Volume : 3 | Issue : 2 | Feb 2014 • ISSN No 2277 - 8160

Research Paper | Physics

On the Use of Function Approximation Potential of Artificial Neural Networks for Predicting Elastic Moduli of Binary Oxide Glass Systems

Dr.R.Sheelarani Vel Tech (Owned By R.S Trust), Avadi, Chennai – 600 062 * correspondent author

Dr. K.T. Arulmozhi DDE, Physics Wing, Annamalai University

ABSTRACT

Function approximation (non-linear regression) is one important application of artificial neural networks (ANNs). Since the elastic moduli of multicomponent materials are dependent on the properties of the constituents in a non-linear way, the ANN can be used to predict the elastic moduli of these systems. In this work a feed forward multiplayer perceptron (FFMLP) has been described to predict elastic moduli of binary oxide glasses using some of the characteristics of the component oxides as predictor variables.

KEYWORDS : Artificial neural network; Non-linear regression; Elastic moduli; Binary oxide glasses

1. Introduction

Elastic, physical and electrical properties are the quality determining factors of multicomponent glass systems. All these properties depend on the characteristics and compositions of the components used to form the system. But this dependency is highly non-linear as well as very much complicated. There are several experimental techniques to determine the elastic moduli of glasses. On the theoretical side there do not exist any formula-like relation to estimate the elastic moduli of polycomponent glass systems. The theories proposed by Bridge et al. and Makishima and Mackenzie to evaluate the elastic moduli of poly-component oxide glasses showed considerable to large deviations in the predicted elastic moduli from those obtained experimentally. In such situations the ANN, as a non-linear regression modelling tool, can be used to predict the elastic moduli utilizing the characteristic parameters of the oxide components as predictor variables. An illustrative list of some earlier works on the use of function approximation ability of ANN in different problems are given [1–9].

2. Methodology

One experimental approach to determine the elastic moduli of glasses is to measure the longitudinal (VL) and shear (VS) ultrasonic velocities. Then using the following relations the longitudinal elastic modulus (L) and shear modulus (S) can be calculated.

$L = rV_L$ … (1)

$S = rV_S$ … (2)

where r is the density of the glass. Here it can be observed that the two moduli are directly dependent on the density (r) of the glass.

The 'bond compression model' was proposed by Bridge et al. [10] for theoretical evaluation of the elastic moduli of oxide glass. The following three parameters used in their model are considered in the present work. They are given by the relations,

number of bonds per unit volume $n_b = \frac{N_a \rho}{M} \sum_i x_i (n_f)_i$ … (3)

average cross-link density $\bar{n}_c = \frac{\sum_i x_i (n_c)_i (N_c)_i}{\sum_i x_i (n_c)_i}$ … (4)

average stretching force constant $F = \frac{\sum_i x_i (n_f)_i f_i}{\sum_i x_i (n_f)_i}$ … (5)

where x is the mole fraction, Na is Avagadro's number, r is density and M is molecular weight of glass, nf is the coordination number, nc is the number of cross-links per cation, Nc is the number of cations per glass formula unit, f is the stretching force constant of the oxide bond and i represents ith oxide. Makishima and Mackenzie [11,12] proposed the semi-empirical model to compute the elastic moduli of polycomponent oxide glasses. The two parameters used in their model are utilized presently. They are

Packing density $VT = \frac{\rho}{M} \sum_i V_i x_i$ … (6)

and Dissociation energy $GT = \sum_i G_i x_i$ … (7)

where x is the mole fraction, Vi is the packing factor of the oxide component and Gi is the dissociation energy of the oxide bond and i represents the ith oxide.

According to Higazy and Bridge [13] the fractional ionic character (FIC) of an oxide is component is important in deciding the network forming capability in the glass system. The FIC of the glass system itself can be computed using the relation

$$FIC_{glass} = \frac{\sum x_i (n_f)_i FIC_i}{\sum x_i (n_f)_i} \quad \text{… (8)}$$

where FIC_i is the fractional ionic character of the ith oxide which can be calculated using Pauling's relation.

