UDC 519.65:519.61:531.7
Original scientific paper
Received: 18.12.2000.

ANN cum mathematical modelling of measurement data

Lavit Rawtani, J.L. Rana and **A.K. Tiwari**
Department of Computer Science & Electronics, Maulana Azad College of Technology, Bhopal 462007, INDIA
e-mail: lavit_rawtani@yahoo.com

SUMMARY

Subject knowledge is often available in the form of data sets, obtained through experiments, with the possibility of some measurement errors. Mathematical modeling of such data is convenient for the use in design. ANN representation of subject knowledge is necessary for including it in the knowledge base of connectionist expert systems which are convenient for decision making. A modeling method is developed which serves as a mathematical as well as an ANN representation, combining the advantages of both. The comparison of the method is carried out with conventional methods like polynomials of different degrees, Hermite cubics with a different number of break points and cubic splines with a different number of knots. For the same error of approximation, the method is found to require a smaller number of free parameters (degrees of freedom).

Key words*: ANN representation, mathematical modelling, knowledge base, expert systems, measurement data.*

1. INTRODUCTION

Expert systems are being designed for a wide variety of applications to perform tasks, which are normally in the domain of human experts. There is an increasing trend to develop connectionist expert systems, wherein the knowledge base is in the form of artificial neural networks [1-3] since they possess the advantages of system integration, parallel processing, error tolerance etc. The scientific knowledge is frequently available in the form of sets of measurement data. The neural network representation of such knowledge is necessary for including it in the knowledge base of a connectionist expert system.

Several methods are available in the literature for data approximation and for finding approximating curves through given data points. More common methods use polynomials, cubic splines, Hermite cubics etc. These methods are described in several texts [4-6]. Improvements in the conventional techniques are being attempted constantly [7-10]. While the designers of connectionist expert systems are constrained to use ANN representation, the others can use any mathematical model which gives better approximation, with a smaller number of free parameters.

In the present investigation a model is developed which serves as a mathematical as well as an ANN representation. The method is applied to sets of data for which the results of data approximation, using conventional methods, are already available in the literature. The data approximation performance of the present method are compared with those obtained by other conventional techniques.

2. MODEL

We consider the case where the data set consists of pairs of values (x_p, y_p) for $p=1, 2,..., n$. A simple network, consisting of one input neuron, one output neuron and a single hidden layer with s neurons, is used, as shown in Figure 1. The i^{th} neuron of the hidden layer receives an input of a_i times x, as the link connecting it to input x has a weight a_i. This neuron has a bias value of b_i and sigmoid activation function. The output of the neuron would be $h_i=sigmoid\ (a_i x-b_i)$. The link connecting the i^{th} neuron of the hidden layer to the output neuron has weight c_i. Through this link the output neuron receives an input of c_i times h_i. The total input received by the output neuron is $[\Sigma c_i\ h_i]$, $i=1$ *to* s. The output neuron has a bias value d and linear

activation function. Thus, the final output of the network is:

$$Y(x) = \lfloor \Sigma c_i h_i \rfloor - d = \lfloor \Sigma c_i \, sigmoid(a_i x - b_i) \rfloor - d, \quad (i = 1 \text{ to } s) \tag{1}$$

where:

$$\eta = sigmoid(\xi) = \frac{1}{(1 + e^{-\xi})} \tag{2}$$

This ANN modeling is, therefore, equivalent to fitting the data in the mathematical form of Eq. (1). The total number of unknown parameters in Eq. (1) is *3s+1* (a_i, b_i, c_i for *i=1* to *s* and *d*), which is smaller than the number of data points (*n*). As such the curve represented by Eq. (1), can not pass exactly through all *n* data points. The parameters are chosen in such a way that they are as close to the data as possible.

3. METHOD

A widely used method for the design of neural networks is back-propagation learning. It is a powerful method, capable of handling problems with several output quantities depending on several input quantities. However, the algorithm is very slow and can get caught in local minima [11, 12]. Constant improvements are taking place in ANN design for different applications [13-16]. For a problem in which there is one output quantity (ordinate *y*) depending on one input quantity (abscissa *x*), a fast algorithm is developed, for finding the parameters (a_i, b_i, c_i for *i=1* to *s* and *d*) appearing in Eq. (1), which are also the weights and bias values of the ANN shown in Figure 1.

