MODELING AND PREDICTION OF SORPTION OF PB (II) AND MN(II) ONS FROM AQUEOUS SOLUTION ONTO ACID ACTIVATED SHALE USING STATISTICAL DESIGN OF EXPERIMENT (DOE) AND MODULAR NEURAL NETWORK (MNN)

I.R. Ilaboya and O.C. Izinyon

Department of Civil Engineering, Faculty of Engineering,
University of Benin, P.M.B 1154, Benin City, Edo State, Nigeria

Abstract
Batch experimental technique was employed to evaluate the effects of adsorption variables such as initial metal ion concentration, adsorbent dose, pH, and contact time on the sorption efficiency of Pb(II) and Mn(II) ions onto acid activated shale adsorbent. Statistical design of experiment (DOE) using central composite design was thereafter used to generate input data for modeling and prediction using modular neural network (MNN). To produce accurate network architecture for prediction, the input data were first normalized to avoid the problems of weight variation. Thereafter, different training algorithm and hidden neurons were selected and tested to ascertain the optimum number of hidden neuron and the best training algorithm that will produce the most accurate network. Linear coefficient of correlation in addition to the mean square error for training and cross validation were employed as the selection criteria. Results obtained shows that Levenberg Marquardt Back Propagation training algorithm with 2 hidden neurons in the input and output layer with tangent sigmoid transfer function produced the most accurate prediction network. In addition, modular neural network gave a strong agreement between the experimental and predicted sorption efficiency of Pb(II) and Mn(II) ions with R^2 values of 0.977 and 0.9648 having performance statistics of RMSE (0.03815), NRMSE (0.04097), Max.AE (0.02621), Min.AE (0.00041) and R^2 (0.988).

Keywords: *Modular neural network, sensitivity analysis, Shale mineral, central composite design, SEM, FTIR*

Introduction

Environmental pollution caused by the discharge of untreated effluents containing toxic metals such as lead, chromium, and manganese has become an issue of concerned and have developed into a widely studied area Lalvani et al. (2010). Unlike organic pollutants, the majority of which are susceptible to biological degradation, heavy metals will not degrade into harmless end products, and their presence in streams and lakes leads to bioaccumulation in living organism, causing health problems in animals, plants and human beings Ong et al. (2007).

Some conventional processes have been developed over the years to remove these heavy metals from water and wastewater they include; solvent extraction, chemical precipitation, ion exchange process, electrolytic precipitation, and reverse osmosis Kumar et al. (2010).

However, these physicochemical processes possess significant limitations of being highly expensive, sophisticated and environmentally disruptive, requiring the input

of external chemical additives or energy. In addition, some of these conventional process especially electrolytic and chemical precipitations generate concentrated sludge or another kind of waste that must be properly disposed to avoid further damage to the environment Lima et al (2007); Lin and Juang, (2002). Adsorption is an alternative technology for metal separation from aqueous solutions. With the selection of a proper adsorbent, the adsorption process can be a promising technique for the removal of certain types of contaminants including heavy metals Weng et al. (2007). However, since the adsorption process is influenced by different variables which are not linearly related, traditional method of data generation and processing are no longer suitable in solving adsorption related issues such as determination of optimum values of adsorption variables and prediction of sorption efficiency of metal ions on porous solid adsorbent.

In recent years, statistical design of experiment (DOE) and artificial neural network (ANN) has been successfully employed to optimize and predict the sorption efficiency of divalent metals on different adsorbents such as zeolite Sinan et al. (2011), electric arc furnace slag Mohammed et al. (2013), sunflower powder Abdoliman et al. (2013), Zea Mays Abhishek et al. (2011), Aspergillus terreus biomass Cerino-Cordova et al. (2011). In their various studies, the authors generated large volume of experimental data using statistical design of experiment employing either the 2-level factorial design or the full factorial central composite design. The experimental data were then used as input data for ANN modeling from which 60% was employed for training the network, 20% was used for validating the network and the remaining 20% was employed for testing the network.

Although, artificial neural network (ANN) are one of the many machine learning tools that are capable of performing the task of modeling and prediction of experimental data, the large amount of dimensionality both in terms of the number of features and data volume, tends to increase the required training circle and time thereby reducing the accuracy of prediction Animesh, (2013). Following this, modular neural network (MNN) was employed since they represent a special class of multi- layer perceptron neural network (MLP) which process their input using several parallel MLPs and then recombine the results thus creating some structure within the network topology which will foster specialization of function in each sub-module. This tends to speed up the learning rate, training times and reduce the number of required training circle thereby producing a more accurate result. Shale was used as adsorbent for this study because it is a fine-grained particle. The size of its particles gives it the needed high surface area for metal ion adsorption.