All the seven characteristic parameters of the glass system namely r, n_b, n_c, F, V_T, G_T and FIC are used as predictor variables in this study. They are applied as inputs to the ANN. The longitudinal modulus (L) and shear modulus (S) are the outputs of the ANN. Using these two elastic moduli values, the other two elastic moduli, bulk (K) and Young's (E) can be computed using the conventional relations.

3. Network architecture

3.1. Design basics

An artificial neural network is analogous to the biological nervous system consisting of a large number of information processing elements. These elements are called nodes or neurons. The most commonly used ANN is the multiplayer perceptron (MLP). It consists of atleast three layers, input layer, output layer and in between them a hidden layer. As a rule, the structure of the ANN is layered comprising of three functional groups. The neurons in the input layer receive signals from outside world and pass on to the hidden layer. The neurons in the hidden layer process the information and the neurons of the output layer generate the results. The number of neurons in the input layer equals the number of predictor variables applied as inputs. Similarly the number of neurons in the output layer equals the output parameters to be predicted. There is no science for the number of hidden layers or for the neurons in them. As generally stated a single hidden layer will be sufficient for most of the work. However the number of hidden neurons can be found only by trial.

3.2. Hidden neurons

The entire job of processing the information is done by the neurons of hidden layer. Two important factors decide the size of the network. First one is 'approximation' which is the ability of the ANN to approximate a given set of input-output training pairs to any arbitrary level of accuracy. In other words it is the flexibility of the network to learn from the training examples. Approximation requires good number of neurons and thus too few neurons cannot learn well. The second one is 'generalization' which means the network performance will persist for data unseen earlier. A good generalization results in well predicted outputs and requires sufficient number of neurons. Too many neurons leads to poor generalization. Thus the number of hidden neurons must be optimal to ensure both good approximation and generalization.

In the present work a three layer MLP network in feed forward configuration is used. Back propagation technique is used to adjust the weights and biases during learning process. The architecture of the designed ANN is shown in Fig. 1.

3.3. Network specifications

Inputs : Seven predictor variables, r, nb, $\overline{n}_c$, F, VT, GT, FIC
Outputs : Two elastic moudli L, S
Hidden layer : One
Training error goal : 1e–003
Training methods : trainscg (18), trainoss (14), trainlm (10), traincgb (16) and number of hidden neurons Transfer functions
: tansig (input, hidden layers), purelin (output layer)

4. Network training

4.1. Data sets

Once the architecture of the ANN is finalized, the next step is to construct a comprehensive database. The success of the ANN depends, to a large extent, on the information contained in the database. The data points should represent a variety of input-output pattern, collected from experimental data, so as to teach the network all the relationships between the predictor variables and the target values. Further the number of input and output parameters should be optimal. Too many input and output parameters will slow down the learning process. There exists different methodologies to divide the data base into sub groups for training, validation and testing. In the present study 85 data points are used for training and 22 data points for testing. The binary oxide glass systems used in the test data and their characteristic parameters used as predictor variables are presented in Table 1. The range of values of input and output parameters used in both training and testing sets are given in Table 2. Since this range differs for each parameter with different magnitudes, the pre-processing (normalization) is carried out. The network performs well on these normalized values. After training post-processing is done to get back the original format of data.

4.2. Training methods

Eighteen back propagation training algorithms are available in the neural network toolbox [14]. Four of these namely, scaled conjugate gradient back propagation (trainscg), one step secant back propagation (trainoss), Levenberg-Marquardt back propagation (trainlm) and conjugate gradient back propagation (traincgb) are used in this study. A comparative study of these functions are made.

4.3. Epochs

One epoch refers to the presentation of entire data to the network. The number of epochs affects the performance of the network. This number depends on the number of training data, number of hidden layers and neurons in them, number of input and output parameters, the training method used and most importantly on the error goal set. When the training is carried out for few epochs the network will not learn all the features of the input-output pattern. For too long training the network overfits the data and looses generalization ability leading to poor predictions for unseen data. Thus the number of epochs is problem specific and should be kept at optimum value. Convergence to goal can be quickly achieved if the error goal is fixed at just low enough value. One has to kept in mind that reduction of training error to a very low value does not guarantee good generalization. In this work an error goal set at $1e^{-002}$ is used which yielded good predictions of outputs.

