Asian Journal of Chemistry Vol. 22, No. 8 (2010), 5937-5950

Neural Network Modeling Used As a Chemometric Tool for Kinetic Investigations

MUSHINI VENKATA SUBBA RAO*, VELURI ANANTA RAMAM†,
VOLETY MURALIDHARA RAO† and R. SAMBA SIVA RAO†
Department of Chemistry, G.M.R. Institute of Technology, Rajam-532 127, India
E-mail: srmushini@rediffmail.com

General introduction on neural networks (NN), including its origin and chronological development are presented. The features of neural networks in terms of architecture, transfer function, training algorithm and the popular neural network packages are presented. Neural network modeling and other polynomial models applied to under stand on the effect of hydrogen ion on the oxidation of kinetic reaction of $Ru(CN)_6^{4-}$ with MnO_4^- in perchloric acid medium. The data set is analyzed using the Trajan software. The obtained residual values with neural network (NN) modeling are far less than that in the polynomial model and also the calculated k_{NN} values are better correlated with k_{obsd} values.

Key Words: Neural networks, Kinetics, Chemometrics, Polynomial model, Residual values.

INTRODUCTION

Chemometrics[1-4] comprises Chemistry, Advanced Mathematics, Statistics and Information Theory. It is an interdisciplinary area emerged in late 1970's. With the advent of chemometrics, multivariate and multi response data acquisition from statistically designed experiments, expert system inferences[5] result in reliable chemical information even in noisy environment.

Kinetometrics, a subfield of chemometrics deals with the knowledge of rates of chemical reaction and mechanistic details. The scope of chemometrics in kinetic investigations is already reported[6-8] and a few popular subfields are envirometrics[9,10], pharmacometrics, synthometrics, qualimetrics and speciometrics[11].

Neural network modeling was used for analysis of environmental samples[9,10] and for preparation of ceramics[12]. Basing on the available reported literature, the neural network (NN) modeling is used in kinetic investigations. Neural network analysis has prospects in predicting the rate constant of a reaction, from a set of k-values for related compounds. For instance, from a pool of rate constants, for the hydrolysis of different esters ($R_1CO_2R_2$), the rate constant for the hydrolysis of an ester which is not investigated, could be accurately predicted.

†School of Chemistry, Andhra University, Visakhapatnam-530 003, India.

Data reduction technique: The variation of response or its function with the magnitude of the influencing factors can be represented by a linear or non-linear model. For instance, in a linear model, $y = a_0 + a_1x$ all the data points are reproducible by the parameters a_0 and a_1. This is a data reduction technique as the number of parameters is far less than the number of data points. Variation of absorbance (Y) with concentration (Beer's law) or log k(Y) with substituent (X) (Hammet equation) is a well known linear model. The slopes and intercepts of these models bear chemical significance. In the case of non-linear models and polynomial models, the number of parameters increases with the order of the polynomial.

Neural network (NN): Introduction of hyphenated instruments which gives output substantially large amounts of data poses new challenges for analysis, graphical representation and interpretation. The extraction of knowledge from chemical data involves various computational procedures using hard or soft models. Hard modeling (linear or non-linear) techniques require knowledge of the model and noise characteristics. The hurdles associated with multivariate technique are co linearity and complexity of data including non-linearity. To surmount the difficulties of deductive methods based on strong assumptions about continuity of parameter space and error structure, natural computations with weak assumptions were proposed in the later part of the 20th century. Uncertainty in data, vagueness of information, missing data points and imprecise goals add to the complexity. Thus, there is a need for a paradigm shift from classical methods to artificial intelligence (AI) tools.

Neuro physiologists and cyberneticists introduced connectionist model to explain vision and tactile senses. The neurons in bundles organized in a very complex manner are instrumental for the functioning of a human brain. Mathematicians proposed perceptron model consisting of neurons, now called processing elements[13] (PE's). That was the beginning of the era of artificial neural networks (ANN)[14-20], which gained momentum in mid 1980's. The ANN is a reality although artificial brains are in the realm of scientific fiction because human brain is so complex that it remains yet to be understood. With 25 years of rigorous depth- and breadth-wise investigations, the word 'artificial' can safely be dropped.

Data processing with neural networks (NNs) is performed either by direct implementation on a chip or by a software. Neural network implementations in software are popular and have been successfully used in predicting stock market, forex, sunspots, onset of diabetes, distinguishing renal cell carcinoma from cyst, diagnosing acute myocardial infraction and classifying iris data.

Neural network (NN) is a data driven imbibing technology. It models multivariate, non-linear data even with discontinuous regions. This technique does not need transformation of data unlike in classical linearization techniques. Neural networks are broadly classified into self-organizing map (SOM) and multi layer perceptron (MLP). Self organising map (SOM) handles only response data in an unsupervised learning form. On the other hand, multi layer perceptron (MLP) (**Chart-1**) requires explanatory variables also and processes them in the supervised learning mode.