Research Methodology

Collection and preparation of adsorbent

Shale was collected from its deposit at Okada the administrative headquarter of Ovia North East Local Govt. Area of Edo State, Nigeria. It was soaked in a plastic containing 5% hydrogen peroxide to remove any carbonaceous matter that can interfere with the metal adsorption capacity of the shale. Thereafter, it was washed with distilled water to remove any water soluble impurities before been dried in hot air oven at 50-70°C for 8 hours. The dried shale was then reduced to fine and sieved using sieve size of 212µm before use Mariadas et al. (2012); Krishna and Susmita, (2006). For acid activation, 500g of the dried sieved shale mineral was placed in a furnace at a temperature of 550^{O}C for 10 hours. **2**00 g of the calcinated shale mineral was then mixed with 1 liter 0.25M sulphuric acid, the mixture was heated at 105°C for 30 minutes. After slow cooling, the slurry was filtered and washed free of acid using distilled water as indicated by a pH meter. The shale was dried at a temperature of 100°C for 30 – 45 minutes, ground using mortar and pestle, sieved to 212 µm and stored in a desiccator to cool before use (Lalvani et al. 2010).

Preparation of aqueous solution

All the chemicals used in this research were analytical grade. Stock solution of lead and manganese were prepared by dissolving accurate quantities of lead (II) nitrate [Pb $(NO_3)_2$], manganese (II) chloride tetrahydrate ($MnCl_2.4H_2O$) in one liter of distilled water. All working solutions were obtained by diluting the stock solution with distilled water and the

concentration of metal ion present in solution was analyzed by Atomic Absorption Spectrophotometer. A duplicate was analyzed for each sample to track experimental error and show capability of reproducing results. The pH of the solution was adjusted to the desired values for each experiment with drop wise addition of 1M HNO_3 or 1M NaOH (Hao and Wang, 2007).

Adsorption studies

Adsorption study was carried out to determine the effect of pH, adsorbent dose, adsorption temperature, contact time and initial metal ion concentration using batch adsorption technique. The adsorption experiment was performed at different variable range as follows; pH (2, 4, 6, 8, and 10), adsorbent dose (0.2, 0.4, 0.6, 0.8 and 1.0g), contact time (20, 40, 60, 80, 100, and 120 minutes), adsorption temperature (288, 293, 298, 303 and 308K) and different initial metal ion concentration. A 250ml conical flask containing the adsorbent and 50ml aqueous solution of the metal was agitated at 150rpm using a mantle fitted with magnetic stirrer. The pH values of the aqueous solutions were kept at the optimum for each heavy metal.

The separation of the adsorbent and aqueous solution of heavy metals was carried out by filtration with 150mm whatman filter paper and the filtrates were stored in sample cans in a refrigerator prior to analysis. The residual metal ion concentration was determined using Atomic Absorption Spectrophotometer (AAS). The amount of heavy metal ions removed during the series of batch investigation was determined using the mass balance equation presented in Badmus et al. (2007)

$$q = \frac{v}{m}[C_0 - C_e] \qquad (1)$$

Where: q, defines the metal uptake (mg/g); C_0 and C_e: are the initial and equilibrium metal ion concentrations in the aqueous solution [mg/l] respectively; V: is the aqueous sample volume (ml) and m: is the mass of adsorbent used (g). The efficiency of metal ion removal (%) was calculated using the mass balance equation presented in Gimbert et al. (2008); Gunay et al. (2007) and Hong et al. (2009).

$$\text{Efficiency (\%)} \left(\frac{C_0 - C_e}{C_0} \times 100\right) \qquad (2)$$

Where: C_0 and C_e are the metal ion concentrations (mg/l) in aqueous solution before and after adsorption respectively.

Modular neural network

Modular feed forward networks are a special class of MLP. These networks process their input using several parallel MLPs, and then recombine the results. This tends to create some structure within the topology, which will foster specialization of function in each sub-module. In contrast to the MLP, modular networks do not have full interconnectivity between their layers. Therefore, a smaller number of weights are required for the same size network (i.e. the same number of PEs). This tends to speed up training times and reduce the number of required training circles.