5. Network performance

The completion of network training is indicated by the convergence of training error to the pre-fixed goal. The predicted outputs show reasonable to high similarities to the target output. If the test data, whose input pattern is unseen earlier, is now presented, the network predicts the outputs namely the longitudinal and shear moduli.

To evaluate the function approximation ability of the designed ANN, the accuracy of the predicted values of elastic moduli are found. It is measured by calculating the mean absolute percentage error (MAPE). It is given by the relation

$$\text{MAPE} = \frac{1}{n}\left[\sum_i \frac{|o_i - t_i|}{o_i} \times 100\right] \quad \ldots \quad (9)$$

where t is the predicted value by ANN and o is the experimentally observed value (taken from earlier works) and n is the number of data points used in the test set. Further the root mean squared (RMS) value and absolute fraction of variance (R^2) are also calculated using the following relations.

$$\text{RMS} = \left[\frac{1}{n}\sum |t_i - o_i|^2\right]^{1/2} \quad \ldots (10)$$

$$R2 = 1 - \left[\frac{\sum |t_i - o_i|^2}{\sum t_i^2}\right] \quad \ldots (11)$$

Suitable program is written using neural network toolbox in MATLAB versions 7.5 [11] environment to present the training data, pre-processing and post-processing and finally the test data. For close monitoring and evaluation of the performance of the network, statements are included in the program to show the scatter plots between the experimental and predicted values of elastic moduli. Further, for precise evaluation of the results provisions are included in the program to show the MAPE and R^2 values at the end of each run.

6. Results and discussion

The results obtained with all the four training algorithms are presented in Table 3. The experimental and ANN predicted values of elastic moduli given in this table, are also shown as scatter plots in Fig. 2(a–h). A survey of these results suggests that all the four methods yielded more or less good predictions. However, as seen from Table 4 the 'trainscg' algorithm showed relatively good generalizations ability with very low values of MAPE (mean) = 6.5916, RMS (mean) = 2.6839 and a high value of R^2 (mean) = 0.9947. The other three training methods exhibit almost equal generalization abilities.

7. Conclusion

The function approximation potential of the ANN has been successfully utilized for the prediction of elastic moduli of binary oxide glass systems. The characteristic parameters of the oxide components served well as predictor variables. All the four training algorithms used in this work showed almost similar performances. However one suggestion to this application of ANN is that by incorporating more input-output data patterns in the training set covering entire range of oxide compositions this neural approach will become a powerful tool for the predictions of accurate values of elastic moduli.

Captions for figures

Fig. 1. A three layer feed forward ANN used in the present work

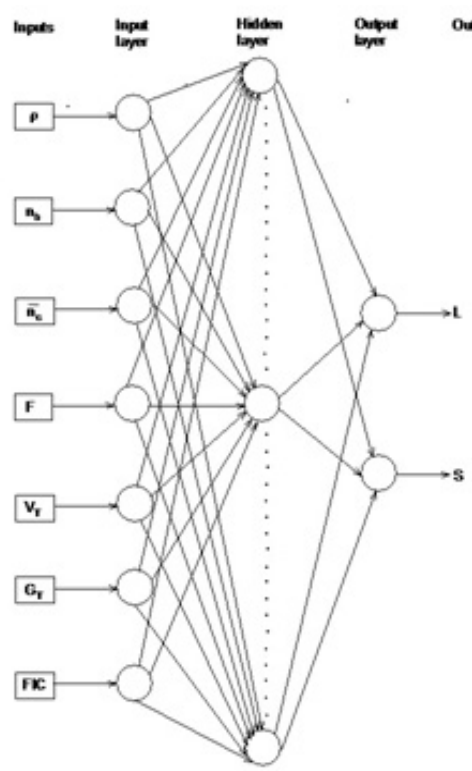

Fig. 2. Linear relationship between a) experimental and predicted L (trainscg), b) experimental and predicted L (traincgb), c) experimental and predicted L (trainoss), d) experimental and predicted L (trainlm), e) experimental and predicted S (trainscg), f) experimental and predicted S (traincgb), g) experimental and predicted S (trainoss) and h) experimental and predicted S (trainlm).

