ICIWRM – 2000, Proceedings of International Conference on Integrated Water Resources Management for Sustainable Development, 19 – 21 December, 2000, New Delhi, India

Temporal disaggregation of rainfall data using artificial neural networks

D. NAGESH KUMAR
Department of Civil Engineering, Indian Institute of Technology, Kharagpur – 271302, India
K. SRINIVASA RAJU
Assistant Professor, Civil Engineering Department, Birla Institute of Technology and Science, Pilani 333 031, Rajasthan, India

Abstract

Seasonal rainfall in many regions is influenced by global climatic parameters such as El Nino Southern Oscillation (ENSO), La Nina. Seasonal rainfall predictions are based on such Global climatic parameters. However, for operational purposes, such as reservoir operation or river basin management in a larger context, rainfall data at a finer time interval are required. In this paper, a temporal disaggregation model based on Artificial Neural Networks is presented for obtaining rainfall for monthly or shorter time periods from the given seasonal rainfall prediction. A feed forward neural network with back-propagation algorithm for learning was used. A sigmoidal function was used for neuron activation. The training error (RMS error) was measured by squaring the difference between the network and training pattern desired output and summing over all outputs and all training patterns. The methodology is applied to obtain monthly rainfall data from Indian monsoon (June-September season) rainfall data for a sub-division for which the relation between climatic indices and monsoon rainfall was already established. Monthly rainfall data for 124 years (1871-1994) for Orissa state, India was used for model application. Rainfall data for the first 100 years was used for ANN training and remaining 24 years data was used for testing. The disaggregated monthly rainfall data compared well with observed monthly data. In addition they preserved all the basic statistics such as summing to the seasonal value, cross correlation structure among monthly flows.

INTRODUCTION

It is well established that global climatic parameters like El Nino, La Nina have good influence on Indian monsoon (Krishna Kumar et al, 1999, Rasmusson and Carpenter, 1983). Monsoon rainfall (seasonal) is predicted by considering these factors at regular intervals (Krishna Kumar et al, 1995). Recently many attempts have been made to make prediction of seasonal rainfall using climatic indices such as ENSO, PDO (Pacific inter Decadal Oscillation), and NAO (North Atlantic Oscillation) index among others. However, for operational purposes, such as reservoir operation or river basin management in a larger context, rainfall data at a finer time interval (say monthly or weekly) are required. There is renewed interest in disaggregation methods as climate related issues (regional ENSO forecasts or downscaling of Climate Change Scenarios) have come to the fore.

The disaggregation models proposed by Valencia and Schaake (1973) have been used to divide annual flows into seasonal flows (Salas et al., 1980) and to divide aggregate basin flows (monthly or annual) into flows at individual sites (Salas et al., 1980). Stedinger and

Vogel (1984) further extended this model to reproduce the correlation between disaggregated flow volumes at different time scales. By and large, these approaches have focused on space or time disaggregation and on annual to seasonal or seasonal to sub-seasonal data. Parametric assumptions of the probability distribution of the underlying streamflow are usually invoked. Recently nonparametric approaches are presented for space or time disaggregation of hydrologic data. Lall et al (1996) and Tarboton et al. (1998), proposed a nonparametric approach for space or time disaggregation based on kernel density estimation. A new algorithm for simultaneously disaggregating monthly to weekly or daily flows at a number of sites on a drainage network is presented in Nagesh Kumar et al. (1999, 2000) by considering K-Nearest Neighbors (KNN) and optimization.

In the present paper, a non conventional approach, Artificial Neural Networks, is used for temporally disaggregating predicted seasonal rainfall data into monthly or shorter time interval rainfall data.

ARTIFICIAL NEURAL NETWORKS

In this study, it is assumed that seasonal forecasts of rainfall (monsoon) are available. The seasonal rainfall data is disaggregated into monthly or shorter time intervals in this study. For this purpose Artificial Neural Networks (ANN) with modified back propagation learning algorithm were used. Artificial Neural Networks are briefly explained before their adoption to the specific problem.