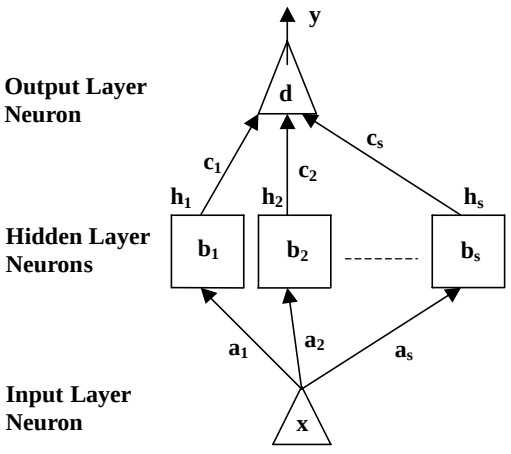

Fig. 1 ANN used for modeling. Hidden layer neurons have Sigmoid activation function. Output neuron has linear activation function.

3.1 Selecting value of *s*

A plot of the given data points, (x_p, y_p) for *p=1, 2,..., n*, arranged in increasing or decreasing order of x_p, is used to locate the turning points. Turning points are those data points, near which the radius of curvature (ρ) of the resulting curve would have a minima. Such points are easily identified visually, as bottoms of cups (kept erect, inverted or tilted). Turning points where slope changes sign, are denoted as T-points, while the other turning points are denoted as t-points. Figure 2 illustrates an example with *n=23* data points, four of which are turning points, with serial numbers f_1*=5* (T-point), f_2*=9* (t-point), f_3*=14* (t-point) and f_4*=19* (T-point). The starting point is f_0*=1* and end point is f_5*=n=23*.

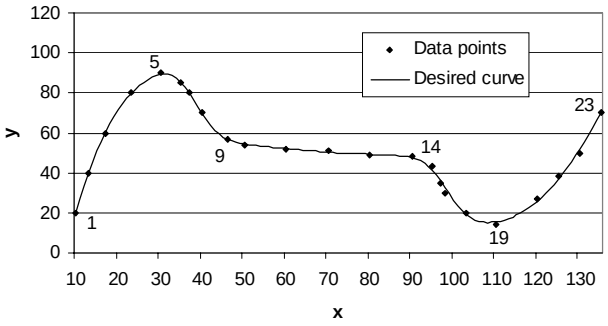

Fig. 2 Example to illustrate visual location of turning points. Serial numbers of turning points are: f_1=5 (T-point), f_2=9 (t-point), f_3=14 (t-point), f_4=19 (T-point). Starting point is f_0=1 and end point is f_s=f_5=n=23, making s=5.

In Eq. (1), the curve is modeled as a sum of *s* sigmoid functions (subjected to scaling and translation transformations). Therefore, the value of *s* should depend on the shape of the curve through the data points and the shape of the sigmoid function curve (Figure 3). The value of sigmoid (*x*), as defined in Eq. (2), increases from 0 to 1 as *x* increases from -∞ to +∞. It has no T-point, as its slope is always positive. It has two t-points at *x=-1.36* and *x=1.36*, with an equal numerical value of ρ.

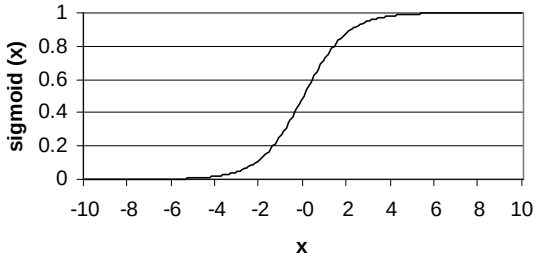

Fig. 3 Variation of sigmoid (x) with x

In general, the turning points on a curve will have different values of ρ. A curve having no T-point but zero or one t-point or two t-points with equal values of ρ, is termed SPR (sigmoid portion resembling) curve. Such a curve may be modeled by a single portion of the sigmoid curve (subjected to the transformation of scaling and translation), i.e. by taking *s=1* in Eq. (1). A curve having one T-point can be modeled as a difference (algebraic sum) of two sigmoid functions (*s=2*). Every additional T-point will require an addition to the value of *s*. Thus, the

value of *s* must be at least one more than the number of T-points on the required curve. To remove the constraint on the number of t-points it is better to take a larger value of *s*, up to one more than the total number of turning points. Thus:

(Number of T-points+1) ≤ s ≤ (Total number of turning points+1)

3.2 SPR case

Consider a case of *n* data points having coordinates (x_p, y_p), $1 \le p \le n$, arranged in increasing or decreasing order of abscissa. A portion of the sigmoid function, defined in Eq. (2), from $\xi=\xi_s$ to $\xi=\xi_e$, after a suitable transformation (scaling and translation), is proposed to be used for fitting the data points. The transformation used is:

$$\xi = ax - b, \quad y = c\eta - d \tag{3}$$

It is insisted that point (ξ_s, η_s) and point (ξ_e, η_e) coincide with the starting point (x_1, y_1) and the end point (x_n, y_n), respectively. Thus:

$$\xi_s = ax_1 - b, \quad \xi_e = ax_n - b, \quad y_1 = c\eta_s - d, \quad y_n = c\eta_e - d \tag{4}$$

Equations (4) together with Eq. (2) yield:

$$\begin{aligned} a &= (\xi_e - \xi_s)/(x_n - x_1), \\ b &= ax_1 - \xi_s, \\ c &= (y_n - y_1)/[sigmoid(\xi_e) - sigmoid(\xi_s)], \\ d &= c\ sigmoid(\xi_s) - y_1 \end{aligned} \tag{5}$$

Combination of Eqs. (2) and (3) yields:

$$y = c\ sigmoid(ax - b) - d \tag{6}$$

The data points are fitted into the curve given by Eq. (6), using Eq. (5) for obtaining the values of *a*, *b*, *c* and *d*. It can be seen that the values of *a*, *b*, *c* and *d* depend on two free parameters ξ_s and ξ_e. Whatever values are chosen for ξ_s and ξ_e the curve of Eq. (6) will pass through the first data point (x_1, y_1) and the last data point (x_n, y_n). The actual values of ξ_s and ξ_e must be chosen so that there is a very close fit at the other data points ($p=2, 3, ..., n-1$). At $x=x_p$ the value of *y* obtained from the approximating curve of Eq. (6) will be $\{c\ sigmoid\ (ax_p-b) - d\}$, as against the given value of y_p. The sum of the squares of errors:

$$E = \sum_{p=1}^{n} \left[\{ c\ sigmoid(ax_p - b) - d \} - y_p \right]^2 \tag{7}$$

has to be minimized. Since the fit at the first and the last points is exact the summation is taken from $p=1$ to n instead of from 2 to $n-1$. The minimization of *E* is achieved as follows:

(a) Choose any initial values of ξ_s and ξ_e;

(b) Keeping the value of ξ_e fixed, change the value of ξ_s in small steps Δ, till the error function *E* gets minimized;

(c) Keeping the value of ξ_s fixed as obtained in step (b), change the value of ξ_e in small steps Δ, till the value of error function *E* is minimized;

(d) Repeat steps (b) and (c), till the change caused in *E* is insignificant (less than a chosen tolerance value).

The values of *a*, *b*, *c* and *d* are obtained from Eq. (5) using these values of ξ_s and ξ_e.

3.3 General case

Consider the general case of *n* data points (x_p, y_p), $1 \le p \le n$, arranged in increasing or decreasing order of abscissa. The visual observation of their plot will give the serial numbers $(f_1, f_2, ..., f_{s-1})$ of the data points, which are turning points. The serial numbers of the beginning and end points are $f_0=1$ and $f_s=n$, respectively.

As the first step, an SPR curve is fitted through data points f_0 to f_1 as explained in Section 3.2. The corresponding values obtained for *a*, *b*, *c* and *d*, are denoted as a_1, b_1, c_1 and d_1. The representative equation of the curve will be similar to Eq. (6) i.e.:

$$F(A_1, x) = c_1\ sigmoid(a_1 x - b_1) - d_1 \tag{8.1}$$

The plot of this equation will be a good match for the portion of the curve from points f_0 to f_1, but there will be vast deviation beyond, as shown in Figure 4.