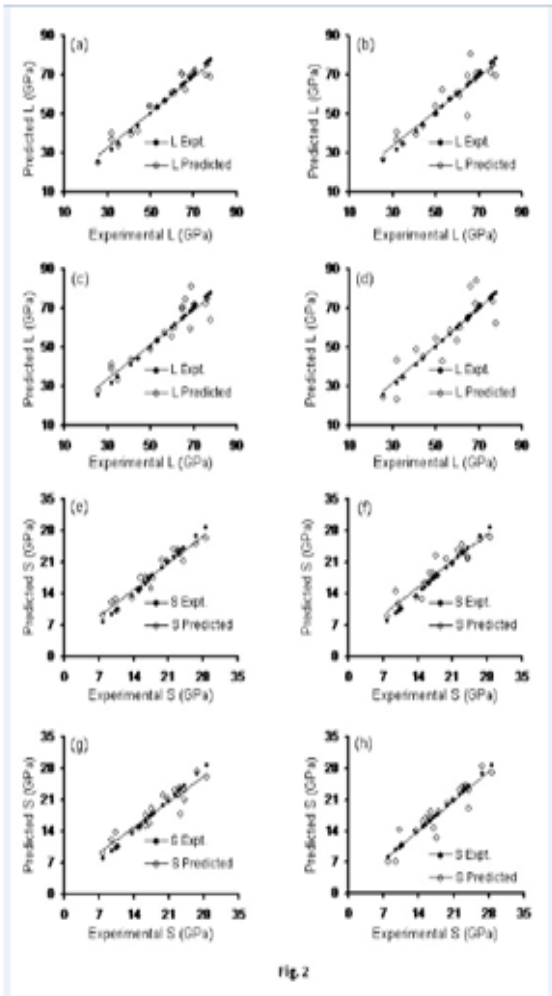

Fig. 2

Table 1 Glass systems used in test data and their characteristic parameters used as predictor variables of ANN

Sample No.	System (% Mol)	Predictor variables (inputs to ANN)						
		r (kg/m^3)	n_b	$\overline{n}_c$	F	V_T	G_T (10^9 kJ/mol)	FIC
1	0.85 TeO_2 – 0.15 BaO	5.613	12.80	4.00	218.3	0.57	51.83	0.4545
2	0.40 ZnO – 0.60 B_2O_3	3.212	10.90	1.75	409.7	0.53	33.52	0.5000
3	0.65 TeO_2 – 0.35 B_2O_3	4.790	11.20	2.44	311.7	0.56	43.01	0.4006
4	0.80 TeO_2 – 0.20 B_2O_3	4.970	11.40	3.00	266.0	0.52	47.72	0.3957
5	0.94 Na_2O – 0.06 B_2O_3	1.972	6.77	1.94	246.7	0.36	31.32	0.7927
6	0.10 Li_2O – 0.90 B_2O_3	1.980	6.00	1.30	623.8	0.44	28.13	0.4814
7	0.11 Na_2O – 0.88 B_2O_3	1.920	5.66	1.33	580.7	0.42	23.62	0.4910
8	0.70 TeO_2 – 0.30 MoO_3	4.912	11.40	4.00	218.7	0.53	58.74	0.3810
9	0.70 B_2O_3 – 0.30 Na_2O	2.340	8.16	1.90	464.6	0.50	25.38	0.5822
10	0.73 V_2O_5 – 0.27 PbO	4.717	7.27	2.00	240.2	0.84	57.57	0.4898
11	0.24 Ce_2O_3 – 0.76 P_2O_5	3.120	4.53	2.47	157.4	0.55	38.42	0.5067
12	0.25 Ce_2O_3 – 0.30 P_2O_5	3.178	4.56	2.50	156.4	0.55	39.25	0.5157
13	0.70 TeO_2 – 0.30 V_2O_5	4.564	8.93	3.08	229.4	0.57	58.65	0.4278
14	0.75 TeO_2 – 0.25 ZnO	5.470	14.10	4.00	215.7	0.51	52.97	0.4350
15	0.50 TeO_2 – 0.50 V_2O_5	3.995	7.76	3.33	243.8	0.58	62.00	0.4673
16	0.90 TeO_2 – 0.10 Sm_2O_3	5.782	11.90	2.71	203.6	0.51	54.91	0.4302
17	0.90 TeO_2 – 0.10 La_2O_3	5.685	11.90	4.18	203.0	0.51	55.18	0.4325
18	0.79 TeO_2 – 0.21 WO_3	5.390	8.22	2.42	211.3	0.50	56.90	0.4385
19	0.90 TeO_2 – 0.1 Nb_2O_5	5.414	11.50	4.00	207.2	0.53	57.14	0.4100
20	0.04 Ag_2O – 0.96 B_2O_3	2.109	5.21	1.12	625.4	0.42	22.78	0.4324
21	0.80 TeO_2 – 0.20 PbO	7.278	18.10	4.00	200.5	0.70	48.26	0.3720
22	0.70 TeO_2 – 0.30 MoO_3	4.900	11.40	4.00	218.7	0.53	58.74	0.3810

