Global Journal of Engineering and Technology Advances
eISSN: 2582-5003
Cross Ref DOI: 10.30574/gjeta
Journal homepage: https://gjeta.com/

(RESEARCH ARTICLE)

Check for updates

Modelling an automated rainfall forecasting system using an optimized intelligent agent

Umeh Constance Nnenna [1,*], Eke James [1] and Eze Ogochukwu Edith [2]

[1] *Department of Electrical and Electronic Engineering, Enugu State University of Science and Technology, Enugu, Nigeria.*
[2] *Instrumentation Division, Centre for Basic Space Science, Nsukka, Enugu State, Nigeria.*

Global Journal of Engineering and Technology Advances, 2023, 15(01), 064–069

Publication history: Received on 09March 2023; revised on 20 April 2023; accepted on 22 April 2023

Article DOI: https://doi.org/10.30574/gjeta.2023.15.1.0077

Abstract

Weather forecasting information is very crucial in decision making process regarding to activities and works, such as in the field of agriculture to determine initial growing season. Recently, climate change causes trouble in weather forecasting. Time series data analysis for forecasting, is one of the most important aspects of the practical usage. Time Series data is large in volume, highly dimensional and continuous updating. Accurate rainfall forecasting with the help of time series data analysis had helped in the field of agriculture, in evaluating drought and flooding situations in advance. The data used within this paper is taken from Automatic Digital Meteorological Station (NECOP Station) in National Space Research and Development Agency (NASRDA-Centre for Basic Space Science, Nsukka, Enugu State Nigeria). Those data include ambient temperature, air pressure, solar radiation, relative humidity, and wind speed etc. In this paper, rainfall forecasting models were developed for Artificial Neural Network (ANN) based on Levenberg-Marquardt training function and Multiple linear regression (MLR) and was used for the prediction. Based on experimental result, it was concluded that prediction using ANN model for NASRDA-CBSS weather data produced prediction with more than 90% accuracy while MLR prediction gives 70%.

Keywords: Rainfall; Agriculture; Forecast; Artificial Neural Network (ANN); Multiple Linear Regression (MLR)

1. Introduction

Devices used by meteorologists to sample the state of the atmosphere, or what it's doing, at a given time are known as weather instruments. Rain gauge used at school, home, or offices are used to measure liquid precipitation [1]. The main important element of the hydrological process is rainfall. It is require being having advance knowledge of actual rainfall in the country like Nigeria, where most of the farmers depend on weather for their crops. Most of the states in Nigeria also suffered from the flood whereas some state is facing the problem of drought. In any of these two situations, an accurate and efficient rainfall prediction model is therefore needed. Rainfall forecasting model will help in better handling worst situations generated due to flood or drought of this kind [3-6]. Accurate rainfall forecasting will help in evaluating drought and flooding situations in advance. Therefore, it is important to have a perfect model for rainfall forecasting. An efficient and accurate rainfall prediction model is therefore needed. This kind of rainfall forecasting model will help in better handling worst situations generated due to flood or drought. Advance prediction of rainfall by this model also gives enough time to makes adequate arrangements for saving lives, transportation, procurement and supply of food and medicines. Data mining is a set of techniques used to extract unknown pieces of information from the large database repository [7-9]. There are various data mining techniques available to extract valuable and useful information from spatial, temporal, sequencing and time series data. Time Series data is a part of temporal data. Time series data generated from scientific data, financial applications, weather data, GPS, Sensor Networks etc. Large in volume, highly dimensional and continuous updating is the nature of time series data [10-12]. Use of time series data in

*Corresponding author: Umer Constance Nnenna

prediction, pattern identification, anomaly detection, motif discovery, clustering, classification, segmentation fetches the attention of data mining researchers.

