Journal of Modern Mathematics and Statistics 5 (3): 66-70, 2011
ISSN: 1994-5388

Comparative Analysis of Rainfall Prediction Using Statistical Neural Network and Classical Linear Regression Model

C.G. Udomboso and G.N. Amahia
Department of Statistics, University of Ibadan, Ibadan, Oyo State, Nigeria

Abstract: Different types of models have been used in modeling rainfall. Since 1990s however, interest has shifted from traditional models to ANN in rainfall modeling. Many researchers found out that the ANN performed better than such traditional models. In this study, we compared a traditional linear model and ANN in the modeling of rainfall in Ibadan, Nigeria. Ibadan is a city in West Africa, located in the tropical rainforest zone, using the data obtained from the Nigeria Meteorological (NIMET) station. Three variables were considered in this study rainfall, temperature and humidity. In selecting between the two models, we concentrated on the choice of adjusted $R^2(\bar{R}^2)$, Akaike Information Criterion (AIC) and Schwarz Information Criterion (SIC). Though, the MSE and R^2 were also used, it was concluded from results that MSE is not a good choice for model selection. This is due to the nature of the rainfall data (which has wide variations). It was found that the Statistical Neural Network (SNN), generally performed better than the traditional (OLS).

Key words: Rainfall, ordinary least squares, Statistical Neural Network (SNN), model selection criteria, OLS, NIMET, Nigeria

INTRODUCTION

Rainfall models play an important role in water resource management planning and therefore, different types of models with various degrees of complexity have been developed for this purpose. Researchers have classified these models into three broad categories, regardless of their structural diversity. They are black box or system theoretical models, conceptual models and physically-based models. Black box models normally contain no physically-based input and output transfer functions and therefore are considered to be purely empirical models. Conceptual rainfall-runoff models usually incorporate interconnected physical elements with simplified forms and each element is used to represent a significant or dominant constituent hydrologic process of the rainfall-runoff transformation. Conceptual rainfall models have been widely employed in hydrological modeling.

There has been a tremendous growth in the interest of application of the Artificial Neural Networks (ANNs) in rainfall modeling sinc, the 1990s. ANNs are usually assumed to be powerful tools for functional relationship establishment or nonlinear mapping in various applications. Cannon and Whitfield (2002) found ANNs to be superior to linear regression procedures. Shamseldin (1997) examined the effectiveness of rainfall-runoff modeling with ANNs by comparing their results with the Simple Linear Model (SLM), the seasonally based Linear Perturbation Model (LPM) and the Nearest Neighbor Linear Perturbation Model (NNLPM) and concluded that ANNs could provide more accurate discharge forecasts than some of the traditional models. Presently, more and more researchers are utilizing artificial neural networks because these models possess desirable attributes of universal approximation and the ability to learn from examples. Artificial neural networks constitute a useful tool to predict and forecast various hydrological variables and are used extensively in water resources research (Tayfur, 2002; Cigizoglu, 2003a, b, 2004; Sudheer, 2005; Cigizoglu and Kisi, 2006; Toprak and Cigizoglu, 2008).

The artificial neural network models are frequently employed for rainfall forecasting (Hsu *et al*., 1997; Kuligowski and Barros, 1998; Hall *et al*., 1999; Silverman and Dracup, 2000; Applequist *et al*., 2002; Ramirez *et al*., 2005; Freiwan and Cigizoglu, 2005). French *et al*. (1992) used a neural network to forecast rainfall intensity fields in space and whilst Raman and Sunilkumar (1995) used the artificial neural network to synthesize reservoir inflow series for two sites in the Bharathapuzha basin, South India. According to Minns and Hall (1996), majority of the early study in this area have been mainly theoretical, concentrating on neural network performance with artificially generated rainfall-runoff data. ANN concept is a mapping technique. The network maps values from the input to the output by

Corresponding Author: C.G. Udomboso, Department of Statistics, University of Ibadan, Ibadan, Oyo State, Nigeria

means of a function, generally known as the transfer function. This study seeks to compare traditional linear model and ANN in the modeling of rainfall in Ibadan, Nigeria. Ibadan is a city in West Africa, located in the tropical rainforest zone. It has two distinct climatic seasons; the wet season from May to November and the dry season from December to April. The peak of the rainfall season is usually August/September. The data used in this study are daily observations obtained from the Nigeria Meteorological (NIMET) station in Ibadan for 33 years (1971-2003).

