ISSN (e): 2250 – 3005 || Volume, 06 || Issue, 07|| July – 2016 ||
International Journal of Computational Engineering Research (IJCER)

Performance improvement of a Rainfall Prediction Model using Particle Swarm Optimization

Sirajul Islam[1], Bipul Talukdar[2]
1 Ex- Research Scholar, Civil Engineering Department, Assam Engineering College, Guwahati
2 Associate Professor, Civil Engineering Department, Assam Engineering College, Guwahati

ABSTRACT
The performances of the statistical methods of time series forecast can be improved by precise selection of their parameters. Various techniques are being applied to improve the modeling accuracy of these models. Particle swarm optimization is one such technique which can be conveniently used to determine the model parameters accurately. This robust optimization technique has already been applied to improve the performance of artificial neural networks for time series prediction. This study uses particle swarm optimization technique to determine the parameters of an exponential autoregressive model for time series prediction. The model is applied for annual rainfall prediction and it shows a fairly good performance in comparison to the statistical ARIMA model.

Keywords: *Particle swarm optimization, exponential autoregression, nonlinear regression, ARIMA*

I. INTRODUCTION

Linear time series models have drawn much attention due to their relative simplicity in understanding and implementation. The autoregressive integrated moving average (ARIMA) model [1] is one such model that has seen many applications in time series forecasting [2, 3, 4]. However, many practical time series including rainfall show nonlinear behaviour due to which nonlinear methods are employed for their prediction. A host of nonlinear statistical models have been described in the literature for predicting volatility changes in time series [5]. Artificial neural networks (ANNs) approach has been suggested as an alternative technique for nonlinear time series forecasting and it has gained immense popularity in the recent time [6, 7, 8]. Support vector regression (SVR) is another new approach being successfully applied in time series prediction [9]. Various hybrid techniques have also been tried in the recent years for efficient prediction. Particle swarm optimization (PSO) approach, in particular, has been utilized in combination with different time series models to improve the performances of these models. PSO has been effectively used by some researchers as an alternative to the Backpropagation (BP) algorithm for training ANN models [10, 11]. Asadi *et al.* [12] combined PSO with ARIMA model and reported that this hybrid method exhibited better prediction results compared to an ARIMA model itself. Cui and Jiang [13] used a binary Particle Swarm Optimization (BPSO) to increase the predictive accuracy of a local linear time series prediction model.
In the present study, an updated PSO is employed for determining the parameters of an exponential regression model so as to improve its prediction accuracy for annual rainfall prediction.

II. PARTICLE SWARM OPTIMIZATION

PSO is a metaheuristic optimization technique originally proposed by Kennedy and Eberhart [14]. This algorithm is inspired by the social behaviour of animals in swarms like bird flock and fish schools. The general idea in the PSO algorithm is that there exists a swarm of particles and each particle resides at a position x_i in the search space. The particles move over the search space with a certain velocity v_i. The velocity is influenced by the global best (*gbest*) position p_{gi} and the best position p_i, a particle has personally found (*pbest*). The velocity is updated and a new position is obtained iteratively using the following equations:

$$v_{i,d}(t+1) = \omega v_{i,d}(t) + r_1 c_1 (p_i(t) - x_{i,d}(t)) + r_2 c_2 (p_{g_i}(t) - x_i(t)) \quad (1)$$

$$x_{i,d}(t+1) = v_{i,d}(t) + x_{i,d}(t) \quad (2)$$

where, ω is the inertial weight; c_1 and c_2 are respectively called the cognitive and social factors; r_1 and r_2 are random numbers in the interval (0,1); t, i and d are respectively the iteration, particle and variable indices. However, the above method of dimension by dimension update of velocity was reported to be biased and leading to premature convergence of the algorithm [15]. Hence the above equations were modified in a geometrical way by Clerc [16] in the following manner.