A neural network in its basic form is composed of several layers of neurons; an input layer, one or more hidden layers and an output layer. Each layer of neurons receives its input from the previous layer or from the network input. The output of each neuron feeds the next layer or the output of the network. This is illustrated in Figure 1, which shows a 3-layers NN (2x3x1). The first layer is an input layer that distributes the inputs to the hidden layer and does not have any activation function. The lines connecting the neurons represent the weights. Also shown in the figure are the bias nodes that are used to shift the neuron transfer function and improves the network performance.

Mathematically the network computes:

1. The output of the hidden layer:

$$s2_j = \sum_i W3_{i,j}\ i_i \qquad\qquad o_j = f(s2_j)$$

2. For the output layer calculate:

$$s3_k = \sum_j W3_{j,k}\ h_j \qquad\qquad o_k = f(s3_k)$$

where

i_i are the network inputs.

o_k are the network outputs.

$W2_{i,j}$ represents the weight connecting neutron i in layer 1 to neutron j in layer 2.
$W3_{j,k}$ represents the weight connecting neutron j in layer 2 to neutron k in layer 3.
$f(x)$ is the neuron transfer function which could be liners, sigmoid or hyperbolic Tan etc.

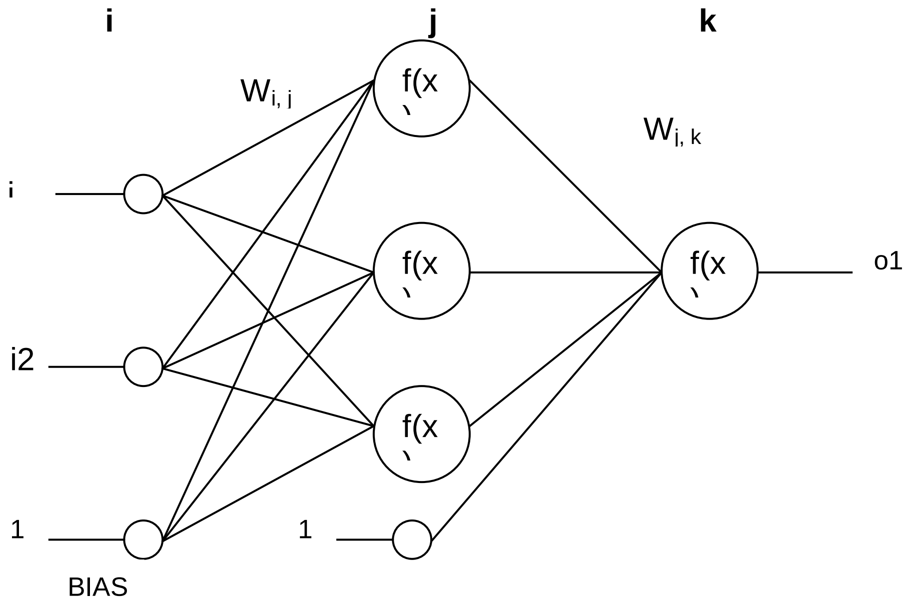

Figure 1. A simple 2x3x1 Neural Network.

In this study sigmoid function is used which is as follows:

$$f(x) = \frac{1}{(1+e^{-x})}$$

Training such a network involves using a data-base of examples which are values for the input and output of the NN. The NN would learn by adjusting the weights to minimize the error of the outputs. The error function is the objective of the minimization procedure and defined as:

$$RMS\,error = \sqrt{\frac{\sum_{p=1}^{\hat{p}}\sum_{k=1}^{\hat{k}}(t_k^p - o_k^p)^2}{\hat{p}*\hat{k}}}$$

where
$o^p{}_k$ is the NN output from neuron k for pattern p.
$t^p{}_k$ is the observed output for training pattern p for neuron k.
In this equation, the summations are over the number of training patterns and the number output neurons.