As the second step consider the points whose coordinates are $[x_p, y_p\text{-}F(A_1, x_p)]$. An SPR curve is passed through points f_0 to f_2 of this curve. The corresponding values obtained for *a*, *b*, *c* and *d*, are denoted as a_2, b_2, c_2 and d_2. The representative equation of the curve will be similar to Eq. (6) i.e.:

$$F(A_2, x) = c_2\ sigmoid(a_2 x - b_2) - d_2 \tag{8.2}$$

The plot of $[F(A_1, x)+F(A_2, x)]$ will be a good match for the portion of curve from points f_0 to f_2, but there will be vast deviation beyond, as shown in Figure 4.

As the third step we consider the points whose coordinates are $[x_p, y_p\text{-}F(A_1, x_p)\text{-}F(A_2, x_p)]$, to get:

$$F(A_3, x) = c_3\ sigmoid(a_3 x - b_3) - d_3 \tag{8.3}$$

The plot of $[F(A_1, x)+F(A_2, x)+F(A_3, x)]$ will be a good match for the portion of curve from points f_0 to f_3.

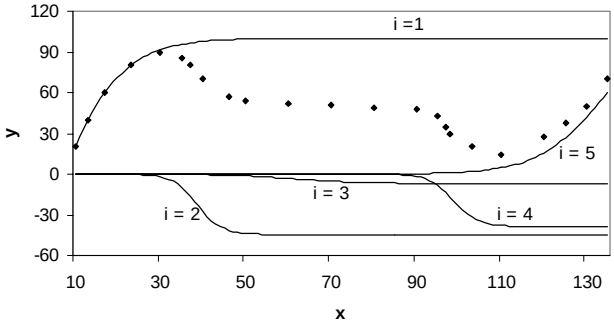

Fig. 4(a) Plot of $F(A_i,x)$ for different values of i, obtained for the data points of the illustrative example shown in Figure 2

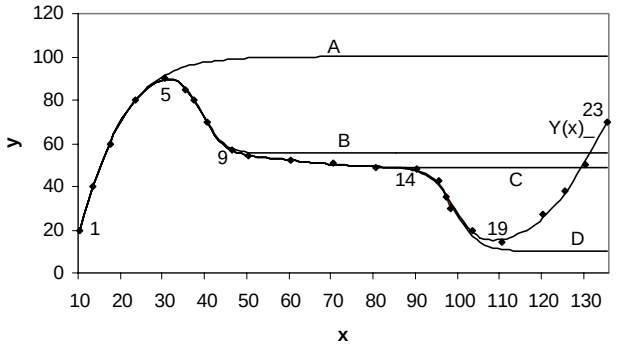

Fig. 4(b) Illustration of the method for data set having n=23, s=5, f_0=1, f_1=5, f_2=9, f_3=14, f_4=19, f_5=23

A - $\Sigma F(A_i,x)$, (i=1 to 1) to provide match from points f_0 to f_1
B - $\Sigma F(A_i,x)$, (i=1 to 2) to provide match from points f_0 to f_2
C - $\Sigma F(A_i,x)$, (i=1 to 3) to provide match from points f_0 to f_3
D - $\Sigma F(A_i,x)$, (i=1 to 4) to provide match from points f_0 to f_4
E - $Y(x)=\Sigma F(A_i,x)$, (i=1 to s) to provide match from points f_0 to f_s

The process is repeated *s* times to get a match for the entire curve from points f_0 to f_s (points *1* to *n*), through the expression:

$$Y(x)=\left[\sum F(A_i,x)\right]=\left[\sum c_1\, sigmoid(a_i x-b_i)-d_i\right],\quad (i=1\ to\ s) \tag{8.4}$$

$$d=\sum d_i,\quad (i=1\ to\ s) \tag{8.5}$$

The general procedure is given below:

* Input the values of x_p, y_p *(p=1, 2, ..., n)* and serial numbers of turning points f_i *(i=1, 2, ..., s-1)*. The beginning and the end points are $f_0=1$ and $f_s=n$.
* Initialize vectors *v* and *Z*, $v_p \leftarrow y_p$ and $Z_p \leftarrow 0$, *(p=1, 2, ..., n)*.
* For *i=1, 2, ..., s*, do
  * Compute the new vector *v* as $v_p \leftarrow v_p - Z_p$, $(p=f_{i-1}, 2, ..., n)$ and $v_p \leftarrow 0$, $(p<f_{i-1})$.
  * Fit an SPR curve through points having coordinates (x_p, v_p) for $(p=1,2,.....,f_i)$, using the method described in Section 3.2 and denote the values obtained for *a*, *b*, *c* and *d* as a_i, b_i, c_i and d_i, respectively.
  * Compute the new vector *Z*, $Z_p \leftarrow c_i\, sig\,(a_i x_p - b_i) - d_i$ for *(p=1, 2, ..., n)*.
* End for
* Compute $d=\Sigma d_i$, *(i=1, 2, ..., s)*

It is to be noted that vectors *v* and *Z* are local to this procedure. In step *i*, the fit is attempted for points (x_p, v_p) and Z_p is the value of $[F(A_i,x_p)]$.

3.4 Second run

The given data set (x_p, y_p), $1 \le p \le n$, gets represented by (x_p, Y_p) by the application of Eq. (1) with a_i, b_i, c_i and *d* determined by the procedure outlined in Section 3.3, causing a small error $e_p=y_p-Y_p$ at $x=x_p$. If the same procedure is run a second time, starting with the imaginary coordinates for data points as $[x_p, y_p+e_p]$ i.e. $[x_p, 2y_p-Y_p]$, the resulting output would be nearer to (x_p, y_p). The output of the second run is compared with the original coordinates of the data points (x_p, y_p) to determine the errors.

4. RESULTS

Rice [6] has considered the following data sets to compare the approximation capabilities of five conventional methods.

The data set A consists of 24 points as given below,

abscissa values

14.00, 13.00, 12.50, 12.44, 12.36,12.28, 12.20, 12.16, 12.12, 12.08, 12.04, 12.00, 11.96, 11.89, 11.80, 11.60, 11.40, 11.20, 11.00, 10.80, 10.60, 10.40, 10.20, 10.00.

ordinate values

4.64, 4.64, 4.64, 4.62, 4.53, 4.27, 3.83, 3.40, 2.89, 2.35, 1.87, 1.52, 1.29, 0.91, 0.74, 0.65, 0.61, 0.58, 0.55, 0.53, 0.52, 0.51, 0.48, 0.42.

The plot of these points is shown in Figure 5. There is no T-point, but there are two *t*-points (point numbers 5 and 15). The results of the present method are given in Table 1, for *s=1* and *s=2*.

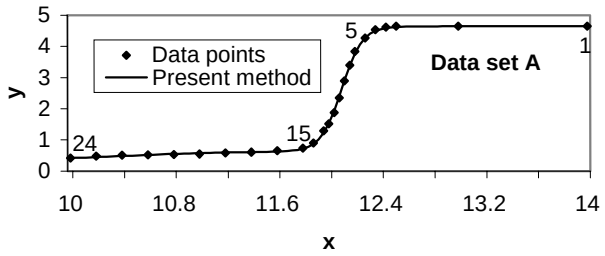

Fig. 5 Results of present method for data set A

Table 1. Values of parameters a_i, b_i, c_i and d obtained for different data sets and rms error during the first and second run

Values of s and f_i	*i*	a_i	b_i	c_i	*d*	*RMS error First run*	*RMS error Second run*
1	*2*	*3*	*4*	*5*	*6*	*7*	*8*
Data set A							
s=2, f_0=1, f_1=15, f_2=24	*1* *2*	*-12.6356* *-2.1812*	*-152.873* *-22.6425*	*-3.99733* *-0.31994*	*-4.64011*	*0.03042*	*0.02868*
s=1, f_0=1, f_1=24	*1*	*-10.9114*	*-131.885*	*-4.22000*	*-4.64000*	*0.10426*	*0.10061*
Data set B							
s=3, f_0=1, f_1=31, f_2=37, f_3=49	*1* *2* *3*	*0.07398* *0.11565* *0.19106*	*64.81545* *104.6302* *181.7535*	*2.61032* *-2.57604* *-0.07027*	*-0.64400*	*0.05253*	*0.03519*
s=2, f_0=1, f_1=31, f_2=49	*1* *2*	*0.07462* *0.10763*	*64.95343* *97.41068*	*2.69610* *-2.73210*	*-0.64400*	*0.06085*	*0.03772*

