



Optimizing the Neural Network Parameters in Chaotic Data Time Series

Shailendra Singh¹, Sanjeev Karmakar², Navita Shrivastava³, R.K. Tiwari⁴

^{1,3}A. P. S. University, Rewa, Madhya Pradesh, India.

²Bhilai Institute of Technology (BIT), Bhilai House, Durg, C.G, India.

⁴Govt. New Science College, Rewa, Madhya Pradesh, India.

Abstract: The Back-Propagation Neural Network is an effective data mining technique in the identification of suitable parameters for long-term rainfall data. BPN model is developed and verified for the training and testing data set to identify climatic parameter for rainfall prediction. In this study the model is used to predict parameters and their optimum values for time series 1970-2014 of vindhya region. The system recommended values for the 12 optimized parameters are observed and the results are analyzed accordingly.

Keywords: Deterministic Forecast, BPN, Learning Rate, Momentum Factor, Input Vector etc.

I. Introduction

The long-term chaotic data time series study over vindhya region for rainfall prediction is studied. Many researchers have forecast for rainfall and other climatic parameters using numerical and statistical methods. These methods are not accurate because of chaotic nature of rainfall data series. The Back Propagation Neural Network in deterministic forecast for long range rainfall is found to be perfect for chaotic rainfall data. Many undefined parameters i.e. number of input vectors, number of hidden layers, learning rate, momentum factor, bias and weights are required to predict rainfall. BPN is best established to forecast chaotic behavior rainfall as well as other climate parameters prediction[1]. BPN is found to be fit for prediction of various climate activities [2]. BPN model predict 99.8% and 94.3% accuracy during the training and testing period respectively for rainfall data [3,4]. Karmakar et al.,[5] have studied the predicted rainfall in Chennai using back propagation neural network model. Present study is very useful to determine optimum values of forecast parameters for rainfall prediction. This model is most suitable for identification of feature vectors of chaotic series by past data time series [6].

Krishnamurthy et al.[7,8], Sahai et al., [9] have found that the statistical technique to forecast monsoon rainfall as well as climate parameters over smaller areas like a district and for a monsoon periods is not appropriate. The poor correlation between the dependent and independent parameters is observed. The long term study show that the architectures of ANN such as BPN, RBFN is best established to forecast chaotic behavior and have efficient enough to forecast monsoon rainfall as well as other climate parameter prediction phenomenon over the smaller geographical region [10]. The mean monthly rainfall is predicted using ANN model. The model performed well both in training and independent periods

In this study, the BPN is used in deterministic forecast for long-range monsoon rainfall for vindhya region of Madhya Pradesh for determination of impact of variance in learning rate and momentum factor in the model.

II. Data Description and Preprocessing

Data time series (x_i) for the first 35 years (1970–2004) are used for developing the model for vindhya region of Madhya Pradesh. Transfer function sigmoid axon is used in the BPN model. The output of sigmoid axon has in close interval 0 to 1. Therefore, model data time series is normalized by using following equation-1 and obtained new normalized data time series (r_i). The eq-2 is used to unnormalize thereafter for actual representation x_i . Remaining 10 years (2005–2014) data time series (x_i) is used to test the model independently for its acceptance.

$$\frac{x_i + \min(x_i)}{x_i + \max(x_i)} \quad (1)$$

$$\frac{\min(x_i) - r_i \cdot \max(x_i)}{(r_i - 1)} \quad (2)$$

Data Normalization and Splitting:

Data Input:

Data entry type: Imported from ExcelSheet

Data series name 1: Year Actual and Normalized Data Series: 45 Years

Data series name 2: Rainfall Number of Training Samples: 35 Years

Units : mm. Number of Testing Samples: 10 Years

Splitting (training and testing dataset):