Back propagation algorithm is used to adjust the weights in an iterative manner i.e., the weight adjustments are summed over all the training patterns in an epoch and then the weights are adjusted. New weights are a function of the derivatives and previous weights, for example

$$\Delta W_{i,j}^{t} = \eta \, dW_{i,j}^{t-1} + \alpha \, \Delta W_{i,j}^{t-1}$$

$$W_{i,j}^{t} = W_{i,j}^{t-1} + \Delta W_{i,j}^{t-1}$$

where η is the learning parameter and α is the momentum. η and α directly relate to the back propagation learning algorithm. In addition random noise is added to both inputs and weights. Adding a random input noise to each input node makes the trained network less sensitive to changes in the input values and can help avoid local minima. Similarly adding a random input noise to each weight makes the network weights constantly "shake" as the training progresses. The purpose of this noise is to help the network jump out form a gradient direction that leads to local minima in the weight surface.

ANN offer the following advantages:
The application of ANN does not require apriori knowledge of the underlying process.
All the existing complex relationships between various aspects of the process under investigation need not be known.
ANN are data driven when compared to conventional approaches which are model driven.

Recently ANN were adopted successfully for various applications in water resources engineering. Schmuller (1990) indicated several potential environmental applications. Aziz and Wong (1992) described the use of neural networks to determine the aquifer parameters in aquifer hydraulics. Smith and Eli (1995) and Lorrai and Sechi (1995) utilized back propagation neural networks for modeling rainfall-runoff process. However, Hsu et al. (1995) also modeled rainfall-runoff processes using linear least squares simplex algorithm. Karunanithi et al. (1994) used neural networks for river flow prediction. In the present work neural networks are employed for temporal disaggregation of seasonal rainfall data for the first time.

PROBLEM STATEMENT

The need for temporal disaggregation of rainfall data is already established. In the present study ANN are used to disaggregate the given or forecasted seasonal rainfall data into monthly rainfall data for that season. Monsoon period in India is considered as season comprising of June, July August and September (JJAS) months of the year. So ANN are employed to disaggregate the monsoon (referred as season) rainfall data into rainfall in the four months. For this purpose a feed forward neural network with one input neuron (seasonal rain) for the input layer and four output neurons (corresponding to 4 months, JJAS) with number of neurons in one or two hidden layers is used. The neural network is trained by giving historic data of both seasonal and monthly rainfall data and improved back propagation algorithm, as explained in the previous section. Once the training is complete, the network weights are frozen and the network is used to disaggregate the given or predicted seasonal rainfall data into JJAS months. The results thus obtained are compared with the observed monthly rainfall for the corresponding period.

APPLICATION AND RESULTS

Rainfall data for Orissa state, India, is used for demonstrating the disaggregation approach. All though rainfall data is available for all the 24 meteorological divisions of India (Parthasarathy et al., 1994) Orissa division is selected for case study as it is one of the few divisions whose rainfall data had good relation with global climatic parameters (Rajagopalan and Krishna Kumar, 1999). In addition, first author is developing operational models for the existing projects in Orissa and it is proposed to utilize the results of this study for that purpose.

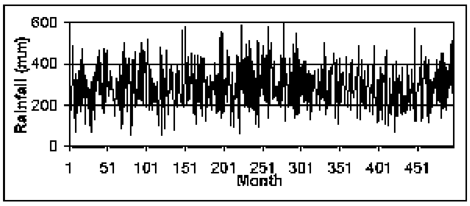

Figure 2. Monthly Rainfall Series for Monsoon Months (JJAS) of Orissa Division for 124 years (1871-1994)

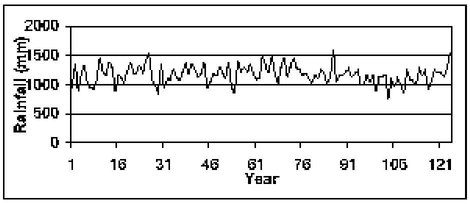

Figure 3. Seasonal Rainfall Series for Monsoon Period (JJAS) of Orissa Division for 124 years (1871-1994)