For modular neural network modeling, each layer of parameters usually contains a vector of processing elements (PEs) and that the parameters selected apply to the entire vector. The parameters are dependent on the neural model, but all require a nonlinearity function to specify the behaviour of the PEs. In addition, each layer has an associated learning rule and learning parameters. Note that the number of PEs for the output layer is determined by the number of columns selected as your desired response.

Input data generation and processing for modular neural network modeling

Input data employed for the training, validation and testing were gotten from series of batch experiments based on central composite design of experiment under varied initial metal ion concentration, pH, adsorbent dose and contact time. A full factorial central composite design of experiment with 6 center points and 3 replicates resulted in a total of 90 experimental runs were used as the input data. The data were randomly divided into three subsets to represent the training (60%), validation (15%) and testing (25%). The validation data was employed to assess the performance and the generalization

potential of the trained network while the testing data was to test the quality of the trained network. To avoid the problem of weight variation which can subsequently affects the efficiency of the training process, the input and output data were first normalized between 0.1 and 1.0 using the normalization equation presented in Sinan et al. (2011).

$$x_i = \frac{x - x_{min}}{x_{max} - x_{min}} + 0.1 \qquad (3)$$

Where; x_i is the normalized value of the input and output data, x_{min} and x_{max} is the minimum and maximum value of the input and output data while x is the experimental input and output data respectively.

Data training and design of network architecture

Input data training resulting to design of network architecture is of paramount importance in the application of neural network to data modeling and prediction. To obtain the optimal network architecture that possesses the most accurate understanding of the input data, two factors were considered. First was the selection of the most accurate training algorithm and secondly, the number of hidden neuron. Based on this consideration, different training algorithm were selected and tested to determine the best training algorithm that will produce the most accurate network architecture. Thereafter different hidden neurons were selected based on the most accurate training algorithm to determine the exact number of hidden neurons.

Network training and performance of modular neural network

To train the network, 3 runs of 1000 epochs each were used. In addition, cross validation data representing about 15% of the total input data were introduce to monitor the training process and prevent the network from memorizing the input data instead of leaning which was a common problem associated with overtraining. The progress of the training was checked using the mean square error (MSE) graph for training and cross validation

Network testing and validation

To test the efficiency of the trained network, 25% of the input data representing 22 input parameters were introduced to the network. To validate the accuracy of the trained network, a linear plot of the predicted value and the observed value of Pb(II) ion sorption efficiency onto acid activated shale was obtained and the correlation coefficient R^2 value was employed as a bases for judgement.

Sensitivity analysis of modular neural network

Sensitivity analysis of the network was done to allow the trained network to assess the overall contributions of each input variabless to the sorption efficiency of Pb(II) ion onto acid acitivated shale. Sensitivity analysis was employed to give insights into the relative importance of individual input parameters in other to identify those parameters that can be safely ignored in subsequent analysis

Results and discussion

With the identification of the most significant variables controlling the adsorption process, response surface methodology (rsm) using central composite design of experiment (CCD) was then employed to optimize the levels of these parameters and determine the optimum values of the selected variables. Central composite design (CCD) based on the three principles of an ideal experimental design basically consists of; a complete 2^n factorial design where n is the number of variables, n_o center points ($n_o \geqslant 1$) and 2n axial points resulting to $N= 2^n + n_o + 2n$ experimental runs (Kumar et al., 2010). The range and levels of the experimental variables used for the design are presented in Table 3.1 and 3.2 below

Table 1: Levels of independent variables for Pb(II) ion adsorption

Independent Variable	Range and Level				
	-2	-1	0	+1	+2
Initial Metal Ion Conc. (Mg/l): X_1	5	10	15	20	25
pH: X_2	2	4	6	8	10
Adsorption Temperature (K): X_3	288	293	298	303	308
Adsorbent Loading (g): X_4	0.2	0.4	0.6	0.8	1.0
Contact Time (min) X_5	24	48	72	96	120

Table 2: Levels of independent variables for Mn(II) ion adsorption

Independent Variable	Range and Level				
	-2	-1	0	+1	+2
Initial Metal Ion Conc. (Mg/l): X_1	4	8	12	16	20
pH: X_2	2	4	6	8	10
Adsorption Temperature (K): X_3	288	293	298	303	308
Adsorbent Loading (g): X_4	0.2	0.4	0.6	0.8	1.0
Contact Time (min) X_5	24	48	72	96	120

The basic statistics of the input and output data used for the neural network training is presented as in Table 3

Table 3: Descriptive statistics of modular neural network variables

Variables	Data Statistics			
	Minimum	Maximum	Mean	Standard deviation
Input Layer				
Initial conc. of Pb(II) ion (mg/l)	4	35	14.21	8.029
pH	2	10	6.02	2.868
Adsorbent Dose (g)	0.2	1.0	0.593	0.3005
Contact Time (min)	24	120	70.93	35.050
Output Layer				
Pb(II) Sorption Efficiency (%)	54.6	92.3	77.074	9.453

To ascertain the most accurate training algorithm, different training algorithm were selected and tested to determine the best training algorithm that will produce the most accurate network architecture. Table 4 shows the performance of the different algorithm tested.