Training Dataset			
Year	Normalized value of Rainfall (mm)	Year	Normalized value of Rainfall (mm)
1970	0.591377	1987	0.598493
1971	0.669092	1988	0.597744
1972	0.631464	1989	0.537075
1973	0.579184	1990	0.63936
1974	0.573279	1991	0.58904
1975	0.632964	1992	0.532094
1976	0.57478	1993	0.581387
1977	0.643612	1994	0.591428
1978	0.651737	1995	0.622093
1979	0.509426	1996	0.610391
1980	0.662376	1997	0.505437
1981	0.600177	1998	0.564736
1982	0.629964	1999	0.636996
1983	0.55652	2000	0.570264
1984	0.605711	2001	0.639425
1985	0.594553	2002	0.594851
1986	0.578683		
Testing Dataset			
Year	Normalized value of Rainfall (mm)		
2005	0.589176		
2006	0.543197		
2007	0.539526		
2008	0.568055		
2009	0.517473		
2010	0.551524		
2011	0.614916		
2012	0.625406		
2013	0.64168		
2014	0.611216		

Table-1: Training data set (1970-2004) and Testing dataset (2005-2014)

Statistics of Data	Training Dataset	Testing Dataset
Min (x_i) of Rainfall	506.6	556.6
Max (x_i) of Rainfall	1498.0	1268.8
Mean value (x_i)	977.0942857142855	885.4799999999999
Standard Deviation (x_i)	244.88219489818215	241.67858269472973
% of Mean(x_i)	25.06228912383474	27.293511168488248
Minimum normalized value (r_i)	0.5054374937643421	0.5174729874428113
Maximum normalized value (r_i)	0.6690921228304405	0.6416799190400463
Mean value (r_i)	0.5956103832774173	0.5802169069399181
Standard Deviation (r_i)	0.0402430487455586	0.04517710183161851
% of Mean (r_i)	6.756606311010968	7.7862436084263615

Table-2: Statistics of training dataset and testing dataset

III. SELECTION OF BPN MODEL PARAMETERS

For the efficient operation of BPN, the appropriate selection of the parameters used for network as follows:

1. **Number of Layers:** This BPN Model has three layers, one at the bottom, one hidden layer at the middle, and one output layer at the top.

2. **Number of hidden layer:** Many researchers have observed that one hidden layer is sufficient, using two hidden layer rarely improves the model.
3. **Number of neurons in hidden layer:** There is two neurons and ten input vectors provided satisfactory result for all climatic data and increase in number of neurons in hidden layer increase Mean Absolute Deviation (MAD) between actual and predicted value.
4. **Number of Input Vectors(n):** The value of 'n' is depending on internal dynamics of data time series.
5. **Learning Rate (α):** A high learning rate leads to rapid learning but the weights may oscillate, while a lower learning rate leads to slower learning.
6. **Momentum Factor (μ):** The main purpose of the momentum factor is to accelerate the convergence of error propagation algorithm during the training period.
7. **Initial Weights:** To get the best result the initial weights are set to random number between 0 and 1. If initial weights(bias) too large the initial input signals to output unit will gone down. If it is too small output unit will approach to zero.
8. **Number of biases in hidden layer:** The initial weights (bias) can be done randomly and moreover there is a specific approach. The faster learning of a BPN can be obtained by using Nguyen-Widrow initialization.
9. **Number of biases in output layer:** The value of this parameter is one since one output unit is used, it is used to improve the learning ability of the output layer.
10. **Transfer function:** Sigmoid function obtained neurons output $f(x)$. The output of the neuron will be in close interval (0,1).
11. **Mean Square Error Level:** BPN based on gradient descent method, it minimizes the MSE of the output computed by the network during feed-forward and back-propagation process.
12. **Number of Epochs:** It is a number of trains in the network. As much epochs increase, MSE is exhibiting an increasing trend.

