Neural, Parallel & Scientific Computations 15 (2007) 539-546

Prediction Of Annual Rainfall By Double Fourier Series And Artificial Neural Network

[1]M. S. Kulshrestha and [2]R. K. George

[1]Assistant Professor (Maths), Department of Agricultural Meteorology, Anand Agricultural University, Anand-388110. Gujarat, India.
Email:kush122003@yahoo.co.in
[2] Reader in Applied Mathematics, Dept. of Applied Mathematics, Faculty of Technology & Engg., M S University of Baroda, Vadodara. Gujarat, India. Email: raju_k_george@yahoo.com

ABSTRACT: Rainfall of India in most of the part is erratic and oscillating year to year. Therefore, Double Fourier Series which contains sine and cosine terms are helpful for prediction of annual rainfall.

An attempt has been made here to predict the annual rainfall (ARF) by Double Fourier Series (DFS). DFS is a Mathematical method to establish functional relation between two inputs and one output. As a case study, annual rainfall (ARF) of the year 2006 of Anand station of the Gujarat State of the year 2006 was predicted. Forty-eight years (1958-2005) of annual rainfall (ARF) data series is used. Used input set was maximum air temperature of month of May (MAT) of the current year and previous year's annual rainfall (PARF) and output was to be occur annual rainfall of the current year. Prediction of annual rainfall by this method is significant to the actual occurred annual rainfall for the year 2002 to 2004. Predicted annual rainfall of the year 2006 (1348.9) mm, which was significant to actual rainfall (1413.8 mm).

Here, Annual rainfall also predicted by artificial neural network using same input data series and results of both the methods are compared by computing Root Mean Square (RMSE) and Percentage of Average Error (PAE). Results obtained by DFS were found better than artificial neural networks.

Related Programmes are developed in MATLAB.

1. INTRODUCTION

The problem of prediction of weather parameters like annual rainfall (ARF), is the primary concern in the field of Meteorology. Processes like rainfall in nature are chaotic in behavior (*Farmer et al.,* 1988) In such processes regularities like periodicities are mixed with noise. This results in large error for long time prediction.

A sum of daily rainfall during a year is known as Annual Rainfall (ARF).

Here, we deal with the problem of prediction of annual rainfall (ARF) at Anand station by using two different methods namely,

i) Double Fourier Series (DFS).

ii) Artificial Neural Network (ANN).

The Figure 1 depicts the known Data Series (DS) of ARF from 1958 to 2005. Figure shows that there is no vivid trend in the ARF series.

1061-5369 $15.00 © Dynamic Publishers, Inc.

We now look for major variables, which affect the amount of ARF. Singh *et al.* (1994) have studied the "Energy exchanges between ocean and atmosphere in relation to south-west monsoon over India" and provided the following conclusions:

(i) If higher evaporation occurs in the month of May over the southern belt of Indian seas (5°-15°N) then it is favorable for the ensuing southwest monsoon.

(ii) If a strong field of momentum flux occurs in the month of May over Indian bay then it generates good monsoon.

Rao *et al.* (1963) and Jagannathan *et al.* (1973) did extensive study to obtain trends of annual rainfall at various stations taking very large number of historical data. Chowdhury *et al.* (1981) have concluded that after the testing of the data series of 60 to 100 years for randomness, it is found that i) There is no trend in yearly, 5 yearly, 10 yearly mean rainfalls for a majority of stations but at some places 2 or 3 years period is found.ii) There is no short period of cycle in annual rainfall and in distribution of rainfall, particularly in arid and semi-arid regions of NW India. ii) There is no increasing or decreasing trend but only oscillation from year to year.

Anand station lies in the semi arid region of India and as per (ii) it has no significant trend for annual rainfall but oscillation occurs from year to year (Fig:1). So, we take the previous year ARF as another factor affecting the current year ARF.

We took two predictors namely May month's Maximum air temperature (MAT) and previous year Annual Rainfall (PARF) and current year Annual Rainfall (ARF) as the predictand. It is found that the correlation between MAT & ARF and PARF & ARF was 11.5% and 12.8 % respectively.

We look for non-linear models to represent the process. Here we tried two methods, namely: Double Fourier Series (DFS) and Artificial Neural Networks (ANNs).