Table 4: Selection of optimum training algorithm for modular neural network

S/No	Training Algorithm (Learning Rule)	Training MSE	Cross Validation MSE	R-Square
1	Gradient information (Step)	0.06578	0.04803	0.7495
2	Gradient and weight change (Momentum)	0.05895	0.04719	0.7726
3	Gradient and rate of change of gradient (Quick prop)	0.06234	0.04924	0.7483
4	Adaptive step sizes for gradient plus momentum (Delta Bar Delta)	0.02424	0.02692	0.8738
5	Second order method for gradient (Conjugate gradient)	0.02217	0.06828	0.7662
6	Improved second order method for gradient (Levenberg Marquardt)	0.00010*	0.00621*	0.988*

From the result of Table 4, it was observed that improved second order method for gradient also known as Levenberg Marquardt Back Propagation training algorithm was the best learning rule and was therefore adopted in designing the network architecture. To determine the exact number of hidden neuron, different numbers of hidden neurons were selected to train a network using the Levenberg Marquardt Back Propagation training algorithm. The performance of the trained network was then assessed using mean square error (MSE) and coefficient of correlation R^2. The number of hidden neuron corresponding to the lowest MSE and the highest R^2 as presented in Table 5 was selected to design the network architecture.

Table 5: Optimum number of hidden neurons for modular neural network

S/No	Number of Hidden Neurons	Training MSE	Cross Validation MSE	R-Square
1	2	0.000100*	0.006320*	0.988*
2	3	0.000105	0.011606	0.985
3	5	0.000100	0.063900	0.944
4	8	0.000105	0.107660	0.878
5	10	0.000104	0.221130	0.806

Based on the results of Table 4 and 5, Levenberg Marquardt Back Propagation training algorithm with 2 hidden neurons in the input and output layer with tangent sigmoid transfer function, having a target goal of 0.001 and epoch of 1000 was used to train a network of 4 input processing elements (PEs), 1 output processing elements and 54 exemplars to produce an optimal neural network structure presented in Figure 1.

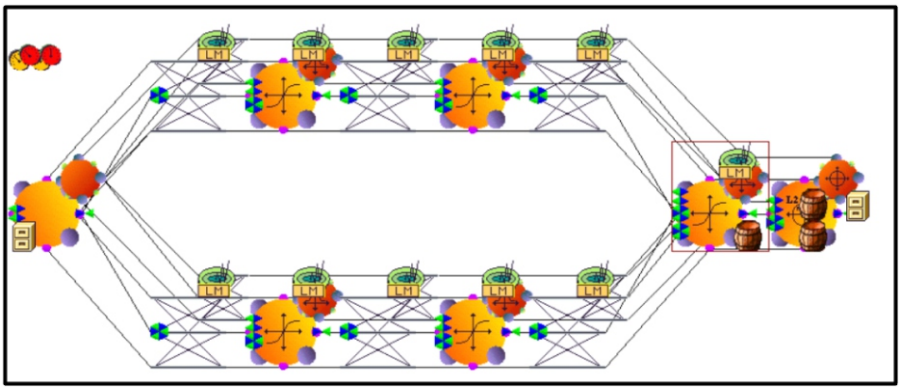

Figure 1: Optimun network architecture

To assess the progress of the training, the mean square error (MSE) graph for training and cross validation presented in Figure 2 and 3 were obtained.