A centre of gravity G_i is defined in the search space such that,

$$G_i(t)\begin{cases} = 0.5(x_i(t) + (c_1(p_i(t) - x_i(t)), & \text{if } p_t(t) = p_{g_i}(t) \\ = x_i(t) + 1/3(c_1(p_i(t) - x_i(t)) + c_2(p_{g_i}(t) - x_i(t))), & \text{otherwise}\end{cases} \tag{3}$$

A random point x'_i is then generated in the hyper sphere $H_i(G_i, ||G_i - x_i||)$ and the velocity update function is obtained as follows:

$$x_{i,d}(t) = \omega\, v_{i,d}(t) + x'_i(t) - x_{i,d}(t) \tag{4}$$

In this updated method, the value of the parameter ω is taken as 1/ 2 log 2 (≈ 0.721) and both c_1 and c_2 have the same value of 0.5 + log2 (≈1.193).

III. PSO-EXPAR MODEL FOR RAINFALL PREDICTION

Among the different statistical methods for time series prediction, ARIMA model is by far the most popular one. However, this method has certain limitations including its inherent assumption of linearity. In most cases, a rainfall time series is nonlinear in nature. A nonlinear model namely, the exponential autoregressive (EXPAR) model [17], is used in this study to predict an annual rainfall time series. This model is capable of handling the non- Gaussian characteristics of a time series [18]. This model can be represented as:

$$y_t = \sum_{i=1}^{p}\left\{\alpha_i + \varphi_i \exp\left(\text{-}\lambda\text{-}\ {}_{t-1}^{2}\right)\right\} y_{t-i} + \varepsilon_t \tag{5}$$

where y_i (i = 1, 2,3......p) is a vector of the predictor variables and p is the order of the model; λ is a scaling constant in the range (0,1) and ε_t is a white noise operator with mean zero and variance one. In the above equation, if y_{t-1} is large then the model turns into an autoregressive model. The coefficients α_i and ϕ_i are linear and hence the model may be termed as linear –in-the-parameters.

Selection of appropriate model parameters is most crucial for efficient prediction of a time series. Further, choosing the best performing model (i.e. choosing the order p in this case) is a cumbersome task. In this study, the robustness of PSO algorithm is used to accomplish both the above tasks. The model parameters are iteratively optimized using the updated PSO algorithm with the objective of minimizing the sum of squared error (SSE) between the observed and predicted values of annual rainfall. The order p of the model is chosen on the basis of Akaike information criterion (AIC) [19] and Bayesian information criterion (BIC) [20].

IV. EXPERIMENTAL RESULTS

Annual rainfall data for Guwahati City in India for the period from 1901 to 2002, obtained from India Meteorological Department (IMD) sources, is used in the present study. A statistical analysis of the data series exhibits a non-Gaussian kernel density. Out of the 102 data points, first 90 data points corresponding to the period 1901 to 1990 are used for model construction, while the rest 12 data points are used for testing the model. In this study, the MATLAB code for PSO developed by Omran [21] is used. This code contains the modified features as suggested by Clerc [16]. Using this code, the optimal values for the coefficients of the regression model in Eq. 5 are obtained. The best fitted model on the basis of minimum AIC and BIC is given below:

$$y_t = [0.7937 + 2.3648 \exp(-0.001 y_{t-1}^{2})]\, y_{t-1} + [0.2413 + 6.4870 \exp(-0.001\, y_{t-1}^{2}]\, y_{t-2} + \varepsilon_t \tag{6}$$

For the sake of comparison, a best fitted ARIMA model is considered. Two common statistical measures of root mean squared error (RMSE) and mean absolute error (MAE) are used for performance comparison (details are not discussed). The observed and predicted rainfall for the periods of model construction and testing are plotted in Fig 1 and Fig 2, respectively. The performance indicators are shown in Table 1. It is clear that the PSO-EXAR model performs better than the ARIMA model for modeling as well as for prediction.