2. Methodology

This research paper uses data obtained from Environmental, Pollution and Soil Monitoring (EPSm) System and with its counterpart station, Nigerian Environmental Climatic Operation Programme (NECOP) station in NASRDA-CBSS Odoru Nsukka. The EPSm system is a smart outdoor system that was designed to comprehensively carry out real time measurement of environmental and soil monitoring. The comprehensive coverage of EPSm system positions it to be a one – stop facility for a wide scenario of environmental and soil monitoring. Meteorological data mining is a form of Data mining concerned with finding hidden patterns inside largely available meteorological data, so that the information retrieved can be transformed into usable knowledge. Useful knowledge can play important role in understanding the climate variability and climate prediction. This understanding can be used to support many important sectors that are affected by climate like agriculture, water resources and tourism. To make an accurate prediction is one of the major challenges facing meteorologist all over the world. The transformation of rainfall into runoff over a catchment is a complex hydrological phenomenon, as this process is highly nonlinear, time-varying and spatially distributed. A number of models have been developed to simulate this process. Depending on the complexities involved, these models are categorized as empirical, black-box, conceptual or physically-based distributed models.

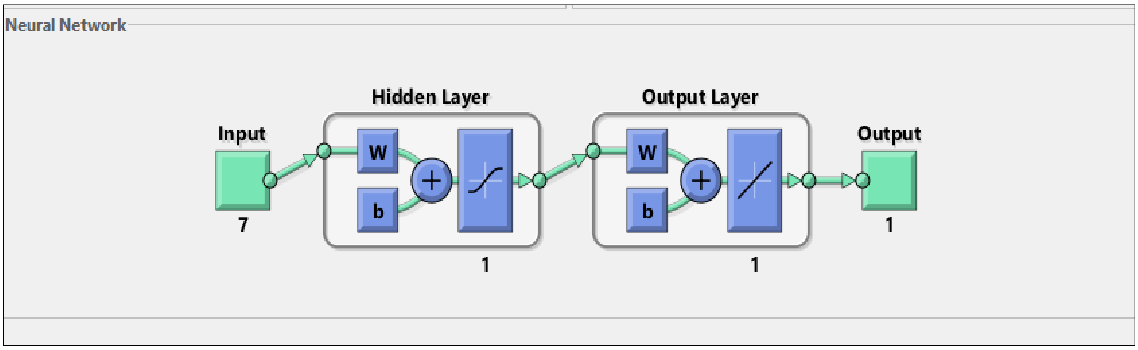

Figure 1 Output of 7 -1 – 1 Hidden layer Neural Network

3. Results and discussion

Figure2 shows the result of ANN training and the values of regression and Mean Square error of the training, validation and testing.