Neural Network	:	MLP (Multi Layer Perceptron)
	:	Feed Forward
	:	Fully Connected
Data	:	Explanatory variable (X) Response (Y)
Architecture	:	I#-H1#-H2#-0
Activation function	:	I - Linear
		H# - [Sigmoid. . . . .]
		O - Linear
Training algorithm	:	[BP, CG, Marquardt - - - -]
Output	:	YTr
		Tr.err Ve.err Te.err

Chart-1: Multi layer perceptron (MLP)

Architecture: An neutral network contains a sequence of layers. Each layer consists of a set of processing elements (PEs). The first and the last are called input and output layers and the processing elements in them correspond to the explanatory and response variables. The maximum number of intermediate layers is three. Since an end user is interested in input and output patterns and not in the intermediate processes, the latter are called hidden layers. Transformation of input to output is in the forward direction and this is referred to as feed forward neural network (FFNN)[21]. The number of PEs in different hidden layers need not be the same. With some exceptions, the usual number of PEs is less than 50 % of the data points. The processes in the hidden layers are vital in understanding the complex patterns at the output layer.

Full or random connections of PEs with different layers result in different NN models. The extent of connection between PEs in two layers is called a weight, which is analogous to synapse strength in neurobiology. They are the parameters of NN and are initially chosen from random numbers. They are refined until convergence or maximum clock time set is over.

Transfer function: Transfer function (TF) is a mathematical equation associated with the PE. The primitive transfer function is multiplication by unity *i.e.* doing nothing. Typical TFs are polynomial (linear, quadratic, cubic, *etc.*), hyperbolic (tanh, sigmoid), kernel (Guassian) and wavelet. Depending upon the nature of the TF, different nets have been proposed. For example, radial basis function NN (RBFNN)[22] and probabilistic neural network (PNN)[23] use radial and probability density functions, respectively.

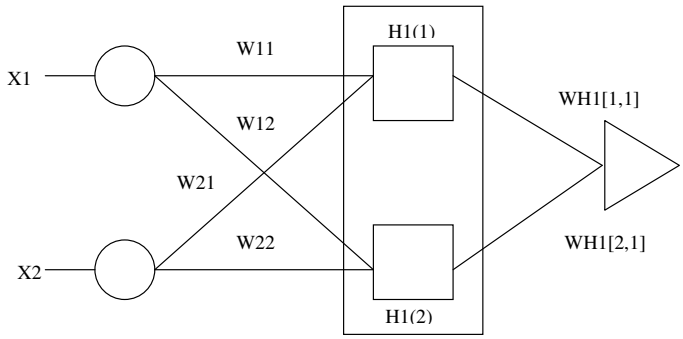
A processing element (PE), in a given hidden layer, receives information from PEs of the previous layer. The TF operates on the information and produces the result, which is passed on to the next layer. Sigmoid (SG) is a popular transfer function. Its output range is zero and +1. One of its two parameters, translation

factor (θ) shifts the entire profile in horizontal direction and the other, scale factor (α) changes the steepness. The α renders sigmoid (SG) profile to be a hard limiter and a straight line. This is responsible for the ability of NN with SGTF to model linear and non-linear profiles.

Wavelets[24,25] as transfer functions (TFs) have good zooming property *i.e.* they explore fine details in highly non-linear multi-response surfaces. Other transfer functions suggested are the binary products of tanh, arctan and sigmoid.

Training: A vector in the data matrix is a pattern. Each pattern is given to the network and the output is compared with the response. Hence, it is a supervised learning. The error function is calculated after all the patterns are presented. The widely employed optimization procedure (learning rule) in 1980's was back propagation (BP), which is a variation of steepest descent algorithm. Recently Marquardt, conjugate gradients (CG), simulated annealing algorithms (SAA), genetic algorithm (GA), *etc.*, have been incorporated in NN software.

Illustration: A three-layered NN (input-one hidden-output layer) is depicted and the formulae for outputs of the neurons are given in **Chart-2a-c**. The response is modeled as a linear combination of sigmoid functions. Non-linearity in I/O mapping increases with the second and third hidden layers. The data set is divided randomly or in a prefixed manner into training, verification and test sets.



X1, X_2-Inputs; W11, W12, W21, W22-weights; H1 (1), H1 **Chart 2a:** Neural network one hidden layer (2)-hidden neurons; WH1[1,1], WH1[2,1]-Outputs

X = [X1, X2]

IH1 = X^T*W

IH1(1) = X^T*W(:,1)=X1*W11+X2*W12

IH1(2) = X^T*W(:,2)=X1*W12+X2*W22

OH1 = $(IH1)^T$*WH

$$\begin{bmatrix} WH1(1,1) \\ WH1(2,1) \end{bmatrix}$$

= [IH1(1) IH1(2)]*

= IH1(1)*WHI(1,1)+IH1(2)*WH1(2,1)

$$= (X1*W11+X2*W12)*WH1(1,1) + (X1*W12+X2*W21)*WH1(2,1)$$

$$OH1 = SG(X^T*W)$$

$$OH1(1) = \frac{1}{1+EXP\{(-X1*W11-X2*W12)*\theta)\}}$$

$$OH1(2) = \frac{1}{1+EXP\{(-X1*W12-X2*W22)*\theta)\}}$$

$$OH1 = \frac{ONE}{ONE+EXP(-X^T*W*\theta)}$$

$$Output = OH1(1)*WH1(1,1)+OH1(2)+WH1(2)$$

$$= \frac{WH1(1,1)}{1+EXP(-X1*W11-X2*W12} + \frac{WH1(2,2)}{1+EXP(-X1*W12-X2*W22)*\theta}$$