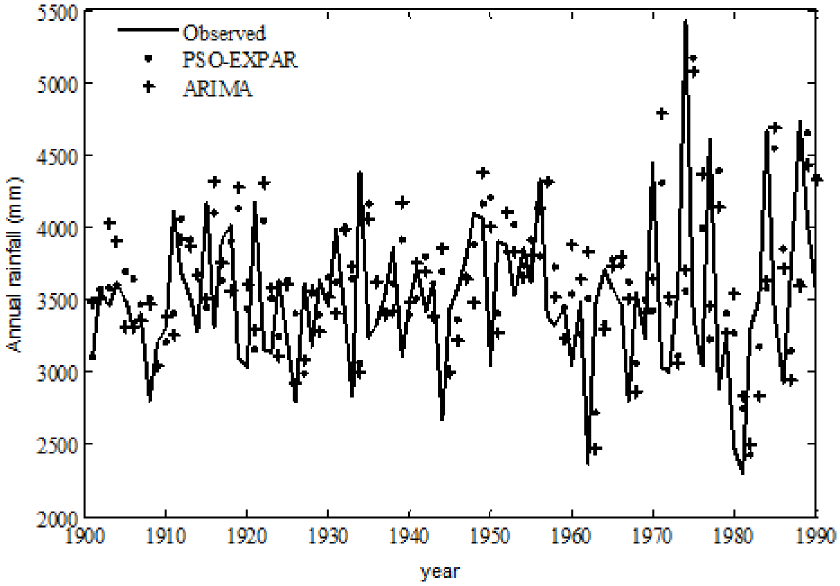

Fig 1: Linear scale plot of observed and predicted rainfall using model dataset

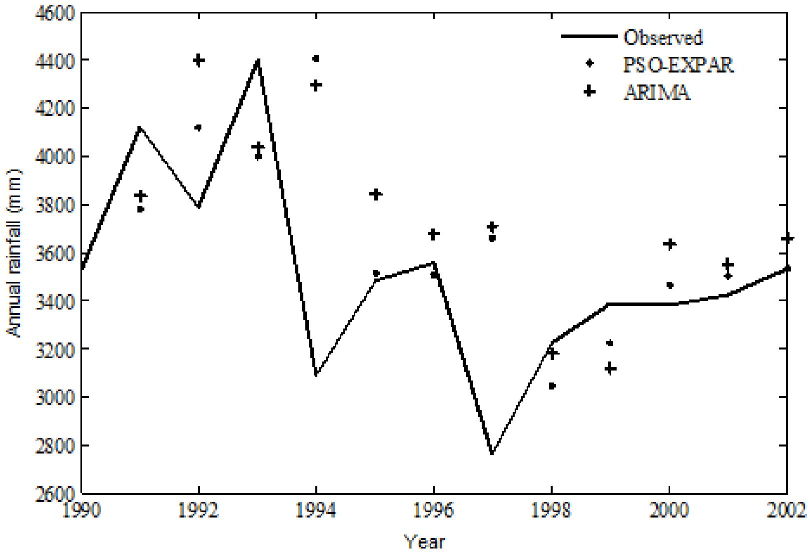

Fig 2: Linear scale plot of observed and predicted rainfall using test dataset

Table 1: Performance measures of best fitted PSO-EXPAR and ARIMA models

Criterion	Performance measures for model dataset		Performance measures for test dataset	
	PSO-EXPAR	ARIMA	PSO-EXPAR	ARIMA
RMSE	688.063	727.867	499.483	520.062
MAE	548.767	589.056	321.829	392.074
AIC	518.773	523.169	72.764	73.185
BIC	522.590	526.986	76.581	77.002

V. SUMMARY AND CONCLUSIONS

An improved Particle swarm optimization algorithm has applied in this study to optimize the parameters of an exponential regression equation for predicting annual rainfall data series. The robustness of PSO has been utilized to obtain the optimal parameters. The method has been found to be efficient and time saving. The numerical results also suggest better performance of the PSO-EXPAR model in comparison to the established statistical ARIMA model. This hybrid method may open avenues for further study in time series prediction using data mining techniques.