Table 2
Range of values of input and output parameters used in training set and test set

Parameter	Range				
	Training set			Test set	
	Min.	Max.		Min.	Max.
Inputs					
r (kg/m^3)	1.88	5.70		1.920	7.278
n_b	14.3	4.48		4.53	18.1
$\overline{n}_c$	1.03	4.00		1.12	4.00
F	194.8	663.6		156.4	625.4
V_T	0.34	0.88		0.36	0.70
G_T (10^9 kJ/mol)	23.12	61.90		22.78	62.00
FIC	0.3750	0.8122		0.3720	0.7927
Outputs					
L (GPa)	23.57	140.73		25.66	77.95
S (GPa)	6.92	35.20		7.85	28.63

Table 3 ANN predicted values of ultrasonic longitudinal (L) and shear (S) moduli values using the four training algorithms

Sample No.	Longitudinal modulus (GPa)					Shear modulus (GPa)				
	Exp.	ANN				Exp.	ANN			
		trainscg	traincgb	trainoss	trainlm		trainscg	traincgb	trainoss	trainlm
1	56.73	56.61	57.15	58.05	58.16	17.90	17.55	18.14	18.19	18.34
2	64.61	70.68	48.39	69.44	64.38	14.88	14.81	12.7	14.80	16.16
3	61.42	60.67	58.84	59.97	60.14	20.84	21.22	20.68	21.30	21.11
4	75.59	69.9	70.8	72.39	74.81	28.63	26.20	26.51	26.04	26.94
5	34.90	34.12	34.37	33.42	34.62	10.63	10.48	10.52	10.54	10.54
6	31.87	40.09	40.24	40.88	43.42	10.22	12.90	11.93	13.71	14.2
7	40.89	39.61	39.46	43.32	48.88	16.36	16.78	18.47	17.71	18.46
8	49.86	53.74	49.9	48.96	54.25	16.28	17.57	16.39	15.13	16.96
9	76.59	75.42	74.78	75.06	73.52	24.02	23.70	22.30	23.09	23.18
10	59.92	60.51	60.52	55.49	53.09	17.06	17.62	18.58	15.77	14.65
11	70.22	71.08	70.11	71.55	71.03	23.17	23.34	23.38	23.65	23.51
12	70.89	72.03	71.04	71.7	70.85	22.15	23.71	23.67	23.27	23.27
13	44.14	41.2	43.99	44.69	44.76	13.63	13.11	13.31	13.99	13.97
14	77.95	68.62	69.38	64.06	62.37	24.05	21.21	21.8	20.97	19.02