Initial Variables of BPN:

A. Weights in hidden layer $V(i,j); i=10, j=2$		B. Biases in hidden layer $Vo(j); j=2$		B. Weights in output layer $W(j); j=2$	
$V(i,1)$	$V(i,2)$	$Vo(1)$	$Vo(2)$	$W(1)$	$W(2)$
0.889864	0.253677	0.390054	0.694807	0.546431	0.356959
0.931142	0.094606				
0.149762	0.079962				
0.83213	0.575017				
0.214287	0.651923				
0.41332	0.629334				
0.444514	0.933841				
0.881214	0.437617				
0.174556	0.732116				
0.435148	0.175127				

Table – 3: Initial Variables of BPN

Optimum Variables of BPN: Optimized weights and biases for Desired Epochs = 10000 and MSE = 0.0013592475951864345

A. Weights in hidden layer $V[i][j], s.t., i=1..10 \& j=1..2$		A. Updated weights $Vo[i], s.t., i=2$		B. Updated weights $W[i], s.t., i=2$	
$V(i,1)$	$V(i,2)$	$Vo(1)$	$Vo(2)$	$W(1)$	$W(2)$
0.887308	0.251494	0.390054	0.694807	0.546431	0.356959
0.928574	0.092413				
0.147197	0.077776				
0.829497	0.572786				
0.211762	0.649765				
0.41081	0.627184				
0.441937	0.931653				
0.878659	0.435436				
0.171959	0.729914				
0.43251	0.172895				
Bias in output layer: 1.95E-05		LPA = 956.74			

Table-4: Optimum Variables of BPN

Optimize Learning Rate (α), Momentum Factor (μ) and MSE:

A high learning rate leads to rapid learning but the weights may oscillate, while a lower learning rate leads to slower learning. Methods suggested for adopting learning rate are as follows:

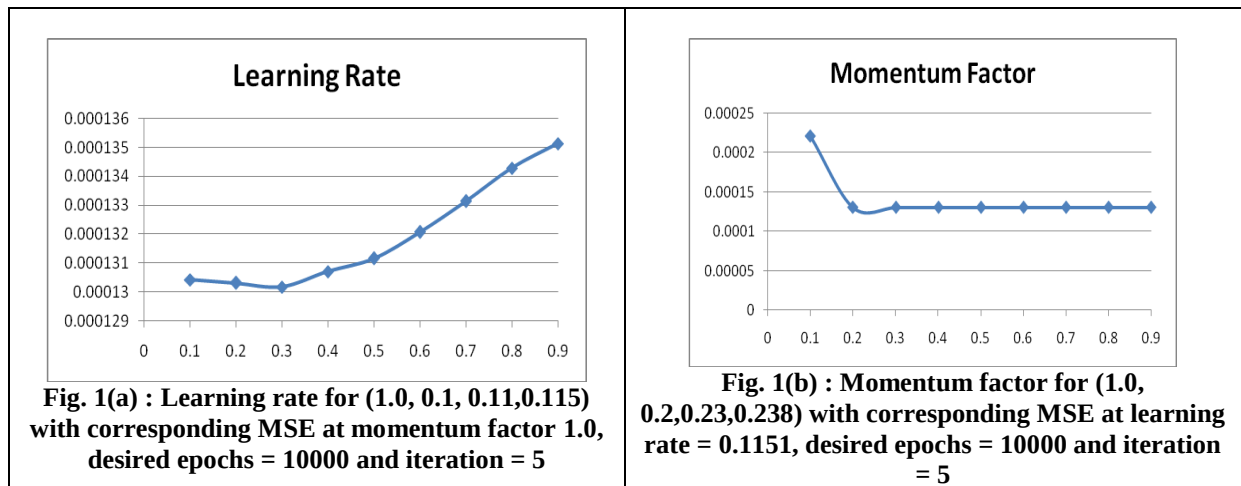
- Start with high learning rate between 0 and 1 and steadily decrease it. Changes in the weight vector must be small in order to reduce oscillation or any divergence.
- Increase the learning rate in order to improve performance and decrease the learning rate in order to worsen the performance.