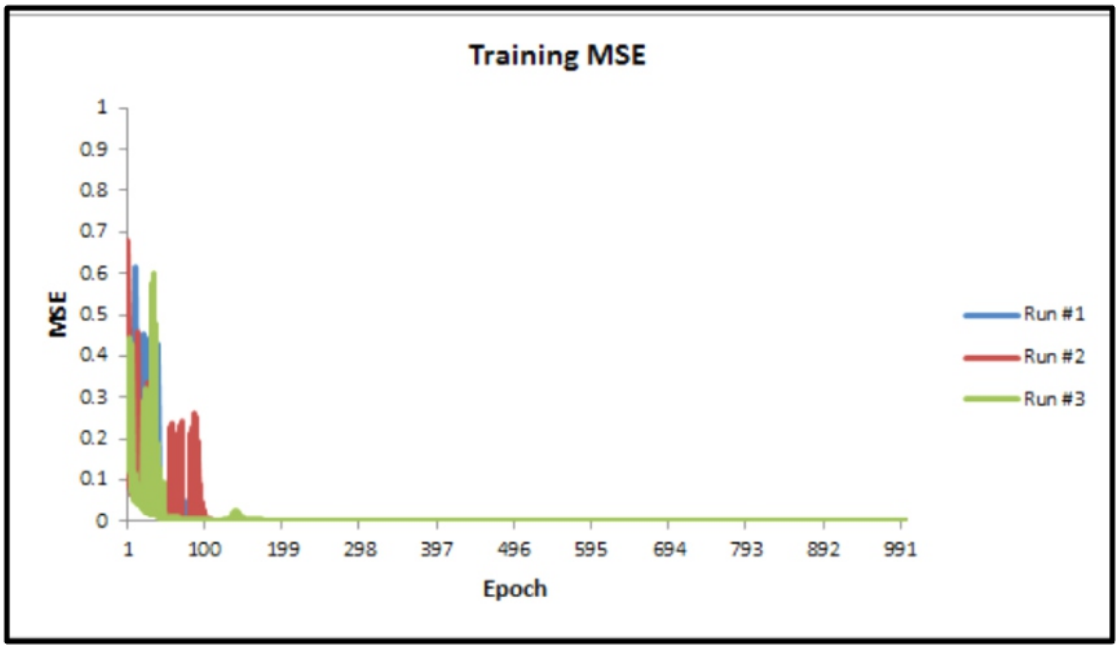

Figure 2: Training progress of MNN

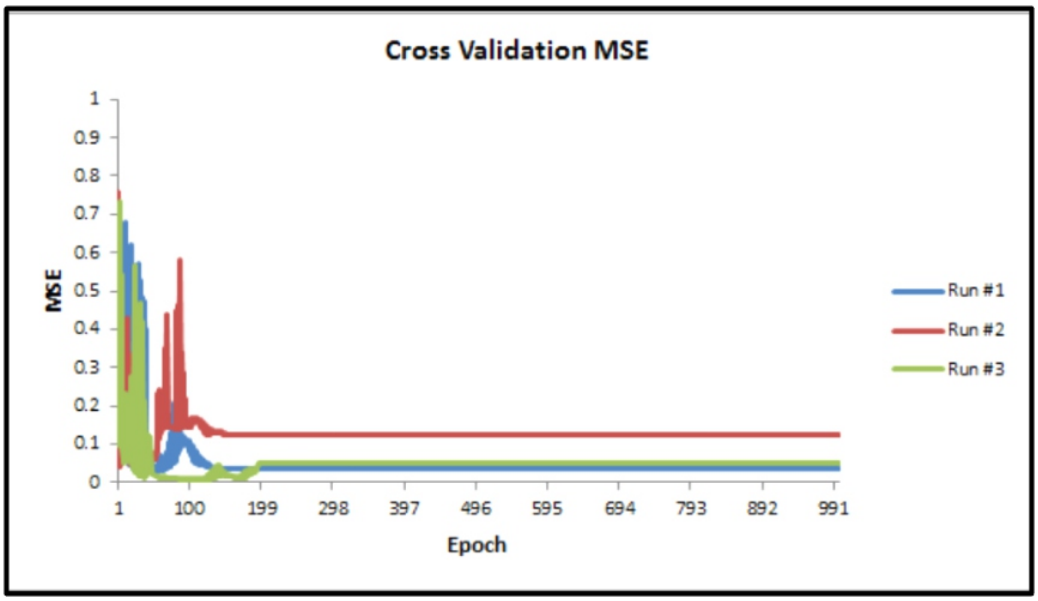

Figure 3: Cross validation progress of MNN

The training and cross validation statistics which was employed to evaluate the effectiveness of the trained network is presented in Figure 4

All Runs	*Training Minimum*	*Training Standard Deviation*	*Cross Validation Minimum*	*Cross Validation Standard Deviation*
Average of Minimum MSEs	0.000104837	0	0.021899809	0.017056541
Average of Final MSEs	0.000104837	3.58365E-18	0.075534031	0.061351226