MATERIALS AND METHODS

Prediction using regression model: The regression model is basically partitioned into two parts, namely the predicted portion having the characteristic that can be ascribed to all the observations considered as a group in a parametric framework. The remaining portion called the residual is the difference between the observed and the predicted values which is ascribed to unknown sources. The model is given as:

$$y_i = f(x_j, \beta) + e_i,\ i = 1, 2, ..., n \tag{1}$$

where, n is the number of observations, y_i is the ith observation, $x_j = (x_{1i}, x_{2i}, ..., x_{ki})$ is the predictor variable vector related to y_i, $\beta = (\beta_0, \beta_1, ..., \beta_p)$ is the parameter vector and e_i is the error associated with ith observation. The matrix form of Eq. 1 is given as:

$$Y = X\beta + \varepsilon \tag{2}$$

Using least squares, the estimate of the parameter β is derived as:

$$\hat{\beta} = (X' X)^{-1} X' Y \tag{3}$$

The predicted model becomes:

$$\hat{Y} = X\hat{\beta} \tag{4}$$

so that the residual is given as:

$$e = Y - \hat{Y} \tag{5}$$

Prediction using SNN model: The Statistical Neural Network (SNN), like the regression model is composed of two parts; the predictive and the residual. It is given as:

$$y = f(X, w) + e_i \tag{6}$$

Where:

$$f(X, w) = \alpha X + \sum_{h=1}^{H} \beta_h g\left(\sum_{i=0}^{I} \gamma_{hi} x_i\right)$$

Thus, Eq. 6 can be written as:

$$y = \alpha X + \sum_{h=1}^{H} \beta_h g\left(\sum_{i=0}^{I} \gamma_{hi} x_i\right) + e_i \tag{7}$$

where, $X = (x_0, x_1, ..., x_I)$ is the vector of the input variable, g(.) is the transfer (or activation) function and $w = (\alpha, \beta, \gamma)$ are the weights (or parameters) associated with the input vector, hidden neuron and the transfer function, respectively while e_i is the error associated with the network. We note that when there is no hidden neuron, the SNN reduces to the ordinary regression model. The weights are estimated using Taylor's 1st order approximation:

$$y = y^0 + \frac{\partial f(x, w)}{\partial w}\Bigg|_{(w=w^0)}^{(w-w^0)} + e \tag{8}$$

Where:

$$y^0 = f(x, w^0)$$

if $\theta = w - w^0$ and;

$$z = \frac{\partial f(x, w)}{\partial w}$$

then, we can write Eq. 6 as:

$$y^* = z\theta + e \tag{9}$$

Where:

$$y^* = y - y^0$$

Using the Least Squares Method (LCM), the estimate of the parameter θ becomes:

$$\hat{\theta}' = (Z' Z)^{-1} Z' Y \tag{10}$$

The estimated model is;

$$\hat{y}^* = z\hat{\theta}$$

while the network error is given as:

$$e = y^* - \hat{y}^*$$

The transfer function used in this study is the symmetric saturated linear transfer function:

$$f(x) = \begin{cases} -1, & x < -1 \\ x, & -1 \le x \le 1 \\ 1, & x > 1 \end{cases}$$

The model formulation is 2-2-1, 2-5-1, 2-10-1, 2-50-1, 2-100-1.

Standardization: We standardized the input variables (converting them to the range (0, 1)) before feeding them into the network. Without this standardization, large values input into an ANN would require extremely small weighting factors to be applied. Otherwise:

- Due to inaccuracies introduced by floating point calculations on microcomputers, one should avoid using the very small weighting values that would be required
- Without using extremely small initial weights, changes made by the backpropagation would be insignificantly small and training would be very sluggish as the gradient of the transfer function at extreme values would be approximately zero. It is this gradient that is used in the adjustment of weights and biases in an ANN during training

All values leaving the network are automatically output in a standardized format. However, these output values must be destandardized to provide meaningful results. This is done by simply reversing the standardization algorithm used on the input nodes. While E-views 4 and SPSS was employed for the OLS part of the analysis, a neural code was written for the analysis of the SNN using Matlab R2009a and interesting results were obtained.