Learning Rate (α)	P1	P2	P3	P4
0.1	0.00146615875876712	0.00146676533328296	0.00146677267806826	0.00146689179681701
0.2	0.00147619343727981	0.00146742589902894	0.00146687406211439	0.00146732998348665
0.3	0.00148629975955298	0.00146885793250091	0.00146583610755944	0.00146746433860021
0.4	0.00149692602064065	0.00147178913361247	0.00146679934678303	0.00146702356684762
0.5	0.00150386514823279	0.00147045667289356	0.00146581908864100	0.00146727312890675
0.6	0.00150828135564350	0.00147226751532258	0.00146745793168144	0.00146707529827236
0.7	0.00152133163921680	0.00147291858197313	0.00146636271465708	0.00146793207735297
0.8	0.00152831225683735	0.00147541478247563	0.00146604878691788	0.00146822988027023
0.9	0.00153437856526002	0.00147554313576873	0.00146710787949013	0.00146706817276775

Table 5 (Learning rate) System Recommended learning rate for (1.0, 0.1, 0.11,0.115) with corresponding MSE at momentum factor 1.0, desired epochs = 10000 and iteration = 5

Momentum Factor (μ)	P1	P2	P3	P4
0.1	0.00146796375612452	0.00145528163042184	0.00145482558945875	0.00145543363974043
0.2	0.00145292013917912	0.00145524981984417	0.00145427976319767	0.00145355282409462
0.3	0.00145453586315322	0.00145349620176701	0.00145417634618025	0.00145436000125180
0.4	0.00145642118900505	0.00145489118242293	0.00145407732553584	0.00145471251330894
0.5	0.00146096123636455	0.00145450127853573	0.00145397521634886	0.00145571230364597
0.6	0.00146155512814190	0.00145446801452127	0.00145535912301330	0.00145398945895608
0.7	0.00146225220931068	0.00145440133444843	0.00145591008410766	0.00145563451077689
0.8	0.00146466037053799	0.00145522835466190	0.00145392476126321	0.00145443213306863
0.9	0.00146523738237253	0.00145502944590897	0.00145447988718255	0.00145663759446584

Table 6. System recommended momentum factor for (1.0, 0.2,0.23,0.238) with corresponding MSE at learning rate = 0.1151, desired epochs = 10000 and iteration = 5



Epoch count	MSE
1	1.63289262934093E-01
100	1.67082747834919E-03
1000	1.67029292416874E-03
10000	1.65516581866368E-03

Table 7. Optimized MSE.

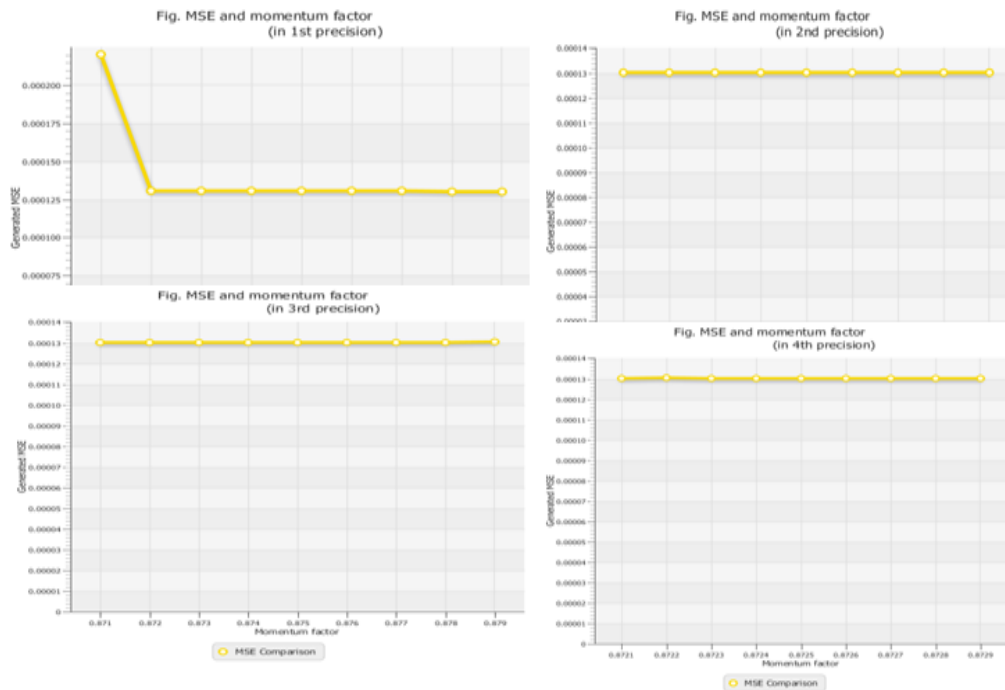


Fig. 2: Minimizing MSE.