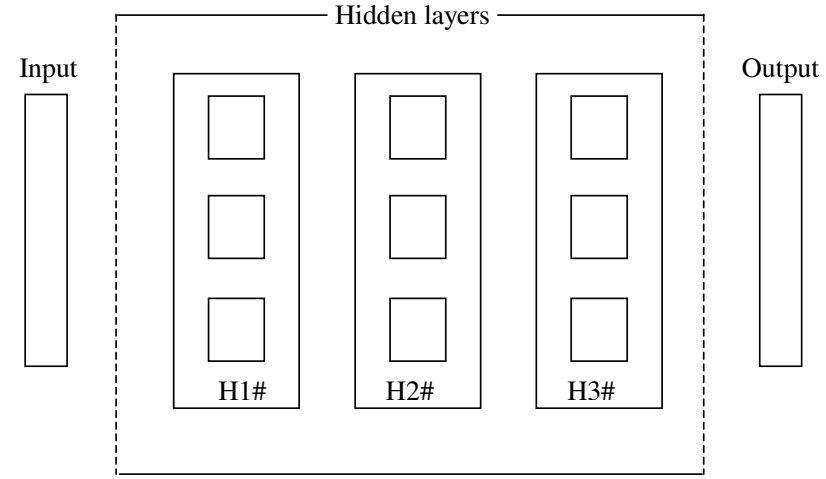
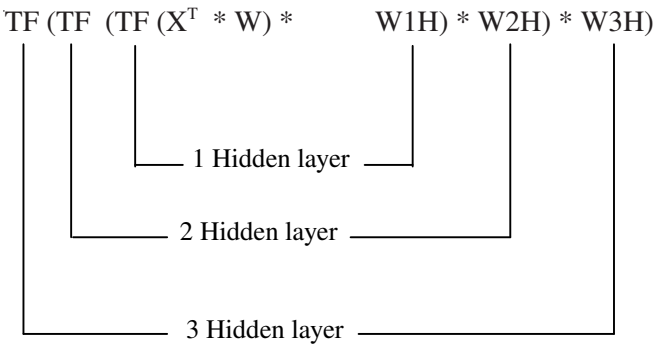


Chart-2b: Neural network-with three hidden layers

The best architecture is chosen by changing the number of hidden layers (1, 2 or 3), hidden neurons in each layer (1 to 50 % of the number of data points), transfer function and learning algorithm. Experimental design has prospects in this activity. A quick heuristic is to propose the best set of models with a minimum training error (Tr.err), verification error (Ve.err) and test error (Te.err). Advanced residuals in the measurement scale and influential statistics have also been handled in some packages by an intelligent module. There are more than 80 commercial packages with unlimited capabilities and a few typical ones are described in **Chart-3**.



W: Scaling or pre-processing; W1H, W2H and W3H: Weights from 1 to H1, H1 to H2 and H2 to H3; TF: Transfer function like sigmoid, hyperbolic and radial basis.

Chart-2c: Neural Network-input to out put mapping in MLP

During the past two decades this technique has been in wide use in chemistry for quantitation, calibration, speciation, parameterization and pattern recognition (PR). In spectroscopic techniques NN is applied in classification[26] and peak identification[27] of chemical compounds. It was used for simultaneous determination of two and four component systems by ion selective electrode[28] (non-Nernstian) data, chemical kinetics and kinetic methods of analysis.

EXPERIMENTAL

Hardware and software: An IBM Pentium II computer is employed and MATLAB (version 4.2c.1 for windows environment) from Math Works Inc. is employed to develop programs and generate graphic outputs. Regression analysis with polynomials of different order was performed within - house programs developed in the laboratory (LINREG, POLREG). Software package Trajan 4.0 is used for neural network analysis. This package has an additional advantage of providing intelligent problem server (IPS) output.

RESULTS AND DISCUSSION

Applications of neural networks in a kinetic reaction: The neural network modeling and other polynomial models has been applied to the data set of literature report[29] reaction of kinetic study of the MnO_4^{2-}–$Ru(CN)_6^{4-}$ in perchlorate acid media. Neural Network modeling and other polynomial models applied to the above reaction to understand the effect of hydrogen ion on the oxidation of $Ru(CN)_6^{4-}$ with MnO_4^{2-}. The study of the MnO_4^{2-}–$Ru(CN)_6^{4-}$ reaction was investigated[29] by stopped flow technique and the ionic strength was maintained with perchlorate acid media (μ = 1.02). The study showed a "bell like" dependence of k_{obs} on log $[H^+]$ in one region and independent of log $[H^+]^{-1}$ in another. Neural network modeling and other