2. DATA

The two input variables data used in the present problem of annual rainfall prediction were highest maximum May month's temperature (MAT) from the year 1959 to 2006 and previous year annual rainfall (PARF) from the year 1958 to 2005.

Figure 2 shows the plotting of 3-D graph of MAT, PARF and ARF data series from 1958 to 2005. This figure shows the continuous surface.

3. DETAILS OF THE METHODS

(a) DOUBLE FOURIER SERIES (DFS)

The theory of Fourier series, (Harmonic Analysis (HA)) is a well established branch of Mathematics. The tools of Harmonic Analysis are extensively used in many engineering fields (Lukaniszyn *et al.*, 2004). The Double Fourier Series is a particular case of the General Fourier Series. Here we consider the Fourier series representation of a function of two variables. Anita *et al.* (2003) have used "A Semi Lagrangian Double Fourier Method for the Shallow water Equations on the Sphere." Vande (2005) has used DFS as a Mapping Tool in Marine Cartography and it has been shown that DFS is a global model to map a projection process against coastal erosion.

We attempt to make use of Double Fourier Series (DFS) to represent the natural process of annual rainfall (ARF) as nature of the rainfall is oscillating (Fig:1).

Let x be the highest May month's maximum air temperature (MAT), y be the previous year annual rainfall (PARF) and z be the current year annual rainfall (ARF). Then by our assumption z is a function of x and y. That is, z=f(x,y).

(a-i) DEFINITION OF DOUBLE FOURIER SERIES (DFS)

DFS contains terms of sine and cosine in combination of sine sine, sine cosine, cosine sine, and cosine cosine .

Let f(x,y) be a function of two variables defined on a rectangle , K ((x,y): $0 < x < P_1$, $0 < y < P_2$) $\subset R^2$, or a function defined for all x and y with period P_1 in x and period P_2 in y. This chosen rectangle can be converted in to $X \in (0,\pi)$ *and* $Y \in (0,\pi)$ by making the substitutions $X=\frac{\pi}{P_1}x$ *and* $Y=\frac{\pi}{P_2}y$; where, $0 < x < P_1$ and $-0 < y < P_2$.

Then the function,

$f\left(\frac{P_1 X}{\pi},\frac{P_1 Y}{\pi}\right)=\phi(X,Y)$ has period in both X and Y.

Fourier series for this function f is given by $z=\phi\ (X,Y)=$

$$\sum_{m=0}^{\infty}\ \sum_{n=0}^{\infty}\lambda_{mn}[a_{mn}\cos mX\cos nY+b_{mn}\sin mX\cos nY+c_{mn}\cos mX\sin nY+d_{mn}\sin mX\sin nY].$$

$$\text{where, }\lambda_{mn}=\begin{cases}\frac{1}{4} & for\ m=n=0\\ \frac{1}{2} & for\ m>0,\ n=0\ or\ m=0,\ n>0\\ 1 & for\ m>0,\ n>0\end{cases}$$

the DFS co-efficient a_{mn}, b_{mn}, c_{mn} and d_{mn} are given by the following formulae.

$$a_{mn}=\frac{1}{\pi^2}\iint_K \phi(X,Y)\ \cos mX\ \cos nY\,dX\,dY,$$

$$b_{mn}=\frac{1}{\pi^2}\iint_K \phi(X,Y)\sin mX\ \cos nY\,dX\,dY,$$

$$c_{mn}=\frac{1}{\pi^2}\iint_K \phi(X,Y)\cos mX\ \sin nY\,dX\,dY$$

$$d_{mn}=\frac{1}{\pi^2}\iint_K \phi(X,Y)\sin mX\ \ \sin nY\ \ dX\,dY$$

The above series converges for a class of functions satisfying certain properties.

(a-c) METHOD IN PROBLEM SOLUTION:

To predict the ARF of the year 2002, DFS coefficients a_{mn}, b_{mn}, c_{mn} and d_{mn} were found by using the historical data namely MAT (1959 to 2001) and PARF (1958 to 2000). These coefficients are used in DFS method to predict the ARF of the year 2002.

In the same manner ARF of the year 2003 to 2006 are found (Table 1).

Computation is carried out in MATLAB and graphs are plotted.