IV. Results and Discussions

The performance of model is observed for training data set of rainfall. The testing data set are also inputted in developed model to check the performance of the model. The model trained with 45 years (1970–2014) dataset of vindhya region. Developed BPN model optimized all the required parameters on the basis of given input. System recommendations are as follows:

System Recommendation Values:

Optimized Parameter(s)	Recommended Values
1. Number of layer	3
2. Number of hidden layer	1
3. Number of neurons in hidden layer	2
4. Number of input vector (x_i s) 'n'	10
5. Learning rate (α)	0.1151
6. Momentum factor (μ)	0.2382
7. Initial weights (v_{ij} s & w_{ij} s)	37
8. Number of biases in hidden layer and values	2
9. Number of Biases in output layer and value	1
10. Transfer function	Sigmoid
11. MSE Level	1.65516581866368E-03
12. Number of Epochs	10000

The statistics of the performance of the BPN in training as well as in testing period is illustrated in Tables.1 to 4. From the fig.1 it is found that with the increase in learning rate, the MSE first decreases and reaches to an optimum value and then start increasing. MSE also found to decrease with the increase in momentum factor upto the optimum value after this it became stable as shown in the figure-2. The momentum factor is rather more stable at higher precision level.

Acknowledgements

The authors thank to MPCST, Bhopal for the financial support. The climate data are received from IMD Pune and also thanks to Department of Computer Science, BIT Durg for academic support.

References

- [1]. Geetha, G. and Selvaraj, R. S. (2011): Prediction of monthly rainfall in Chennai using back propagation neural network model, *Int. J. Eng. Sci. Tech.*, 3,211–213.
- [2]. Guhathakurta, P. (1998): A hybrid neural network model for long range prediction of all India summer monsoon rainfall, in *Proceedings of WMO International workshop on dynamical extended range forecasting*, Toulouse, France, November 17–21, 1997, PWPR No. 11, WMO/TD 881, 157–161.
- [3]. Guhathakurta, P., Rajeevan, M. and Thapliyal, V. (1999): Long range forecasting Indian summer monsoon rainfall by hybrid principal component neural network model, *Meteorol. Atmos. Phys.*, 71, 255-266.
- [4]. Guhathakurta, P. (2006): Long-range monsoon rainfall prediction of 2005 for the districts and sub-division Kerala with artificial neural network, *Curr. Sci. India*, 90, 773–779.
- [5]. Karmakar, S., Kowar, M. K. and Guhathakurta, P.(2009):Artificial neural network skeleton in deterministic forecast to recognize pattern of TMRF, *CSVTV Res. J.*,2(2), 41–45.
- [6]. Karmakar, S., Kowar, M. K. and Guhathakurta, P.(2012): Application of neural network in long range weather forecasting: In the context of smaller geographical region (i.e. Chhattisgarh State, India). *Lambert Academic Publishing, Germany*, 57–87.

- [7]. Krishnamurthy, Vand Kinter III, J. L.(2002): The Indian monsoon and its relation to global climate variability, in Global climate – Current researchuncertainties in the climate system, edited by Rodo, X. and Comin, F. A., Springer Berlin Heidelberg, 186–236.
- [8]. Krishnamurthy, V. and Kirtman, B. P. (2003): Variability of the Indian Ocean: Relation to monsoon and ENSO, Q. J. Roy. Meteor. Soc., 129, 1623–1646.
- [9]. Sahai, A. K., Grimm, A. M., Satyan. V. and Pant, G. B.(2002): Prospects of prediction of Indian summer monsoon rainfall using global SST anomalies, IITM Research Report No. RR-093.
- [10]. Shrivastava G. and Karmakar S. (2013) BPN model for long-range forecast of monsoon rainfall over a very small geographical region and its verification for 2012, Geofizika UDC 551.509.331 Volume 30.