Package	Type of network	Algorithm	Salient features
Trajan	Kohonen	Levenburg Marquardt Conjugate Gradient BP Quick Propagation Delta-Bar-Delta, SVD K-Means, K-Nearest, Weigend Weight	PCA, Neuro-Genetic Input Selection, Automatic Network Design
Neural Works	ART 1 Kohonen Modular NN General Regression Fuzzy ART Probabilistic NN LVQ Boltzmann BSB SPR	BP	Neural Works Explorer Developers Package Neural Works Professional
Direct Executing Simulation in Real time (DESIRE NEUNET)		Fuzzy logic	Dynamic-system simulation (6,000 differential equations, 20,000 difference equations, 13 integration rules) for control, aerospace, chemical engineering, physiological modeling, ecolosy
Neuroshell	Ward Nets Recurrent Kohonen, Probabilisti c NN General	BP	Dupont in making safety glass Texaco in process control in oil refineries
STATISTIC A: neutral Networks	Kohonen PNN GRNN	BP Levenburg Marquardt Conjugate Gradient Quick Propagation Delts-Bar-Delta, Linear SVD K-Means, K-Nearest, Neighbour Isotropic Deviation	Genetic Input Selection, Automatic Network Designer
Matlab; :Neural Network Toolbox	Perceptron RBF Hopfield LVQ Competitive NN Kohonen Elman Hebb	BP Levenberg-Marquard	MATLAB Signal Processing Nonlinear control Financial modeling

Chart-3: Features of Neural Network packages

TABLE-1
THE k_{obsd} VALUES[29] (AT 25 ºC) AS A FUNCTION OF HYDROGEN ION CONCENTRATION AND CALCULATED RATE CONSTANTS (k_{NN}) AND RESIDUAL VALUES BY NN MODEL

S. No.	$[H^+]$ M	k_{obsd} (M^{-1} S^{-1})	Calculated rate constant by NN (k_{NN})	Residual value by NN
1.	0.008	78.0	77.674	0.30914
2.	0.013	42.4	44.918	-2.5137
3.	0.015	57.2	63.391	-6.1879
4.	0.017	69.8	80.357	-10.554
5.	0.016	97.2	76.415	20.79
6.	0.019	98.0	93.964	4.0337
7.	0.022	111.0	108.80	2.204
8.	0.024	109.0	116.01	-7.0128
9.	0.026	118.0	121.80	-3.8047
10.	0.029	125.0	128.68	-3.6814
11.	0.033	142.0	135.63	6.3652
12.	0.037	144.0	140.85	3.1482
13.	0.040	133.0	143.87	-10.866
14.	0.042	148.0	145.50	2.502
15.	0.050	158.0	149.41	8.5932
16.	0.061	153.0	149.44	3.5613
17.	0.070	146.0	146.52	-0.51869
18.	0.081	127.0	141.19	-14.187
19.	0.100	135.0	130.96	4.0383
20.	0.122	127.0	120.41	6.5887
21.	0.130	117.0	117.06	-0.060395
22.	0.160	106.0	106.30	-0.3005
23.	0.200	89.6	94.716	-5.1157
24.	0.240	87.5	84.996	2.5039
25.	0.281	75.5	76.722	-1.2223
26.	0.320	74.8	70.485	4.3152
27.	0.340	64.5	67.889	-3.3887
28.	0.360	67.4	65.672	1.7281
29.	0.460	54.7	58.704	-4.0044
30.	0.480	69.2	57.638	11.562
31.	0.500	40.4	56.444	-16.044
32.	0.560	65.3	51.233	14.067
33.	0.620	37.8	44.268	-6.4682
34.	0.640	42.5	42.144	0.35574
35.	0.660	37.5	40.291	-2.7913
36.	0.700	37.7	37.478	0.2218
37.	0.720	35.7	36.477	-0.7775
38.	0.740	33.6	35.696	-2.0963
39.	0.780	37.1	34.633	2.4672
40.	0.900	35.7	33.466	2.2337

polynomial models are applied to the data of this reaction for calculating of the rate constant (k_{NN}) and the residual values (residual 1, 2.3, 4 and 8) of a reaction.

The calculated rate constants (k_{NN}) and the residual values by using NN modeling are presented in the Table-1. The reported data of the rate constants (k_{obsd}) as a function of hydrogen ion concentration are also given[29] in the Table-1.

A graph (Fig. 1) is drawn in between the k_{obsd} values and hydrogen ion concentration. This scatter plot indicates the complex dependence of k_{obs} on H^+. Another graph is drawn (Fig. 2) in between the log k_{obsd} and log $[H^+]$, it also yields a peculiar behaviour. In view of this, the rate constants have been evaluated by using linear, quadratic, cubic models and with higher order polynomials. In the Fig. 3 a plot is drawn in between the residual 1 and H^+ concentration and also another plot is drawn in between the k_{obsd} values against concentration of $[H^+]$ in the same Figure. The top plot of the figure shows k_{obsd} values as a function of $[H^+]$. That is in the graph, the bottom plot represents the residuals in k_{obsd} when a linear model is considered with $[H^+]$. From that it is seen that the residuals (1) are quite large, showing that the model fails in totality.