Model selection criteria: In this study, we discuss several criteria that have been used to choose between the two models. Several criteria are used for this purpose. In particular, we discuss these criteria: R^2, adjusted $R^2(\bar{R}^2)$, Akaike Information Criterion (AIC) and Schwarz Information Criterion (SIC). All these criteria aim at minimizing the residual sum of squares (SSE). However, except for the 1st criterion, criteria R^2, $R^2(\bar{R}^2)$ and SIC impose a penalty for including an increasingly large number of predictors. Thus, there is a tradeoff between goodness of fit of the model and its complexity (as judged by the number of predictors). The variables included in the study include rainfall, temperature and humidity. A close observation of the data shows that there is a wide range of variation, especially in the rainfall data. This is expected as rain does not fall every day. The data were further aggregated as monthly data. Rainfall is the predicted variable while temperature and humidity are the predictor variables. The entire data were split into three, n_1-n_3 where, $n_2 = 2\,n_1$ and $n_3 = 3n_1 = n_1+n_2$.

RESULTS AND DISCUSSION

Figure 1 compares the time plot for the variables under study. Except for temperature that seems to be stable, rainfall and humidity shows a lot of irregularity. It can be seen from the irregular pattern that there is high humidity when there is low rainfall. The pattern exhibited by the three variables is cyclical.

It is however, noticed that while the MSE for the OLS reduces as the sample increases, it is on the contrary for that of SNN (Table 1). However, generally as the hidden neuron increases, the MSE for the SNN becomes reduced. This is explained by the sensitivity of the neural network to data. Discrepancies not captured in the traditional method affects the network at very low hidden neurons. Increasing the number of hidden neuron reduces the biases in the weights. This explains the reason for the low values of the MSE in higher neurons. It is important to

Table 1: Model selection based on OLS and SNN

	OLS					SNN					
n	MSE	R^2	$\bar{R}^2$	AIC	SIC	HL	MSE	R^2	$\bar{R}^2$	AIC	SIC
132	8.00	0.25	0.237	8.368	8.934	2	3.40	0.03	0.02	3.558	3.799
						5	2.41	0.31	0.30	2.522	2.693
						10	2.60	0.26	0.25	2.721	2.905
						50	2.53	0.28	0.27	2.648	2.827
						100	2.39	0.32	0.31	2.501	2.671
264	6.78	0.28	0.275	6.933	7.220	2	5.91	0.06	0.05	6.046	6.297
						5	5.46	0.13	0.12	5.586	5.817
						10	4.44	0.29	0.29	4.542	4.730
						50	3.53	0.44	0.44	3.611	3.761
						100	0.76	0.89	0.89	0.778	0.810
396	6.70	0.29	0.289	6.795	7.003	2	9.43	0.00	-0.01	9.574	9.867
						5	8.88	0.06	0.06	9.016	9.292
						10	7.98	0.17	0.17	8.102	8.350
						50	5.41	0.43	0.43	5.493	5.661
						100	2.48	0.74	0.74	2.518	2.595

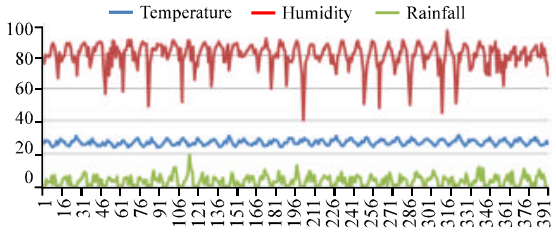

Fig. 1: Time plot for rainfall, temperature and humidity

Table 2: Model selection based on MSE and R^2

	Model selected	
n	MSE	R^2
132	2-2-1	2-5-1
	2-5-1	2-10-1
	2-10-1	2-50-1
	2-50-1	2-100-1
	2-100-1	
264	2-2-1	2-10-1
	2-5-1	2-50-1
	2-10-1	2-100-1
	2-50-1	
	2-100-1	
396	2-50-1	2-50-1
	2-100-1	2-100-1

Table 3: Model selection based on $\bar{R}^2$AIC and SIC

	Model selected		
n	$\bar{R}^2$	AIC	SIC
132	2-5-1	2-2-1	2-2-1
	2-10-1	2-5-1	2-5-1
	2-50-1	2-10-1	2-10-1
	2-100-1	2-50-1	2-50-1
		2-100-1	2-100-1
264	2-10-1	2-2-1	2-2-1
	2-50-1	2-5-1	2-5-1
	2-100-1	2-10-1	2-10-1
		2-50-1	2-50-1
		2-100-1	2-100-1
396	2-50-1	2-50-1	2-50-1
	2-100-1	2-100-1	2-100-1

note here that the MSE is not a good criteria for model selection (Table 2). The model performances of the two models (OLS and SNN) is compared using the adjusted $R^2(\bar{R}^2)$, Akaike Information Criterion (AIC) and Schwarz Information Criterion (SIC) (Table 3). The R^2 shows the performances of the individual models. We therefore, notice that at higher neurons, the better the SNN model. The AIC and SIC results show that the SNN a better model to the traditional OLS. However, SNN models for n_2 and n_3 at 100 hidden neuron was not good enough.

