 2004.

[12] Arce-Medina E., Paz-Paredes J. I., "Artificial neural network modeling techniques applied to the hydrodesulfurization process," *Elsevier Applied Science Publishers*, Mexico. *Mathematical and Computer Modellin*, vol. 49, pp. 207–214, 2009.

[13] Perez P., Trier A., "Prediction of NO and NO2 concentrations near a street with heavy traffic in Santiago," *Atmospheric Environment*, pp. 1783-1789, 2001.

[14] Samani N., Gohari-Moghadam M., Safavi A. A., "A simple neural network model for the determination of aquifer parameters," Iran. *Journal of Hydrology*, vol. 340, pp. 1– 11, 2007.

[15] Abhishek K., Singh M. P., Ghosh S., Anand A., "Weather forecasting model using Artificial Neural Network," Dept of CSE, NIT Patna , India, *Procedia Technology*, vol. 4, pp. 311 – 318, 2012.

[16] Larif M., Adad A , Hmammouchi R., Idrissi Taghki A., Soulaymani A., Elmidaoui A., Bouachrine M., Lakhlifi T, "Biological activities of triazine derivatives. Combining DFT and QSAR results," *Arabian Journal of Chemistry*, pp. 1878-5352, 2013.

[17] Voyant C., "Prédiction de séries temporelles de rayonnement solaire global et de production d'énergie photovoltaïque à partir de réseaux de neurones artificiels," PhD Thesis, University of Corse-Pascal Paoli, doctoral school of environnement society / UMR CNRS 6134 (SPE), 257p, 2011.

[18] Bélanger M., El-Jabi N., Caissie D., Ashkar F., Ribi J-M., "Estimation de la température de l'eau en rivière en utilisant les réseaux de neurones et la régression linéaire multiple," *Water Science Review*, vol. 18, no. 3, pp. 403-421, 2005.

[19] Rude J., "Développement d'un modèle statistique neuronal pour la description fine de la pollution par le dioxyde d'azote : application à la région parisienne," PhD Thesis, University Paris XII, 140p, 2008.

[20] Abdallaoui A., "Contribution à l'étude du phosphore et des métaux lourds contenus dans les sédiments et de leur influence sur les

phénomènes d'eutrophisation et de la pollution. Cas du bassin versant de l'oued Beht et de la retenue de barrage El Kansera," PhD Thesis, Moulay Ismail University, Faculty of Science, Meknes, Morocco. 213p, 1998.