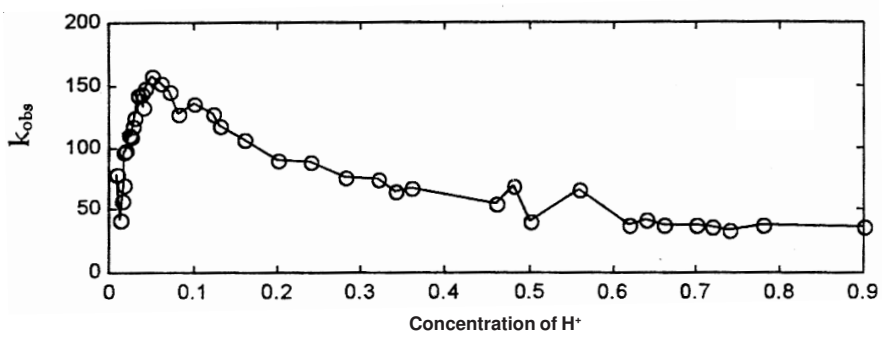


Fig. 1. Plot of k_{obsd} values *vs.* H^+ concentration

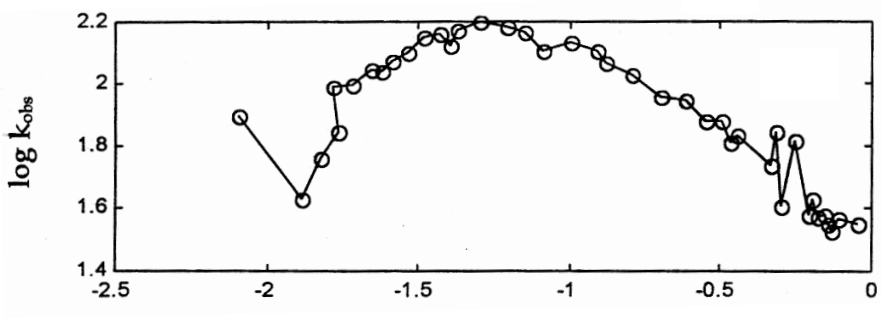


Fig. 2. Plot of log k_{obsd} *vs.* log $[H^+]$

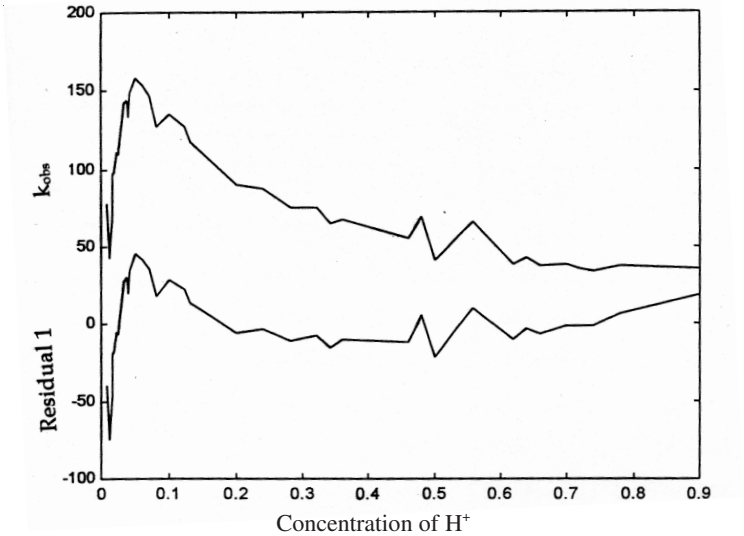


Fig. 3. Plot of residual (1) *vs.* concentration of H^+ and plot of k_{obsd} *vs.* concentration of H^+

Assuming a quadratic model and higher order polynomial models (up to 4th order), the residuals have been calculated and are represented in the Fig. 4. This plot also reveals that the model is inadequate as residuals are quite significant compared to the k_{obsd} values. This is further supported by the fact that the standard deviation in parameters is much higher than the parameters in Table-2.

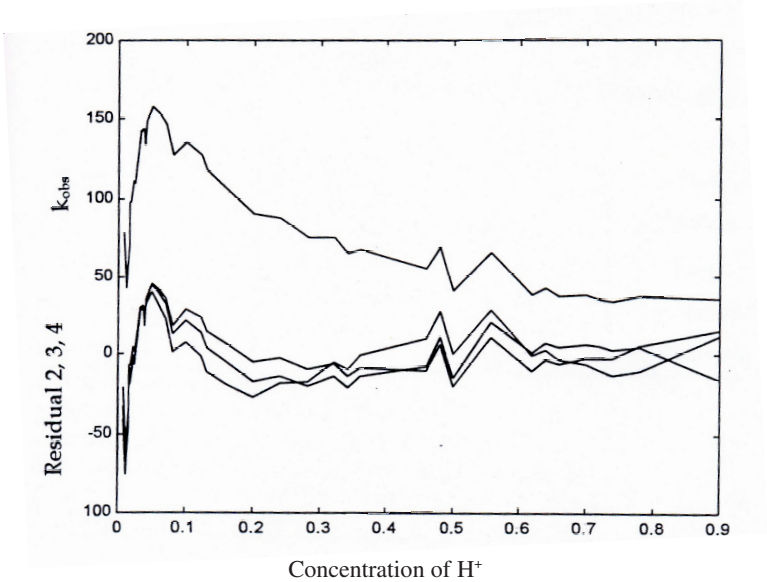


Fig. 4. Plot of residual values 2, 3, 4 *vs.* concentration of H^+ and plot of k_{obsd} *vs.* concentration of H^+