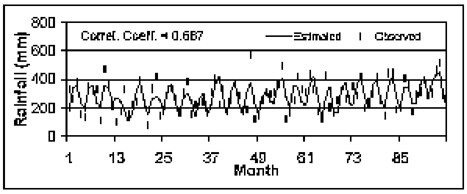

Figure 4. Comparison of observed and disaggregated rainfall using ANN (1,10,4) for the 4 monsoon months for 24 years (1971-1994)

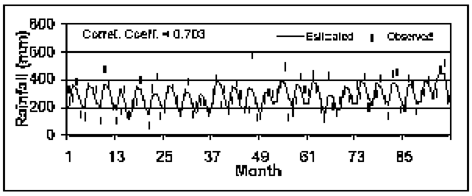

Figure 5. Comparison of observed and disaggregated rainfall using ANN (1,5,5,4) for the 4 monsoon months for 24 years (1971-1994)

Monthly Rainfall data is available for 124 years (Parthasarathy et al., 1994) from 1871-1994. As mentioned earlier only monsoon rainfall is considered for disaggregation. Monthly rainfall data for monsoon months (JJAS) for all the 124 years is presented as time series in Figure 2. Seasonal rainfall data for the 124 years is presented as time series for the 124 years in Figure 3. Although it is not clearly distinguishable, long term cyclicity of monthly data may be noticed. So hundred years of rainfall data is taken for training and the remaining 24 years data is used for testing. Of the several network configurations tried, results from two network configurations are presented here. They are (i) ANN with one hidden layer comprising of 10 neurons denoted as ANN(1,10,4) and (ii) ANN with one hidden layer comprising of 10 neurons denoted as ANN(1,5,5,4). After the network is successfully trained the resulting weights are frozen and are used for disaggregating the seasonal rainfall data of the remaining 24 years (or 96 months). A comparison of disaggregated rainfall data using ANN(1,10,4) with the corresponding observed rainfall data is shown in Figure 4. Correlation coefficient between the observed and disaggregated rainfall is 0.687. A better comparison can be noticed with results obtained from ANN(1,5,5,4) as shown in Figure 5. Here the correlation coefficient is 0.703.

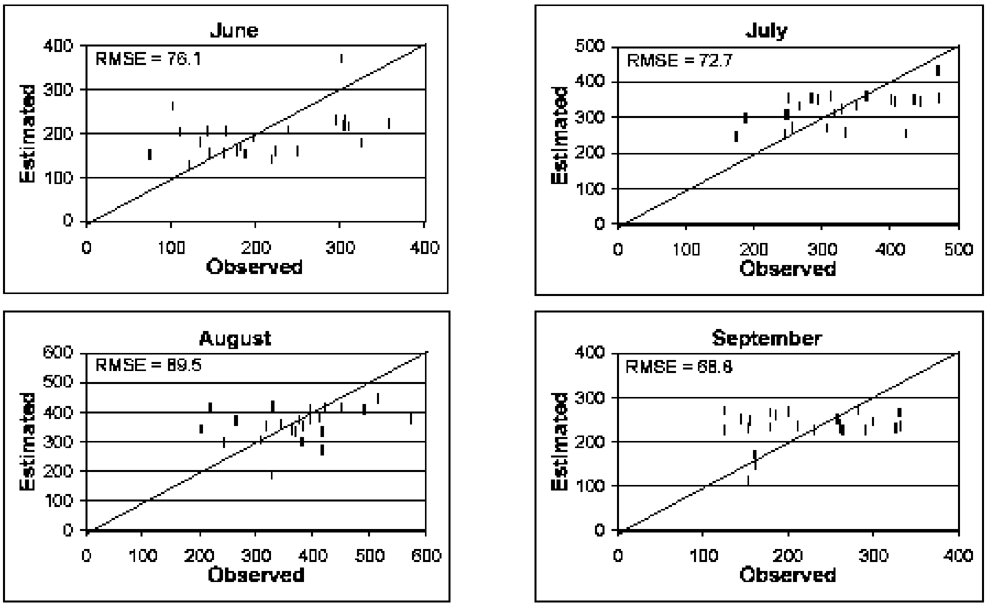

Figure 6. Scatter Plot of Observed versus Disaggregated Rainfall (in mm) using ANN (1,10,4) for each Monsoon Month for 24 years (1971-1994)