(b) ARTIFICIAL NEURAL NETWORK (ANN) APPROACH

Hall *et al.* (1993), Hsu *et al.* (1993, 1995) have applied artificial neural network for rainfall- runoff modeling. Goswami *et al* (1996, 1997) have used ANN with three layers, namely, input layer, hidden layer and output layer for experimental forecasts of all India Summer Monsoon Rainfall for 2002 and 2003.

Here an attempt to represent the rainfall process in terms of a single –hidden layer Feed Forward Neural Network (Figure 3) was made.

(b-i) NETWORK ARCHITECTURE:

It was considered that an ANN consists of an input layer with three inputs namely, maximum air temperature of May month (MAT) , Previous year annual rainfall (PARF) and number of the year, one hidden layer and an output layer with one output namely, current year amount of annual rainfall (ARF). Number of nodes in the input and output layer is equal to the number of variables of inputs and output respectively. Each neuron is connected by feed forward network (Fig: 3).

In the present analysis number of neurons in hidden layer and number of epochs are given in the Table 2. Selected learning rate and momentum is 0.001 and 0.5.

In ANN input nodes have known values and these values passes to the next layer (hidden layer) after multiplying with the weight of the connection. Hidden neuron get these weighted inputs and applies a sigmoidal function to determine other neuron fires or remains dormant. These neurons group determines the importance of that particular input to the overall prediction. The sigmoid function is of S-shaped. Here, $F(x) = \frac{1}{1+\exp(-x)}$ is considered as activation function, where, x is the sum of all weighted inputs coming to the node, that is, x = netj

During the training, a learning algorithm namely, Back Propagation with momentum was used to iteratively modify the weights of the connections to minimize the total error in the approximation.

(b-ii) LEARNING OF THE NETWORK

Here learning is the type of supervised learning (Werner, 1994). In the network, supervisor is the observed output of ARF data series. Training is equally hard for larger as well as differently configured networks. The time required for training depends on the value of the parameters, like momentum, learning rate and error ratio. It also, depends on number of hidden neurons. It also appears that some times larger nets that are with more number of hidden nodes do not give any significant change in accuracy. Therefore, always it is desirable to decide by trial and error that how many numbers of hidden nodes are required. There is no standard method is developed but Chu *et al.* (1964) have proved that 'if the number of binary input cells is N (i.e. N-bit) for perceptron networks, the number of functional link cells that need to be generated to make the network learn successfully is at most 2^N –N-1'.

Here to train the networks, normalizations for all the inputs variables and actual output is done by dividing them by their norm. Training is done under the supervision of actual occurred ARF from 1960 to 2001. That means that training of the ANN is done by minimizing the error that is taking difference between actual ARF and output given by the ANN at each epoch. At given minimum error the training is stopped. Here computer error is 0.0007.This training required 22230 numbers of epochs (Table 2). To achieve the desired error, 147 numbers of hidden neurons is used. Less number of hidden nodes like 8, 10 or 15 does not give good performance to achieve the error goal and more then 147 numbers of hidden nodes took very large time to converge the aim or actual out put.

Value of the learning rate will be 0.1 or 0.01 enlarges the training period of ANN.

4. RESULTS AND DISCUSSION

(a) DOUBLE FOURIER SERIES (DFS)

The ARF is predicted by using DFS, selecting suitable values of m and n .It has been found that the predicted ARF for the years 2002 to 2006 have non-significant difference to the actual one. However the predicted ARF (750.9mm) of the year 2005 has large difference with occurred ARF (1693.0mm)

Therefore, the year 2005 prediction is eliminated from the computation of RMSE and PAE.

Fig 4 shows the prediction of the year 2002. ARF of 2002 to 2003 are 423.1mm and 928.7 mm respectively. These are one step a head prediction in the DS (Table 1). Predicted ARF for the year 2004 and 2006 are also found significant with actual rainfall . Table 1 shows the values of Root Mean Square Error (RMSE) was 56.28 and Percentage of Average Error 5.78 %. This shows that all the four prediction by DFS are significant.

It is concluded that

(i) Predicted annual rainfalls of the year 2002 to 2006 (excluding the year 2005) were significant to the actual rainfall.

ii) Computation time was less in DFS approach.