Best Networks	*Training*	*Cross Validation*
Run #	1	1
Epoch #	100	77
Minimum MSE	0.000104837	0.003058134
Final MSE	0.000104837	0.045040441

Figure 4: Evaluation statistics for MNN

To evaluate the performance of the trained network, comparison between the predicted sorption efficiency of Pb(II) ion onto acid activated shale using modular neural network (MNN) and the experimental values of Pb(II) ion sorption efficiency by acid activated shale was obtained and presented in Figure 5

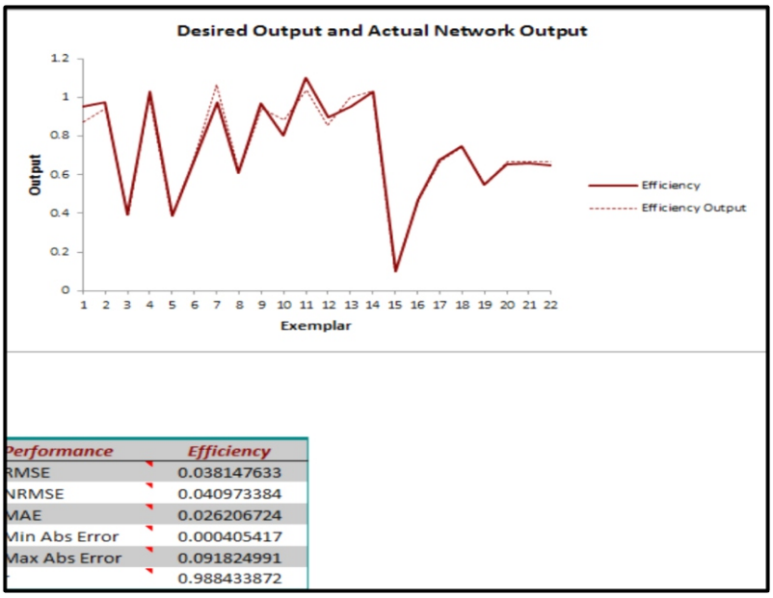

Figure 5: predicted versus observed sorption efficiency of Pb(II) ion using MNN

Result of figure 3.5 revealed a strong agreement between the experimental sorption efficiency data and modular neural network predicted data with a performance statistics of RMSE (0.03815), NRMSE (0.04097), Max.AE (0.02621), Min.AE (0.00041) and R^2 (0.988). To validate the accuracy of the trained network, a linear plot of the predicted data and the experimental data of Pb(II) ion sorption efficiency onto acid activated shale was obtained and the correlation coefficient R^2 value was employed as a bases for judgement. The linear plot is presented in Figure 6

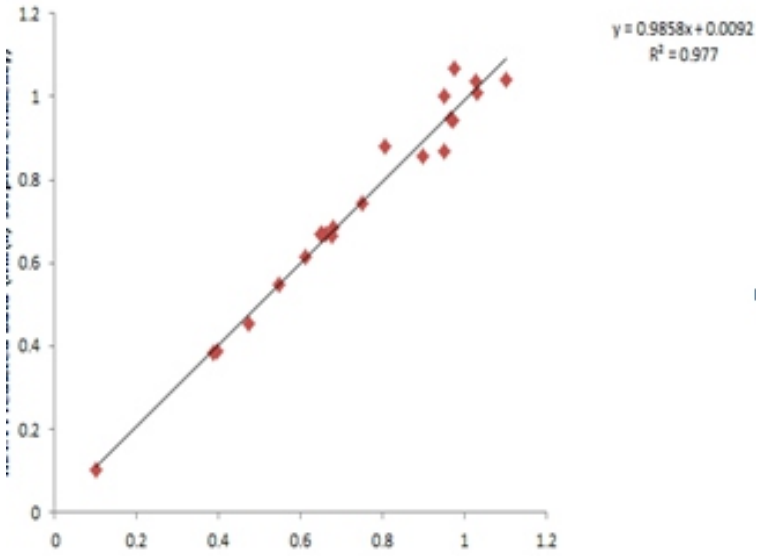

Figure 6: Accuracy of MNN prediction

Coefficient of determination of 0.977 as observed in Figure 3.6 was employed as a bases to justify the suitability of modular neural network in predicting the sorption efficiency of Pb(II) onto acid activated shale. Sensitivity analysis was employed to give insights into the relative importance of individual input parameters in other to identify those parameters with the highest significant contributions towards the sorption of Pb(II) ion onto acid activated shale. Result of the sensitivity analysis is presented in 7.