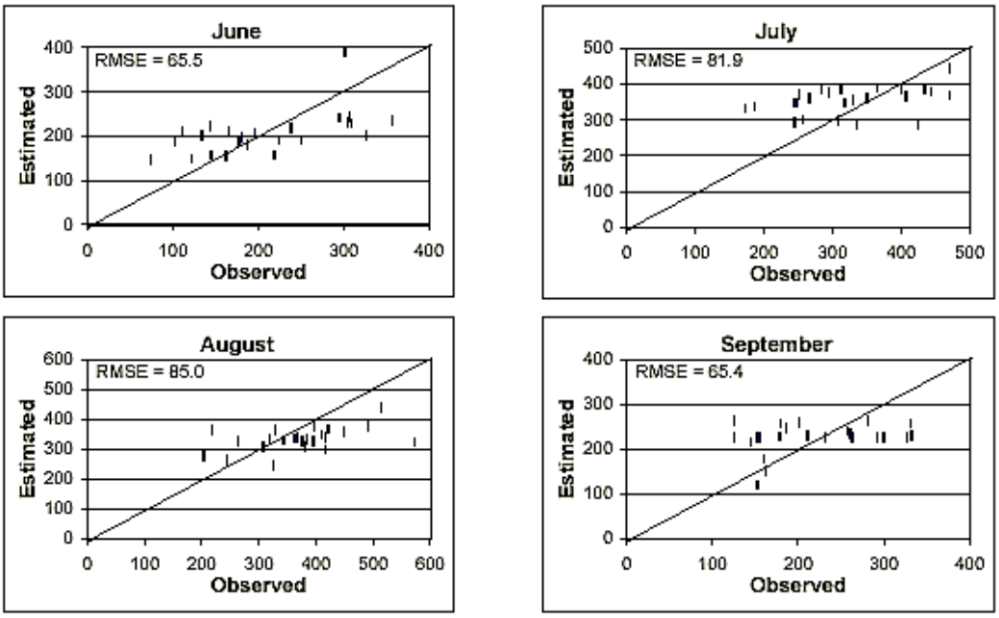

Figure 7. Scatter Plot of Observed versus Disaggregated Rainfall (in mm) using ANN (1,5,5,4) for each Monsoon Month for 24 years (1971-1994)

To further analyze the results obtained, observed and disaggregated data are compared for each month (JJAS) in a scatter plot. Figure 6 shows this comparison for each month for the results obtained from ANN(1,10,4). Similar scatter plots for the results obtained from ANN(1,5,5,4) are shown in Figure7. Root Mean Square Error (RMSE) between observed and disaggregated rainfall data for each month are also shown in these figures. As can be seen from Figures 6 and 7, ANN(1,5,5,4) performed better in disaggregating the seasonal rainfall data.

CONCLUSIONS

Feed forward neural networks with back propagation algorithm are employed for disaggregating the seasonal rainfall data into shorter time intervals (months). This approach is used to disaggregate the monsoon rainfall of Orissa division, India. Preliminary results obtained from the approach are quite encouraging. This approach can be used for short term planning and operation of river basin projects and development activities from the qualitative prediction of seasonal rainfall data based on global climatic parameters. This will further enhance the planning for sustainable development of river basin.

Acknowledgements
The first author wishes to thank Department of Science and Technology, Govt. of India for awarding Young Scientist project entitled “Artificial Neural Networks and their Adaptability to Stochastic Hydrology” (Project no. HR/OY/O-02/97) under which this research work is conducted.