The data set B consists of 49 points having abscissa values $x_p=575+10p$, $1\le p\le 49$. Corresponding ordinates y_p are given below,

0.644, 0.622, 0.638, 0.649, 0.652, 0.639, 0.646, 0.657, 0.652, 0.655, 0.644, 0.663, 0.663, 0.668, 0.676, 0.676, 0.686, 0.679, 0.678, 0.683, 0.694, 0.699, 0.710, 0.730, 0.763, 0.812, 0.907, 1.044, 1.336, 1.881, 2.169, 2.075, 1.598, 1.211, 0.916, 0.746, 0.672, 0.627, 0.615, 0.607, 0.606, 0.609, 0.603, 0.601, 0.603, 0.601, 0.611, 0.601, 0.608.

The plot of the data points is given in Figure 6. There are three turning points at serial number 24 (t-point), 31 (T-point) and 37 (t-point). The results of the present method are given in Table 1, for *s=2* and *s=3*.

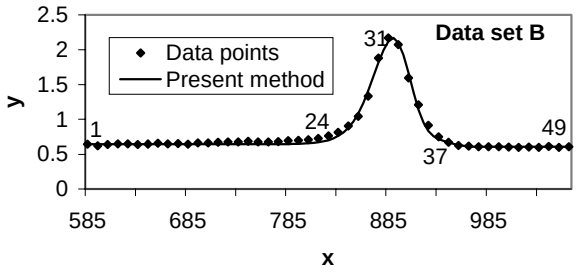

Fig. 6 Results of present method for data set B

5. COMPARISON OF METHODS

The data points (x_p, y_p) get approximated as (x_p, Y_p) for $(p=1, 2, \ldots, n)$. The maximum error of approximation is, $E_{max}=max\,|y_p-Y_p|$, for $(p=1, 2, \ldots, n)$. Rice [6] has applied the following five models to these data sets and computed the maximum error (E_{max}):

1. Polynomials of degrees 6, 8 and 11 (7, 9 and 12 parameters);
2. Cubic splines with 5, 7 and 10 equidistant knots (7, 9 and 12 parameters);
3. Cubic splines with 5, 7 and 10 data dependent knots (7, 9 and 12 parameters);
4. Hermite cubics with 5, 7 and 10 equidistant break points (10, 14 and 20 parameters);
5. Hermite cubics with 5, 7 and 10 data dependent break points (10, 14 and 20 parameters).

These methods are referred to as conventional methods 1, 2, 3, 4 and 5, respectively.

A realistic comparison of different models must consider not only the maximum error but also the number of free parameters (degrees of freedom) involved. The results of comparison are given in Table 2 and plotted in Figures 7 and 8, with the maximum error as the ordinate and the number of parameters as the abscissa, for data sets A and B, respectively. A better method should give points nearer to the abscissa axis (less error of approximation), as well as nearer to the ordinate axis (smaller number of free parameters). It can be seen

Table 2. Comparison of results obtained by the present method with those obtained by five other models by Rice

Representation Model	*Number of Parameters*	*Max. error in data set A*	*Max. error in data set B*
(1)	*(2)*	*(3)*	*(4)*
1. Polynomial			
degree=6	*7*	*0.82*	*0.87*
degree=8	*9*	*0.64*	*0.73*
degree=11	*12*	*0.53*	*0.52*
2. Cubic Splines with equidistant knots			
knots=5	*7*	*1.03*	*0.89*
knots=7	*9*	*0.80*	*0.58*
knots=10	*12*	*0.30*	*0.46*
3. Cubic Splines with data dependent knots			
knots=5	*7*	*0.59*	*0.44*
knots=7	*9*	*0.072*	*0.30*
knots=10	*12*	*0.053*	*0.10*
4. Hermite cubics with equidistant break points			
break points=5	*10*	*0.43*	*0.61*
break points=7	*14*	*0.26*	*0.14*
break points=10	*20*	*0.14*	*0.076*
5. Hermite cubics with data dependent break points			
break points=5	*10*	*0.12*	*0.099*
break points=7	*14*	*0.034*	*0.077*
break points=10	*20*	*0.016*	*0.0091*
6. Present Method			
s=1	*4*	*0.200*	-
s=2	*7*	*0.065*	*0.094*
s=3	*10*	-	*0.081*

that the results of the present method are better than the conventional methods, for the same number of degrees of freedom, in respect of both data sets.