TABLE-2
STANDARD DEVIATION IN PARAMETERS

	Polynomial model				
	1	2	3	4	8
a_0	118.06	118.95	109.52	93.38	8.9
(Sda_0)	148.26	183.52	201.64	204.46	135.9
a_1	112.49	-125.22	112.30	687.76	5969.5
(Sda_1)	388.39	1516.1	314.06	5029.60	8752.3
a_2		17.057	752.56	-4050.90	-88337.0
(Sda_2)		1959.3	9306.10	25941.00	1.6864e+005
a_3			61.73	6627.10	5.9112e+005
(Sda_3)			7321.42	45590.20	1.4376e+006
a_4				-3411.20	-2.1702e+006
(Sda_4)				25653.00	6.369e+006
a_5					4.6318e+006
(Sda_5)					1.569e+007
a_6					-5.7176e+006
(Sda_6)					2.1625e+007
a_7					3.7789e+006
(Sda_7)					1.5576e+007
a_8					-1.0331e+006
(Sda_8)					4.5592e+006

Model 1 : $k = a_0 + a_1[H^+]$
Model 2 : $k = a_0 + a_1[H^+] + a_2[H^+]^2$
Model 3 : $k = a_0 + a_1[H^+] + a_2[H^+]^2 + a_3[H^+]^3$
Model 4 : $k = a_0 + a_1[H^+] + a_2[H^+]^2 + a_3[H^+]^3 + a_4[H^+]^4$
Model 8 : $k = a_0 + a_1[H^+] + a_2[H^+]^2 + a_3[H^+]^3 + a_4[H^+]^4 + a_5[H^+]^5 + a_6[H^+]^6 + a_7[H^+]^7 + a_8[H^+]^8$

Modeling with 8th order polynomial is also attempted (Fig. 5), comparing the residual values 4 and 8; and this also failed, suggesting inadequacy of polynomial model for variation of k_{obsd} with H^+.

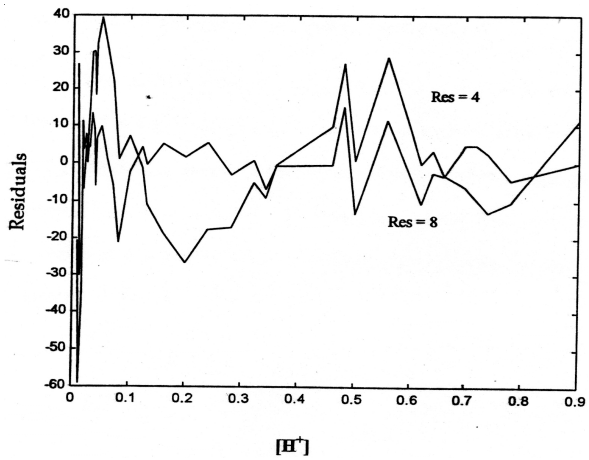


Fig. 5. Plot of residual values of 4, 8 *vs.* $[H^+]$

Using quadratic model and higher order polynomial models (up to 8th order), the residuals have been calculated. In all these cases these models are inadequate and as the residuals are quite significant compared to the k_{obsd} values. So the NN modeling is applied to the reported literature data set for calculating the rate constant (k_{NN}) as well as the residual values (Table-1). The data set is analyzed using the Trajan software. The output of the data set is presented in Fig. 6. This Figure shows data, the training error (Tr.er), network architecture, training algorithm in tiled windows.

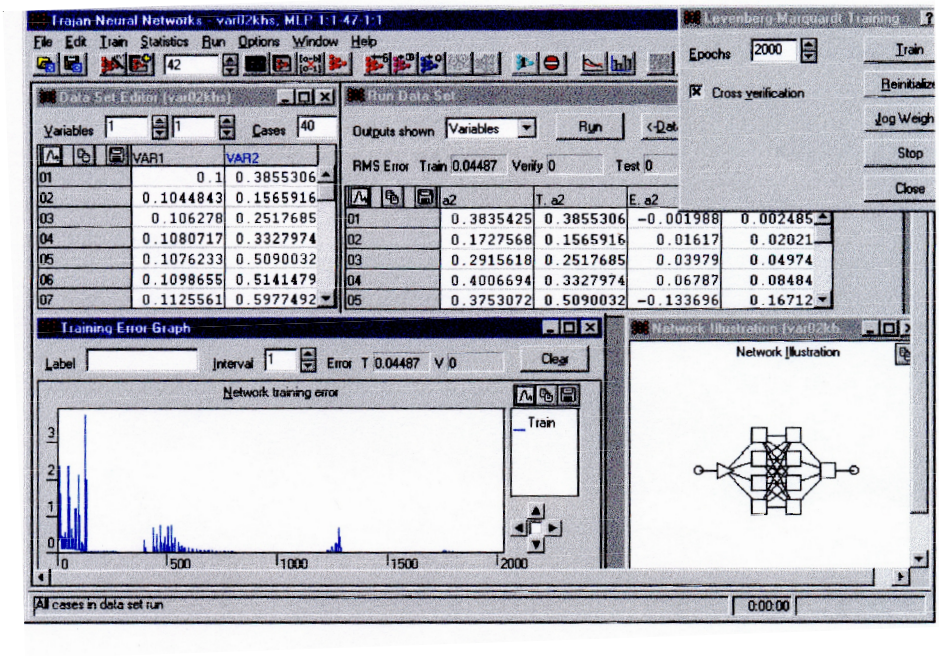


Fig. 6. Trajan output

The Multi Layer Perceptron (MLP) architecture (1-4-4-1) gave a training error of 0.04487 for 2000 epochs. The training algorithm used is Leven Berg Marquardt algorithm (LM).The model parameters of 1-4-4-1 NN (Table-2) are pictorially displayed in Fig. 7.