References

Aziz, A. R.A., and K.-F.V. Wong, (1992) A neural-network approach to the determination of aquifer parameters, Ground Water, 30, 164–166.

Hsu, K.-L., H. V. Gupta, and S. Sorooshian, (1995) Artificial neural network modeling of the rainfall-runoff process, Water Resour. Res., 31, 2517–2530.

Karunanithi, N., W. J. Grenney, D. Whitley, and K. Bovee, (1994) Neural networks for river flow prediction, J. Comput. Civ. Eng., 8, 201–220.

Krishna Kumar, K., M.K. Soman, and K. Rupa Kumar (1995), “Seasonal forecasting of Indian summer monsoon rainfall: A review”, Weather, 50, pp. 449-466.

Krishna Kumar, K., R. Leeman, M.A. Cane and B. Rajagopalan (1999), “Epochal changes in Indian monsoon – ENSO precursors”, Geophysical Research Letters, 26(1), pp. 75-78.

Koutsoyiannis, D. (1992) A nonlinear disaggregation method with a reduced parameter set for simulation of hydrologic series. Water Resour. Res., 28(12), pp. 3175-3191.

Koutsoyiannis, D., and A. Manetas (1996) Simple disaggregation by accurate adjusting procedures. Water Resour. Res., 32(7) , pp. 2105-2117.

Lall, U., B.Rajagopalan, and D. G. Tarboton, (1996) A nonparametric wet/dry spell model for resampling daily precipitation. Water Resour. Res., 32(9), pp. 2803-2823.

Lorrai, M., and G. M. Sechi, (1995) Neural nets for modelling rainfall-runoff transformations, Water Resour. Manage., 9, 299–313.

Nagesh Kumar, D., U. Lall and M. Peterson, (1999) “Space-time disaggregation of streamflow data using k-nearest neighbor patterns and optimization”, Proceedings of Xth World Water Congress, Melbourne, Australia, in CD-ROM.

Nagesh Kumar, D., U.Lall and M. Peterson, (2000) Multi-site Disaggregation of Monthly to Daily Streamflow. Water Resour. Res., 36(7), pp. 1823-1833.

Parthasarathy, B, A.A. Munot and D.R. Kothawale (1994), “All-Indian monthly and seasonal rainfall series: 1871-1993”, Theoretical and Applied Climatology, 49, pp.217-224.

Rajagopalan, B, Krishna Kumar, K., (1999), “Personal Communication”, IRRI, Columbia Univ, New York, USA.

Rasmusson, E.M., and T.H. Carpenter (1983), “The relationship between eastern equatorial Pacific seas surface temperature and rainfall over Indian and Sri Lanka”, Monthly Weather Review, 111, pp. 517-528.

Salas, J. D., J. W. Delleur, V. Yevjevich, and W. L. Lane (1980) Applied Modeling of Hydrologic Time Series. Water Resour., Littleton, Colorado.

Schmuller, J., (1990) Neural networks and environmental applications, in Expert Systems for Environmental Applications, edited by J. M. Hus-hon, pp. 52–68, Am. Chem. Soc.

Smith, J., and R. Eli, (1995) Neural-network models of rainfall-runoff process, J. Water Resour. Plan. Manage., ASCE, 121, 499–508.

Srikanthan, R., Sequential generation of monthly streamflows. J. Hydrol., 38, pp. 71–80, 1978.

Stedinger, J. R., and R. M. Vogel (1984) Disaggregation procedures for generating serially correlated flow vectors. Water Resour. Res., 20(1), pp. 47–56.

Tarboton, D. G., A. Sharma and U. Lall (1998) Disaggregation procedures for stochastic hydrology based on nonparametric density estimation. Water Resour. Res., 34(1), pp. 107-119.

Valencia, D. R., and J. L. Schaake, Jr. (1973) Disaggregation processes in stochastic hydrology. Water Resour. Res., 9(3), pp. 580-585.