REFERENCES

[1] Box, G.E.P. and Jenkins, G.M. Time Series Analysis, Forecasting and Control. Holden-Day, CA, San Francisco, 1970.

[2] Leite, S. , Mand, J.P. and Peixoto, S. The autoregressive model of climatological time series an application to the longest time series in Portugal. Intern. Journ. of Climatology, 16, 1165-1173, 1996.

[3] Chander, R.E. and Wheater, H.S. Analysis of rainfall variability using generalized linear models: A case study from the west of Ireland. Water Resour. Res., 38(10), 1– 11, 2002.

[4] Shamsnia, S.A. , Amiri, S.N. and Pirmoradian, N. Drought simulation in Fars Province using standardized precipitation index and time series analysis (ARIMA Model). Intern. Journ. Appl. Math. (IJAM), 22(6), 869-878, 2009.

[5] Parrelli, R. Introduction to ARCH & GARCH models. Optional TA Handouts, Econ 472 Department of Economics, University of Illinois, 2001.

[6] French, M. N., Krajewski, W. F., and Cuykendall, R. R. Rainfall forecasting in space and time using neural network, J. Hydrol., 137, 1–31, 1992.

[7] Abrahart, R. J. and See, L. Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecast in two contrasting catchments. Hydrol. Proc., 14, 2157–2172, 2000.

[8] Hung, N. Q., Babel, M.S., Weesakul, S. and Tripathi, N.K. An artificial neural network model for rainfall forecasting in Bangkok, Thailand. Hydrol. Earth Syst. Sci., 13, 1413–1425, 2009.

[9] Samui, P., Mandla, V. R. , Krishna A. and Teja T. Prediction of Rainfall Using Support Vector Machine and Relevance Vector Machine. Earth Sc. India, 4(IV), 188 – 200, 2011.

[10] Jha, G.K., Thulasiraman, P. and Thulasiram, R.K. PSO based neural network for time series forecasting. IEEE International Joint Conference on Neural Networks (IJCNN), June 14–19, USA, 1422–1427, 2009.

[11] Adhikari, R. and Agrawal, R.K. Effectiveness of PSO based neural network for seasonal time series forecasting. Indian International Conference on Artificial Intelligence (IICAI), Tumkur, India, 232–244, 2011.

[12] Asadi, S., Tavakoli, A., and Hejazi, S.R. A new hybrid for improvement of auto-regressive integrated moving average models applying particle swarm optimization. Expert Systems with Applications , 39, 5332–5337, 2012.

[13] Cui, C. and Jiang, M. Chaotic Time Series Prediction Based On Binary Particle Swarm Optimization. AASRI Procedia, 1, 377–38, 2012.

[14] Kennedy J, and Eberhart R.C. Particle swarm optimization. In: Proceedings of the IEEE conference on neural networks. Perth: Piscataway, 4, 1942-48, 1995.

[15] Monson, C.K., and Seppi, K.D. Exposing Origin-Seeking Bias in PSO. GECCO'05, Washington, DC, USA, 241-248, 2005.

[16] Clerc, M. Standard Particle Swarm Optimisation. Particle Swarm Central, Tech. Rep., http://clerc.maurice.free.fr/pso/SPSO descriptions, 2012.

[17] Haggan, V. and Ozaki, T. Modeling nonlinear vibrations using an amplitude dependent autoregressive time series model. Biometrika, 68, 189- 196, 1981.

[18] Gurung, B. Fitting Nonlinear Time-series Model Using Swarm Optimization Technique. Advance in Electronic and Electric Engineering, 4(6), 537-540, 2014.

[19] Akaike, H. A new look at the statistical model identification. IEEE Trans. Autom. Control, AC-19,716-723, 1974.

[20] Rissanen, J. Modeling by short data description. Automation , 14, 467-471,1978.

[21] Omran, M.G.H. SPSO 2001, MATLAB. http://www.particleswarm.info/ Programs. Html, 2011.