CONCLUSION

We have compared the ordinary least squares regression and the statistical neural network to estimate rainfall events in Ibadan, Nigeria from 1971-2003. Both methods attempt to minimize the error sum of squares between observations and predicted values. Regression requires an explicit function to be defined before the least squares parameter estimates could be computed while a neural network depends more on training data and the learning algorithm. We have restricted the variables for the models to rainfall, temperature and humidity as measured by the Nigeria Meteorological (NIMET) station in Ibadan. Comparing model prediction in both cases show that the statistical neural network performs better than the regression model. Researchers note here that if we consider other variables like sunshine hour, wind speed and solar radiation, the neural network is likely to have an added advantage over the traditional OLC.

ACKNOWLEDGEMENT

Researchers are deeply indebted to Professor Isaac Kwame Dontwi of the Department of Mathematical Sciences of the Kwame Nkrumah University of Science and Technology, Kumasi, Ghana for the scholarly review of this research.

REFERENCES

Applequist, S., G.E. Gahrs, R.L. Pfeffer and X.F. Niu, 2002. Comparison of methodologies for probabilistic quantitative precipitation forecasting. Weather Forecast., 17: 783-799.

Cannon, A.J. and P.H. Whitfield, 2002. Downscaling recent streamflow conditions in British Columbia, Canada using ensemble neural network models. J. Hydrol., 259: 136-151.

Cigizoglu, H.K. and O. Kisi, 2006. Methods to improve the neural network performance in suspended sediment estimation. J. Hydrol., 317: 221-238.

Cigizoglu, H.K., 2003a. Estimation, forecasting and extrapolation of flow data by artificial neural networks. Hydrol. Sci. J., 48: 349-361.

Cigizoglu, H.K., 2003b. Incorporation of ARMA models into flow forecasting by artificial neural networks. Environmetrics, 14: 417-427.

Cigizoglu, H.K., 2004. Estimation and forecasting of daily suspended sediment data by multi-layer perceptrons. Adv. Water Resour., 27: 185-195.

Freiwan, M. and H.K. Cigizoglu, 2005. Prediction of total monthly rainfall in Jordan using feed forward backpropagation method. Fresenious Environ. Bull., 14: 142-151.

French, M.N., W.F. Krajewski and R.R. CuykendalL, 1992. Rainfall forecasting in space and time using a neural network. J. Hydrol., 137: 1-31.

Hall, T., H.E. Brooks and C.A. Doswell, 1999. Precipitation forecasting using a neural network. Wea. Forecasting, 14: 338-345.

Hsu, K., H. Gao, S. Soroshin and H. Gupta, 1997. Precipitation estimation from remedy sensed information using artificial neural networks. J. Appl. Met., 36: 1176-1190.

Kuligowski, R.J. and A.P. Barros, 1998. Localized precipitation from a numerical weather prediction model using artificial neural networks. Weather Forecasting, 13: 1195-1204.

Minns, A.W. and M.J. Hall, 1996. Artificial neural networks as rainfall-runoff models. J. Sci. Hydrol., 41: 399-417.

Raman, H. and N. Sunilkumar, 1995. Multivariate modelling of water resources time series using artificial neural networks. Hydrol. Sci. J., 40: 145-163.

Ramirez, M.C.V., H.F. de Campos Velho and N.J. Ferreira, 2005. Artificial neural network technique for rainfall forecasting applied to the Sao Paulo region. J. Hydrol., 301: 146-160.

Shamseldin, A.Y., 1997. Application of a neural network technique to rainfall-runoff modelling. J. Hydrol., 1999: 272-294.

Silverman, D. and J.A. Dracup, 2000. Artificial neural networks and long-range precipitation prediction in California. J. Appl. Meteor., 39: 57-66.

Sudheer, K.P., 2005. Knowledge extraction from trained neural network river flow models. J. Hydrol. Eng., 10: 264-269.

Tayfur, G., 2002. Artificial neural networks for sheet sediment transport. Hydrol. Sci. J., 47: 879-892.

Toprak, F. and H.K. Cigizoglu, 2008. Predicting longitudinal dispersion coefficient in natural streams by artificial neural networks. Hydrol. Processes, 22: 4106-4129.