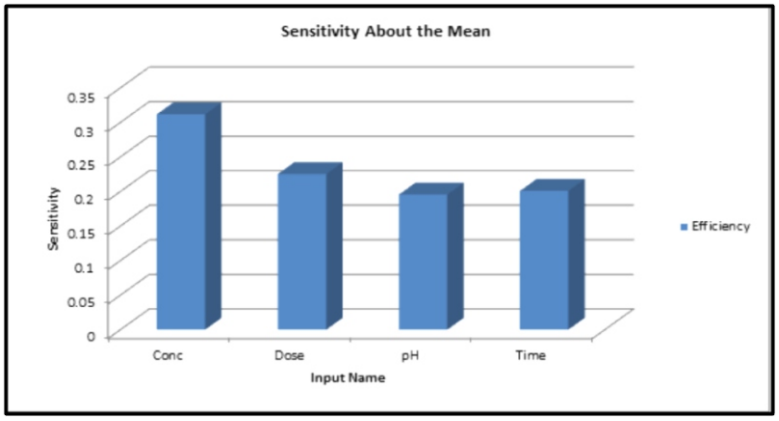

Figure 7: Sensitivity analysis of MNN

From the result of Figure 7, it was observed that initial Pb(II) ion concentration was the most important input variable affecting the overall process of adsorption. To test the ability of the network to predict the sorption efficiency of Mn(II) ion onto acid activated shale, 30 input data generated from a response surface CCD design of experiment were normalized and presented to the network as production data. Comparison between the predicted sorption efficiency of Mn (II) ion and the experimental values is presented in Figure 8.

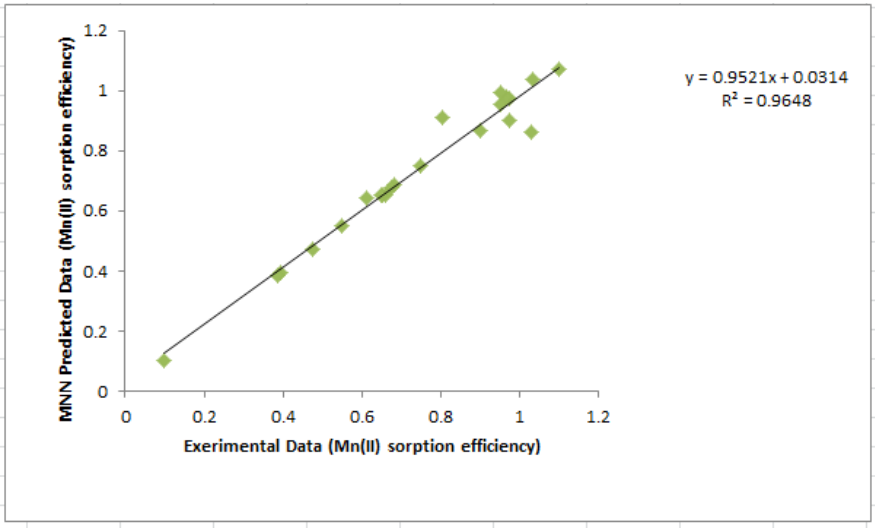

Figure : Performance of MNN on new input data

Correlation coefficient of 0.9648 as observed in Figure 8 was good enough to justify the strength and accuracy of modular neural network as a tool for modeling and prediction of sorption processes.

Conclusion

The suitability of modular neural network in modeling and prediction of metal ion adsorption onto porous solid adsorbent was investigated and found to be highly effective. Modular neural network gave a strong agreement between the experimental and predicted sorption efficiency of Pb(II) and Mn(II) ions with R^2 values of 0.977 and 0.9648 having performance statistics of RMSE (0.03815), NRMSE (0.04097), Max.AE (0.02621), Min.AE (0.00041) and R^2 (0.988).

References

Abdoliman, A; Ali, A.A and Fatemeh, A (2013), A study of cadmium removal from aqueous solution by sunflower powder and its modeling using artificial neural network (ANN), Iranian Journal of Health Sciences, vol. 1(3), pp; 28-34

Abhishek, K; Kumar, R.R; Jyoti, K.A and Shalini, S (2011), ANN modeling on prediction of biosorption efficiency of Zea Maize for the removal of Cr(III) and Cr(VI) from wastewater, International Journal of Mathe-matics Trends and Technology, July/August Issue, pp; 23-29

Animesh, D. (2013), Modular approach to big data using neural network, unpublished M.Sc. dissertation submitted to the department of computer science, San Jose State University, pp 23.56.