A comparison of residuals obtained with NN model and that of polynomial order 4, shows that the residuals of NN model are very small.

The rate constant values are calculated by using 4th order polynomial and neural network model. Further an attempt is made, these values are compared to the reported values (k_{obs}). The k_{obs} and k calculated values are correlated better in NN model compared to that obtained with a 4th order polynomial model (Fig. 8). So NN modeling values correlated better with that of reported values.

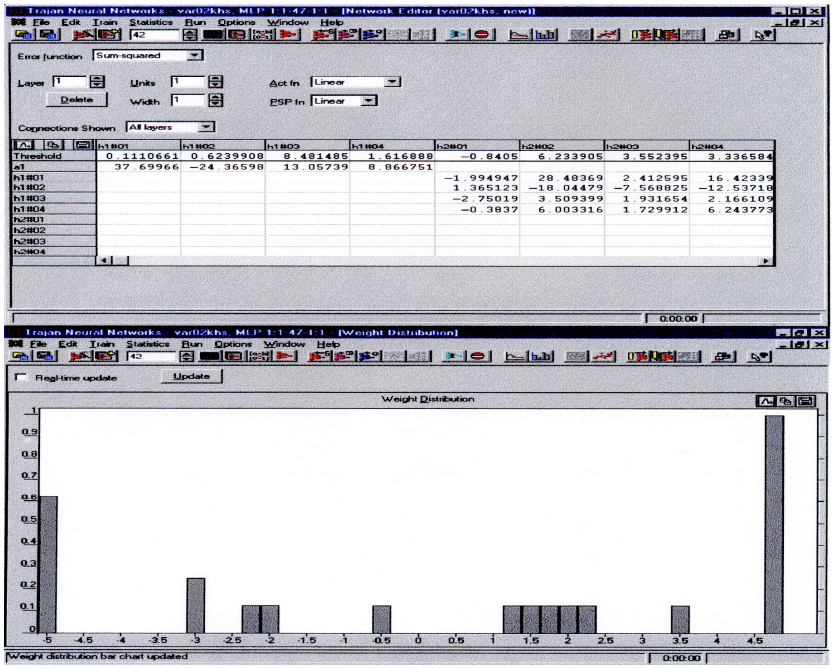


Fig. 7. Weight distribution

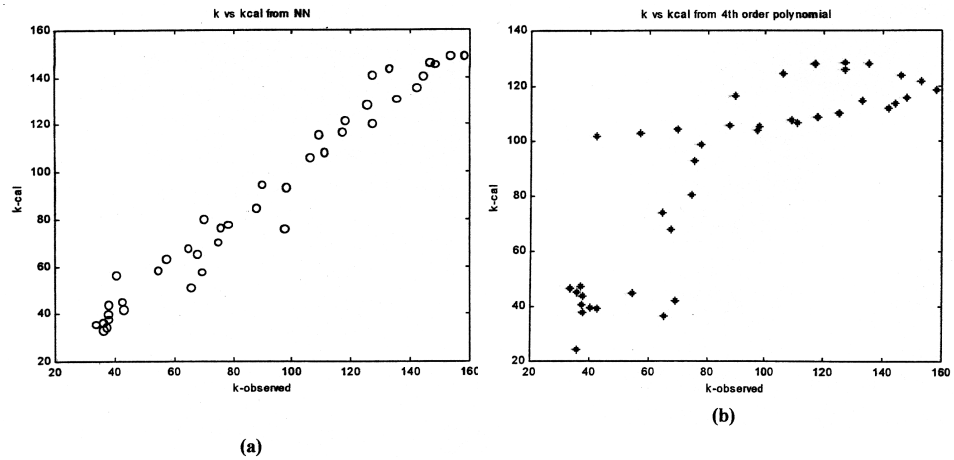


Fig. 8. Comparison of k_{obsd} *vs.* k_{cal} by NN model (a) and comparison of k_{obsd} *vs.* 4th order polynomial (b)

Finally, it is observed that the residuals with NN modeling are far less than that in the polynomial model. The residuals of the k_{obsd} with NN model (Table-1) are very small indicating the capability of NN model highly complex non-linear behaviour. So these results show the adequacy of the NN modeling in kinetic investigation.