The first data set has been described in [6], as typical of real world behavior, with inadequate data between $x=12.5$ and $x=14$, difficult to model mathematically. Referring to Figure 7, if the tolerance in the maximum error of approximation is taken as 0.2, the present method requires 4 free parameters, as compared to 9, 20 and 10 free parameters required for conventional methods 3, 4 and 5, respectively. The conventional methods 1 and 2 are not able to achieve this tolerance limit, even with 12 parameters. Thus, the present method has given remarkably better results as compared with the other five methods.

Rice [6] has described the second data set as *titanium data*, which is well known as a physical data set, containing moderate uncertainty and difficult to represent with a mathematical model. Referring to Figure 8, for tolerance of 0.1 in maximum error of approximation, the conventional methods 3, 4 and 5 require 12, 20 and 10 free parameters, respectively. Conventional methods 1 and 2 are nowhere near this tolerance limit, with 12 free parameters. Better results are obtained by the present method, as this tolerance limit is achieved with 7 free parameters.

The conventional methods 2 and 3 employ cubic splines with several knots. In between two successive knots, the modeling is by piecewise polynomials. Same is the case with conventional methods 4 and 5, which

employ Hermite cubics with several break points. In conventional method 1, the polynomial expression is valid over the entire region. In the present method, the data set is approximated by Eq. (1), which is valid over the entire range.

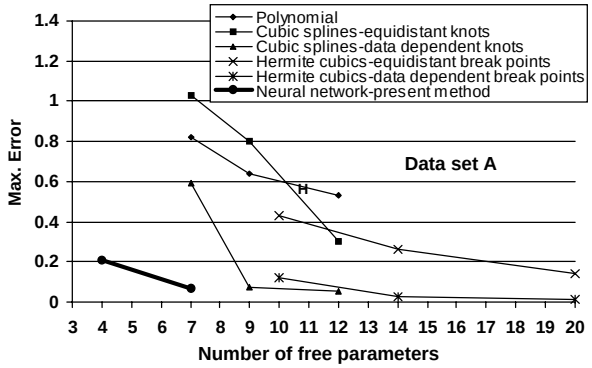

Fig. 7 Variation of maximum error with degrees of freedom for data set A

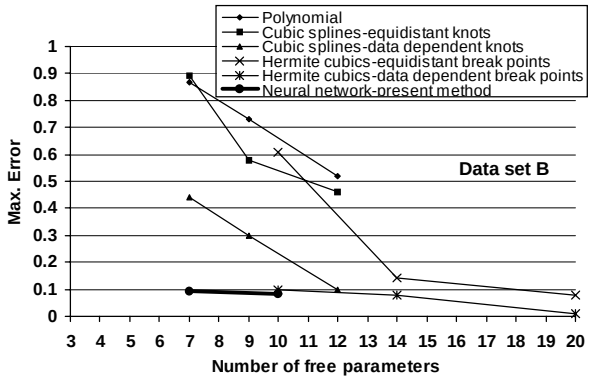

Fig. 8 Variation of maximum error with degrees of freedom for data set B

The Hermite cubics used in conventional methods 4 and 5, yield first derivative continuity at break points. The cubic splines, used in conventional methods 2 and 3, ensure the continuity of the first two derivatives at knots. The present modeling Eq. (1) is valid from the beginning to the end of the region and has an infinite number of continuous derivatives.

6. CONCLUSIONS

The study carried out indicates that the present method gives better data approximation than the other conventional methods, especially when the data contain moderate measurement errors. For the same maximum error of approximation, the present technique requires a smaller number of parameters. The Hermite cubics yield the first derivative continuity at break points. The cubic splines ensure continuity of the first two derivatives at knots. The present modeling Eq. (1) is valid from the beginning to the end of the region and has an infinite number of continuous derivatives. The algorithm used for the determination of parameters in Eq. (1), is very fast, as compared to other learning algorithms such as back-propagation.