Badmus, M.A.O.; Audu T.O.K and Anyata, B.U, (2007), Removal of lead ion from industrial wastewaters by activated carbon prepared from periwinkle shells, Turkish Journal of Engineering and Environmental Science, pp; 251 – 263

Cerino-Cordova, F.J; Garcia-Leon, A.M; Garcia-Reyes, R.B; Garza-Gonzalez, M.T; Soto-Regalado, E; Sanchez-Gonzalez, M.N and Quezada-Lopez, I (2011), Response surface methodology for lead biosorption on Aspergillus Tesseus, International Journal of Environmental Science and Technology, vol. 8(4), pp; 695-704

Gimbert, F.; Morin-Crini, N.; Renault, F.; Badof, P.M and Crini, G., (2008), Adsorption isotherm models for dye removal by cationized starch-based materials in a single component system; Error analysis, Journal of Hazardous Materials, Vol. 157, pp 34-46

Gunay, A.; Arslankaya, E and Tosun, I., (2007), Lead removal from aqueous solution by natural and pretreated clinoptilolite; adsorption equilibrium and kinetics, Journal of Hazardous Materials, Vol. 146, pp 362-371

Hao, C and Wang, A., (2007), Kinetic and isothermal studies of lead ion adsorption onto polygorskite clay, Journal of Colloid and Interface Science, vol. 307, pp 309 – 316

Hong, Z.; Donghong, L.; Yan, Z.; Shuping, L., and Zhe, L., (2009), Sorption isotherm and kinetic modeling of aniline on Cr-bentonite, Journal of Hazardous Materials, vol. 167, pp 141 – 147

Krishna, G.B and Susmita, S.G (2006), Adsorption of chromium (VI) from water by clays, Industrial Engineering and Chemical Resource Journal, vol. 45, pp; 7232-7240

Kumar, N.K; Reddy, D.S.R and Venkateswarlu, P (2010), Application of response surface methodology for optimization of chromium biosorption from aqueous solution onto Syzigium Cumini (Java) seed powder, Microbial and Biochemical Technology, vol. 2(1), pp; 020-027

Lalvani, S.B.; Wiltowski, T., and Weston, A., (2010), Metal ions removal from wastewater by adsorption, Journal of Engineering Science, Vol. 2, pp 877-879

Lima, E.C; Betina, R; Vaghetti, J.C.P; Brasil, J.L; Nathalia, M.S; Dos-Santos Jr., A.A; Pavan, F.A; Silvio, L.P.D; Edilson, V.B and Da-Silva, E.A (2007), Adsorption of Cu(II) on Araucaria Augustifolia wastes; determination of the optimal conditions by statistical design of experiment, Journal of Hazardous Materials, vol. 140(1), pp; 211-220

Lin, S. H., and Juang, R.S., (2002), Heavy metal removal from water by sorption using surfactant-modified montmorillonite,

Journal *of Hazardous Materials, vol. B92,* pp 315 – 322

Mariadas, K.; Kalyani, G.; Joga, H.R.; Kumar, P. Y and King, P., (2012), The removal and equilibrium studies of cadmium by natural clay as adsorbent, International Journal of Scientific Engineering Research, Vol. 3(8), pp 1-6

Mohammed, Y; Farahnaz, E.B; Sepideh, K.B; Bahador, D.J; Soraya, H; Mohammad, R.M and Lugman, A. C (2013), Removal of Ni(II) from aqueous solution by an electric arc furnace slag using artificial neural network (ANN) approach, Advance in Environmental Biology, vol. 7(9), pp; 2303-2310

Ong, S.; Seng, C., and Lim, P., (2007), Kinetics of adsorption of copper (II) and cadmium (II) from aqueous solution on rice husk and modified rice husk, Electronic Journal of Environmental Agriculture and food Chemistry, Vol. 6(2), pp 1764-1774.

Sinan, M.T; Beytullah, E and Asude, A (2011), Prediction of adsorption efficiency for the removal of Ni(II) ions by zeolite using artificial neural network (ANN) approach, Fresenius Environmental Bulletin, vol. 20(12), pp; 3158-3165

Weng, C.H; Tsai, C.Z; Chu, S.H and Sharma, Y.C., (2007), Adsorption characteristics of copper (II) onto spent activated clay, Separation Purification Technology Journal, vol. 54, pp; 187-197