(b) ARTIFICIAL NEURAL NETWORK

From the student's t-test for two tails at 95% of confidence interval predicted ARF by ANN, is found to be significant to the actual rainfall except for the year 2005. For the outliers like 1693.0 mm in the year of 2005 in the data series ANN is unable to predict the ARF accurately. All the predicted values are shown in the Fig 5 and Table 3. Predicted Annual Rainfall (ARF) of the year 2002 is 520.9mm and actual 479.2mm.Difference between these two ARF is 41.7mm.RMSE and PAE are computed for the four years (Table 2). Found RMSE was 63.01 and PAE is 6.5 %.

Predicting ARF of the year 2005, training of the NNs becomes difficult due to extreme ARF that is 1693.0mm. Estimated ARF is found 1319.5 mm. Here, error goal is 0.00089. If we increases the error goal or increase the numbers of hidden nodes, ANN is not convergent and predicting the non-significant ARF.

From the predicted ARF of the year 2006 predicted ARF is 1485.1mm (Table 2; Fig: 5). Actual ARF is 1428.7 mm.

It is concluded that

i) The results show the applications of ANNs to the rainfall analysis for predictions of ARF of the years 2002 to 2006 by ANN is significant to actual annual rainfall except 2005.

ii) Prediction has been done for dependent random variable having non -linear relationships with predictors.

iii) ANN gives freedom to use multi-inputs and multi outputs without using any Mathematical or statistical model.

Prediction of ARF by two methods, ANNs and DFS are mentioned in the Table 3. Here comparison between two methods has been done by obtaining their Root Mean Square Error (RMSE). Percentage of Average Error (PAE). Computed RMSE and PAE for ANNs & DFS method are 63.011 & 56.28 and 6.5 & 5.78 (Table 3) respectively.

Comparing these two errors of two methods it is found that DFS model gives less error in comparison to ANNs. But that difference is not very large.

Prediction of annual rainfall (ARF) for 2002 to 2006 is found significant by both DFS and ANN methods except for outliers (2005).

Obtained PAE By both the methods are less than 10 % and therefore, predictions are significant.

Further, predictions are also checked by Student t -test for two tails gives non - significant difference with actual Annual Rainfall (ARF).

5. ACKNOWLEDGEMENT:

This work is not become possible without the guidance of Dr A M Shekh, principal & Dean of the Agriculture faculty, Anand Agricultural University, Anand and providing data and necessary facilities by Department of Agricultural Meteorology, Anand Agricultural University, Anand

Table 1 Details of the parameter values used in DFS

Sr. No.	Predicting Year	Number of Data Point	m and n values m	n	Predicted ARF P_i (mm)	Actual ARF A_i (mm)	$\|P_i-A_i\|$ (mm)	RMSE	PAE (%)
1	2002	43	21	12	423.0	479.2	56.2		
2	2003	44	22	9	1175.4	1135.4	40.0		
3	2004	45	24	12	928.7	866.0	92.4	56.28	5.78
4	2005	46	23	12	*750.9.	1693.0	*720.7		
5	2006	47	23	15	1348.9	1428.7	79.8		

- Excluded from computation of RMSE and PAE.

Table 2 Details of the parameter values used in ANN training.

Sr. No.	Predicting Year	Number of epochs used	No. of neurons	Error goal	Predicted ARF I (mm)	Actual ARF II (mm)	Difference |I-II| (mm)	RMSE	PAE (%)
1	2002	22230	147	0.0007	0520.9	0479.2	41.7		
2	2003	237540	147	0.00001	1054.0	1135.4	081.4		
3	2004	103684	147	0.00008	0816.4	0866.0	049.6	63.01	6.5
4	2005	18362	160	0.00089	*1319.5	1693.0	*373.5		
5	2006	5942	147	0.0075	1485.1	1428.7	56.4	------	

- NOT INCLUDED IN THE COMPUTATION OF RMSE AND PAE.