REFERENCES

[1] P. Bhattacharya and L. Rawtani, Learning concepts in connectionist expert systems, Proc. of National Conference on Neural Networks and Fuzzy Logic, Anna University, Madras India, pp. 159-169, 1995.

[2] S.I. Gallant, *Neural Network Learning and Expert Systems*, MIT Press, Cambridge MA., 1993.

[3] L. Jing, K. Nezu and Xu Jinwu, Extracting fuzzy rules from neural network in connectionist expert system, Proc. 35th SICE Annual Conf., Tottori, Japan, pp. 1383-1387, 1996.

[4] G. Farin, *Curves and Surfaces for Computer Aided Geometric Design*, 4th edition, Academic Press, 1996.

[5] P. Lancaster and K. Salkauskas, *Curve and Surface Fitting*, Academic Press, 1986.

[6] J.R. Rice, *Numerical Methods, Software and Analysis*, McGraw Hill Int. Book Company, 1983.

[7] G. Eiber, I.K. Lee and M.S. Kim, Comparing offset curve approximation methods, *IEEE Computer Graphics Applications*, Vol. 17, No. 3, pp. 62-71, 1997.

[8] B. Hamann, Data point selection for piecewise linear curve approximation, *Computer Aided Geometric Design*, Vol. 11, pp. 289-301, 1994.

[9] C.G. Lim, A universal parameterization in B-spline curve and surface interpolation, *Computer Aided Geometric Design*, Vol. 16, pp. 407-422, 1999.

[10] E. Sarkar and C.H. Meng, Parameter optimization in approximating curves and surfaces to measurement data, *Computer Aided Geometric Design*, Vol. 8, pp. 267-290, 1994.

[11] M.H. Hassoun, *Fundamentals of Artificial Neural Networks*, Prentice Hall of India, New Delhi, original edition published by MIT Press, Cambridge, 1998.

[12] K. Mehrotra, C.K. Mohan and S. Ranka, *Elements of Artificial Neural Networks*, MIT Press, Cambridge, 1997.

[13] Z. Lozina and I. Duplan~i}, Application of the artificial neural network to the numerical modelling of the experimental data and simulations of the extrusion processes, *Int. Journal for Engineering Modelling*, Vol. 11, No. 3, pp. 91-96, 1998.

[14] Y. Fukuoka, H. Matsuki, H. Minamitani and A. Ishida, Modified back-propagation method to avoid false local minima, *Neural Networks*, Vol. 11, No. 6, pp. 1059-1072, 1998.

[15] G. Ritter, Efficient estimation of neural weights by polynomial approximation, *IEEE Trans. Information Theory*, Vol. 45, No. 5, pp. 1541-1550, 1999.

ANN CUM MATEMATI^KO MODELIRANJE PODATAKA MJERENJA

SA@ETAK

Znanje o nekom predmetu istra`ivanja ~esto je dostupno u obliku skupa podataka koji se dobivaju eksperimentima, pri }emu postoji mogu}nost pojave nekih gre{aka pri mjerenjima. Matemati~ko modeliranje ovih podataka mo`e se korisno primijeniti u projektiranju. ANN prikaz znanja o predmetu istra`ivanja je potreban kako bi se to znanje uklju~ilo u bazu znanja veznih ekspertnih sustava koji su prikladni za donošenje odluka. Razvijena metoda modeliranja slu`i i kao matemati~ki prikaz i kao ANN prikaz, time što kombinira prednosti oba prikaza. Ova metoda uspore|uje se s klasi~nim metodama kao što su polinomi razli~itih stupnjeva, Hermite-ove kubne jednad`be s razli~itim brojem to~aka prekida i kubi~ni spline-ovi s razli~itim brojem ~vorova. Za istu grešku aproksimacije dokazano je da razvijena metoda zahtijeva manji broj parametara (stupnjeva slobode).

***Klju~ne rije~i**: ANN prikaz, matemati~ko modeliranje, baza znanja, ekspertni sustavi, podaci mjerenja.*