REFERENCES

1. B.R. Kowalski, Chemo metrics: Theory and Application, ACS Symposium Series, American Chemical Society, USA (1977).
2. (a) B.K. Lavine, *Anal. Chem.*, **74**, 2763 (2002); (b) B.K. Lavine, C.E. Davidson and D.J. Westover, *J. Chem. Inf. Comp. Sci.*, **44**, 1056 (2004); (c) B.K. Lavine and J.R. Workmann, *Anal. Chem.*, **76**, 3365 (2004); (d) B.K. Lavine, *Anal. Chem.*, **72**, 91R (2000); (e) B.K. Lavine, *Anal. Chem.*, **70**, 209R (1998); (f). S.D. Brown, S.T. Sum and F. Despagne, *Anal. Chem.*, **68**, 21R (1996); (g) S.D. Brown, T.B. Blank, S.T. Sum and L.G. Weyerl, *Anal. Chem.*, **66**, 315R (1994); (h) S.D. Brown, R.S. Bear and T.B. Blank, *Anal. Chem.*, **64**, 22R (1992); (i) S.D. Brown, *Anal. Chem.*, **62**, 84R (1990); (j) S.D. Brown, T.Q. Baker, R.J. Larivee, S.L. Monfre and H.R. Wilk, *Anal. Chem.*, **60**, 252R (1988); (k) L.S. Ramos, K.R. Beebe, W.P. Carey, E.M. Sanchez, B.C. Erickson, B.E. Wilson, L.E. Wangen and B.R. Kowalski, *Anal. Chem.*, **58**, 294R (1986); (l) M.F. Delaney, *Anal. Chem.*, **56**, 261R (1984); (m) I.E. Frnak and B.R. Kowalski, *Anal. Chem.*, **54**, 232R (1982); (n) B.R. Kowalski, *Anal. Chem.*, **52**, 112R (1980).
3. Windows on chemo metrics (monthly), The Royal Society of Chemistry, (Milton road, Cambridge, CB4 4WF, U.K). E-mail: MARKETING @ RSC.ORG (Internet).
4. G.N. Rao, K.S. Rao, R.S. Rao and L.H. Abdel Rehman, *J. Indian Coun. Chem.*, **17**, 48 (2000).
5. G.P. Brunda, G.N. Rao and R.S. Rao, *J.T.R. Chem.*, **4**, 29 (1997).
6. G.N. Rao, V.A. Ramam, S.V.V.S. Rao and R.S. Rao, *J. Indian Chem. Soc.*, **75**, 236 (1998).
7. V.A. Ramam, G.N. Rao, S.V.R. Sastry and R.S. Rao, *Indian J. Chem.*, **36A**, 964 (1997).
8. G.N. Rao, V.A. Ramam and R.S. Rao, *Bull. Soc. Kinet. (India)*, **19**, 1 (1997).
9. A. Braibanti, G.N. Rao, J.S. Babu, D. Sudarsan and R.S. Rao, *Ann. Chim.*, **90**, 1 (2000).
10. V.V.P. Rao, V.A. Ramam, G.N. Rao and R.S. Rao, Proceedings of National Seminar on Environmental Geology & Waste Management, Department of Geology, Andhra University, Visakhapatnam, p. 125 (August 2000).
11. A.R. Babu, D. Muralikrishna and R.S. Rao, *Atlanta*, **40**, 1873 (1993).
12. R.S. Rao and K.S. Rao, *J. Korean Phys. Soc.*, **32**, S1850 (1998).
13. J. Zupan and J. Gasteinger, *Anal. Chim Acta*, **248**, 1 (1991).
14. C.T. Leonde, Neural Network Systems Techniques and Applications, Academic Press, New York, Vol. 1 to 7 (1998).
15. D.T. Pham and L. Xing, Neural Networks for Identification, Prediction and Control, Springer Verlag, Berlin, Heidelberg (1995).
16. J.W. Kay and D.M. Titterington, Statistics and Neural Networks, Advances at the Interface, Oxford University Press (1999).
17. W.S. McCulloch and W. Pitts, *Bull. Math. Biophys.*, **5**, 115 (1945).
18. H.S. Stern, *J. Am. Stat. Assoc.*, **38**, 205 (1996).
19. B. Cheng and D.M. Titterington, *Stat. Sci.*, **9**, 2 (1994).
20. S. Raudys, Statistics and Neural Classifiers, An Integrated Approach to Design, Springer (2001).
21. T.A. Plate, J. Bert, J. Grace and P. Bond, *Neural Comput.*, **12**, 1353 (2000).
22. J. Moody and C.J. Darkin, *Neural Comput.*, **1**, 281 (1989).
23. D.F. Specht, Probabilistic Neural Networks, Joint Conference on Neural Networks, Singapore, November (1990).
24. Ch. Cai and P.B. Harington, *J. Chem. Inf. Comput. Sci.*, **38**, 1161 (1998).
25 Y.Y. Tang, L.H. Yang, J. Liu and H. Ma, Wavelet Theory and Its Application to Pattern Recognition, Word Scientific New-Jersey (2000).
26. (a) E.W. Robb and M.E. Munk, *Mikrochim. Acta*, **1**, 31 (1990); (b) E.W. Robb and M.E. Munk, *Mikrochim. Acta*, **2**, 505 (1991).
27. E.J. Wythoff, S.P. Levine and S.A. Jomellini, *Anal. Chem.*, **62**, 2702 (1990).
28. M. Bos and W.E. Vandel linden, *Anal. Chim Acta*, **233**, 31 (1990).
29. K.W. Hicks and G.A. Chappelle, *Inorg. Chem.*, **19**, 1623 (1980).

(*Received*: 3 September 2009; *Accepted*: 30 April 2010) AJC-8644