Table 3 Comparison of the results by two methods DFS and ANN

Sr. No	Year	Predicted ARF (mm) By ANN	By DFS	Actual ARf (mm)	Difference with Actual ARF(mm) By ANN	By DFS	RMSE (mm) By ANN	By DFS	PAE By ANN	By DFS
1.	2002	520.9	478.1	479.2	41.9	1.1				
2.	2003	1054.0	1165.4	1135.4	81.4	30.0				
3.	2004	913.6	857.4	866.0	47.6	8.9	63.01	56.28	6.5	5.78
4.	2005	*1319.5	*1152.4	1693.0	* 373.5	*540.6				
5.	2006	1485.1	1413.8	1428.5	56.4	14.9				

- NOT INCLUDED IN THE COMPUTATION OF RMSE AND PAE.

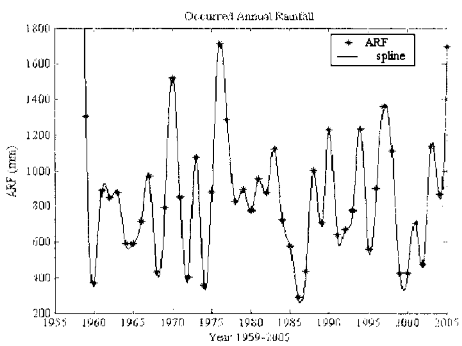

Fig:1

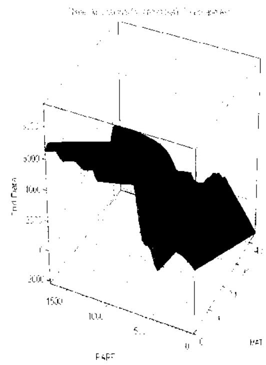

Fig:2

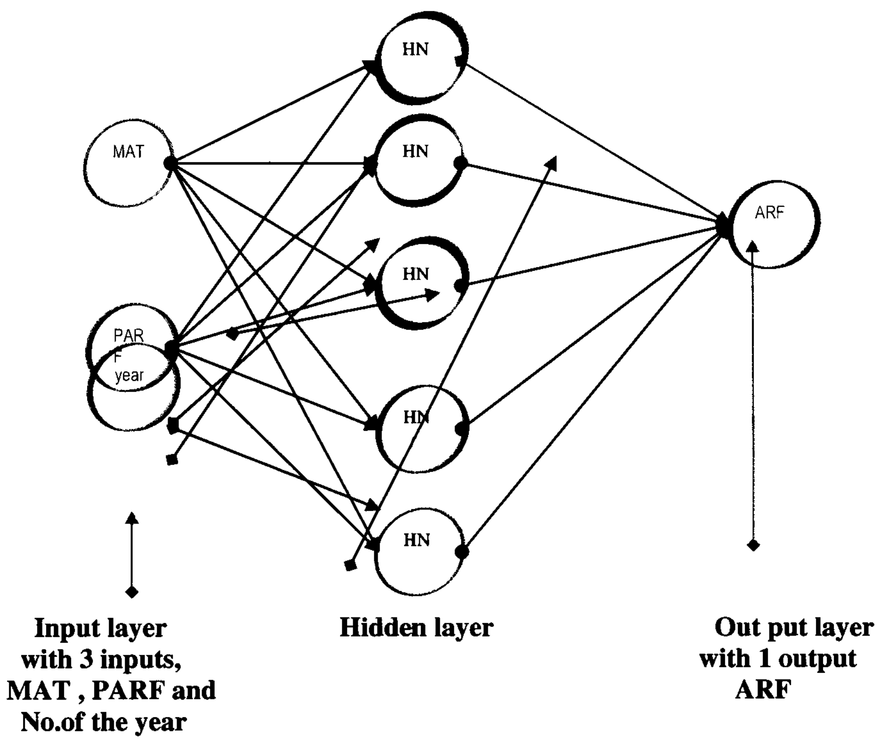

Fig:3

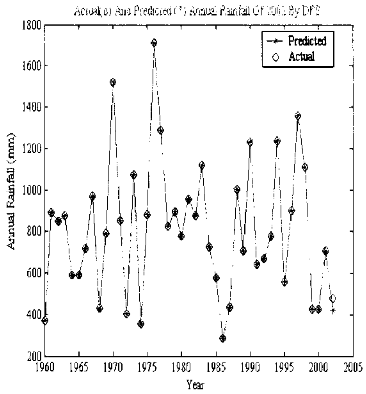

Fig: 4

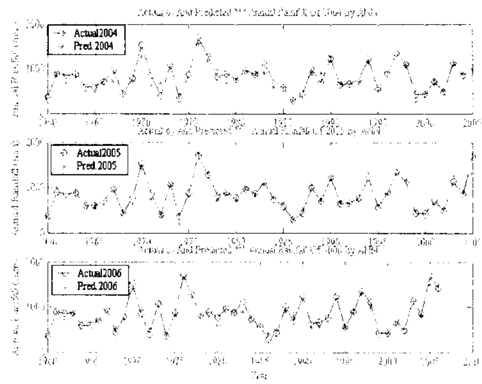

Fig: 5

REFERENCE

[1] Anita, T. L and William, F. S., A semi lagrangian double fourier method for the Shallow water Equations on the Sphere". Journal of Computational Physics. 189 1 (2003) 180-196.

[2]Chowdhury,A and Ganesan,H.R., Man and Climatic Variability. Vayu Mandal, Bulletin of the India Meteorological Society. January-June, (1981) 9-13.

[3] Chu, Y. P. and Hsieh, C M. An artificial neural network model with modified perceptron algorithm, IEEE, 93h , 68120. Parallel Comput. 18, 1992 9 (1964) 983-996.

[4] Farmer, J. D. and Sidorowich, J. J., Exploiting chaos to predict the future and reduce noise. LA-UR, (1988) 88-901.

[5] Goswami, P. and Pradeep Kumar, Experimental annual forecast of all India mean summer monsoon rainfall for 1997 using a neural network model. Curr. Sci. 72, (1997) 781-782.

[6] Goswami, P. and Srividya, A novel neural network design for long range prediction of rainfall pattern.Curr. Sci. 35, 70 (1996) 447-457.

[7] Hall, M. J. and Minns, A. W., Rainfall - runoff modeling as a problem in artificial intelligence experience with a neural network. Proc. 4th National Hydrology Symposium, British Hydrological Soc., Credit, (1993) 5.51-5.57