15	53.37	53.26	62.5	53.26	42.98	17.55	15.41	22.48	19.09	12.26
16	68.66	69.22	70.23	80.92	83.77	26.68	24.92	26.01	27.29	28.51
17	66.30	62.47	80.11	74.4	80.92	22.90	23.62	24.8	22.08	23.95
18	68.34	68.37	67.82	59.3	72.21	23.32	22.80	23.07	17.94	24.1
19	64.96	69.95	69.14	70.37	64.79	19.77	21.31	21.72	22.05	20.38
20	25.66	24.99	26.53	27.58	24.49	7.85	9.19	8.94	9.25	7.09
21	31.80	36.63	34.61	38.94	23.59	9.55	12.08	14.47	11.98	7.05
22	49.86	53.75	49.9	48.96	54.25	15.39	17.57	16.39	15.13	16.96

Table 4
Evaluation parameters of the network performance for the four training methods

Evaluation parameter	Training algorithm											
	trainscg			traincgb			trainoss			trainlm		
	L	S	Mean	L	S	Mean	L	S	Mean	L	S	Mean
MAPE	5.5179	7.6654	6.5916	6.4788	9.8498	8.1643	8.5194	8.7249	8.6222	9.9101	9.8498	9.8799
RMS	3.8901	1.4778	2.6839	5.8205	2.1478	3.9842	6.0287	1.8820	3.9554	7.3046	2.1478	4.7265
R^2	0.9956	0.9938	0.9947	0.9902	0.9874	0.9888	0.9894	0.9890	0.9892	0.9845	0.9874	0.9859

REFERENCES

[1] Aongxing Y, Wei Z, Chengzhi L. Optimal design an techno economic analysis of a hybrid solar-wind power generation system. Appl Energy 2009;86:163–9. | [2] Altien F, Kisi O, Aydin K. Predicting the compressive strength of steel fiber added light weight concrete using neural network. Comput Mater Sci 2008;48:259–65. | [3] Rafiq MY, Bugmann G, Easterbrook DJ. Neural network design for engineering applications. Comput Struct 2001;79:1541–52. | [4] Esteban LG, Fernandez FG, Palacios PD. MOE prediction in Abies Pinsapo Boiss. timber: application of an artificial neural network using non-destructive testing. Comput Struct 2009;87:1360–5. | [5] DeGiorgi MG, Ficarella A, Tarantino M. Error analysis of short term wind power prediction models. Appl Energy 2011;88:1298–1311. | [6] Demir F. A new way of prediction elastic modulus of normal and high strength concrete-fuzzy logic. Cement Concrete Res 2005;35:1531–8. | [7] Lande, PS. and Gadewar, AS. Application of artificial neural networks in prediction of compressive strength of concrete by using ultrasonic pulse velocities. IOSR J. Mech. Civil Eng 2012;3: 34-42. | [8] Pishgar-Komleh, SH., Keyhani, A., Mostofi-Sarkari, MR. and Jafari, A. Predicting the relationship between grain-combine travel, cylinder speed and harvesting losses by applying artificial neural networks. Int. J. Appl. Sci. Eng. Res 2012;1: 405-414. | [9] Kamanli, M., Yasar Kalatakci, M., Bahadir, F., Balik, FS., Husnu Korkmaz, H., Sami Donduren, M. and Tolga Cogurcu, M. Predicting the flexural behaviour of reinforced concrete and lightweight concrete beams by ANN. Indian J. Eng. Mater. Sci 2012;19: 87-94. | [10] Bridge B, Patel ND, Waters, DN. Elastic constant and structure of the vitreous system. Phys Status Solidi 1983;77:655. | [11] Makishima A, Mackenzie JD. Calculation of bulk modulus, shear modulus and Poisson's ratio of glass. J Non-Cryst Solids 1975;17:147–57. | [12] Makishima A, Mackenzie JD. Direct calculation of Young's modulus of glass. J Non-Cryst Solids 1973;12:35–45. | [13] Higazy AA, Bridge B. Elastic constants and structure of the vitreous system CO3O4-P2O5. J Non-Cryst Solids 1985;72:81–108. | [14] Demuth H, Beale M. Neural network toolbox for use with MATLBE, Mathworks Inc. 1994, p. 158. |