[8] Hsu, K.L., Gupta, H.V. and Sorooshian, S. *"Artificial Neural Network modeling of the rainfall-runoff process"*. Water Resources Research, 29 4 (1993) 1185-1194

[9] Hsu, K.L., Gupta, H.V. and Sorooshian, S. Artificial neural network modeling of the rainfall-runoff process". Water Resources Research, 31 (1995) 2517-2530

[10] Jagannathan, P. and Parthsarthy,B., Trends and periodicities of rainfall over India." Mon.Weather Rew, 101(1973) 371-375.

[11] Lukaniszyn, M L; M Jagiela; R Wrobel; K Latawiec. 2D harmonic Analysis of the cogging torque in synchronuous parmanent magnet machines. COMPEL . Int J For Computation and Maths.in Electrical Eng., Volume 23 (3) (2004) 774-782.

[12] Pant, G. B. and Rupa Kumar, K., "Climates of South Asia." http.//www.wiley.com John Wiley & Sons Ltd. Baffins Lane, Chichester, West Sussex PO19 1UD, England (1997) 133-176

[13] Parthsarthy,B. and Sonatake, N.A., " El-Ninao/SST of Puerto Chicama and Indian summer monsoon rainfall-statistical relationships." Geofis.Int., 27 (1988) 37-59.

[14] Rao, K. N. and Jagannathan, P. "Climatic changes in India." Proc. Synop. Changes of climates. UNESCO, (1963) 53-66.

[15] Shukla , J. and Paolino, A. The southern oscillation and long-range forecasting of the summer monsoon rainfall over india." Monthly Weather Review,111 9 (1983) 1830–1837

[16] Singh,O.P. and Joshi,S.P., and Sundari,G., Pressure fields over Indian seas and their relationships with Indian summer monsoon rainfall, Mausam,45,2 (1994)149- 154.

[17] Vande, W., "Double Fourier Series as a Mapp.ing Tool in Marine Cartography Analyses and Use." Journal of Geographic Information and Decision Analysis, 4 1 May (2005) 29-44.

[18] Walker, G.T., Correlations in seasonal variations of weather, VIII. A preliminary study of world wether I". memories India Meteorology Department., 24 (1923) 275-332

[19] Werner, K., *"Neural Network"*. Galgotia Publication,New delhi, India, (1994).