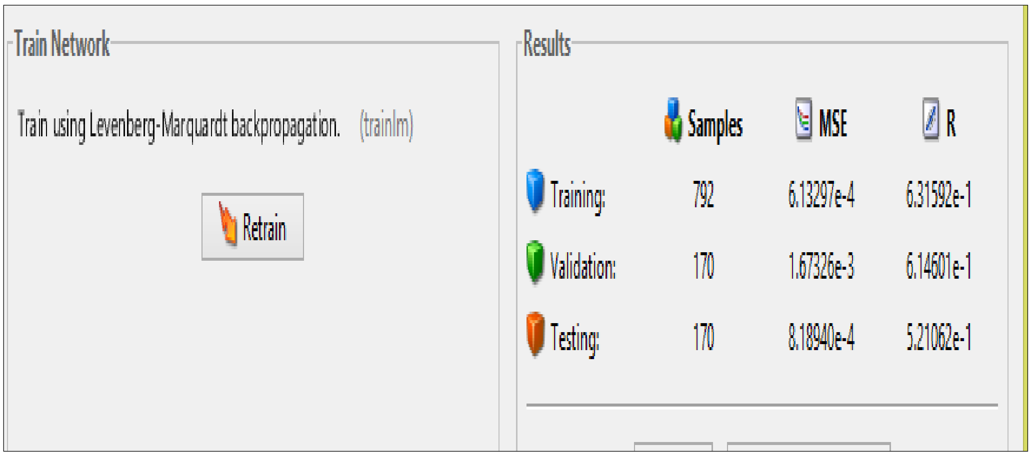

Figure 2 Result of NN training

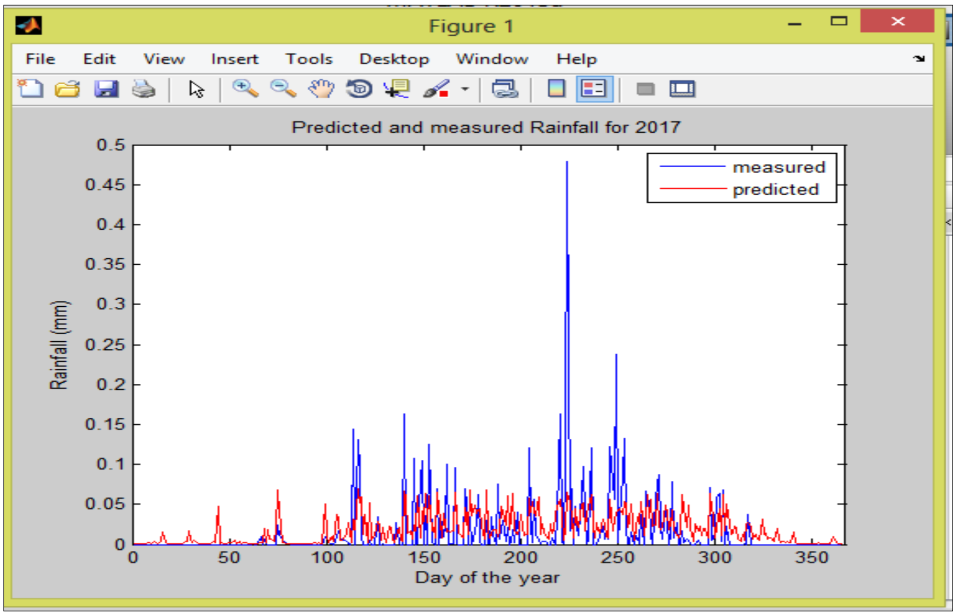

Figure 3 Predicted rainfall image

Table 1 Multiple regression model for predicting rainfall

VARIABLES	COEFFICIENTS	P - VALUE
Intercept(B0)	1.062049	5.28E-09
Atmospheric Temperature (AT)	-0.00458	1.84E-10
Relative Humidity (RH)	0.000821	2.26E-15
Wind Speed (WS)	0.011137	0.006491
Wind Direction (WD)	9.62E-05	0.001282
Atmospheric Pressure (AP)	-0.00106	3.88E-08

AT is B_1, RH is B_2, WS is B_3, WD is B_4 and AP is B_5.

Where Y is the predicted Rainfall, B_0 is the intercept or constant, AT (Atmospheric Temperature), RH (Relative Humidity), WS (Wind Speed), WD (Wind Direction), AP (Atmospheric Pressure).

From table 1, the estimates of B_0 is 1.062049 (its P value is 5.28E - 09 < 0.05), B_1 is -0.00458(P value = 1.84E-10 < 0.10), B_2 is 0.000821 (P value is 2.26E-15 < 0.10), B_3 is 0.011137 (P value is 0.006491 < 0.10), B_4 is 9.62E-05 (P value is 0.001282 < 0.10), B_5 is -0.00106 (P value is 3.88E-08 < 0.10). The multiple regression models are significant at 10% level. The model has a multiple $R^2 = 0.213$ which shows that the residuals are correlated.

Table 2 Summary of the outputs of MLR data Analysis in Excel

Regression Statistics	
Multiple R	0.461930663
R Square	0.213379937
Adjusted R Square	0.209886953
Standard Error	0.03133502
Observations	1132

Here, the square regression is 0.213 which is the same with the MLR value of the scattered graph in figure 3.6

Table 3 The intercept and coefficient values for the rainfall parameters with their Predicted values (P-Values)

	Coefficients	Standard Error	tStat	P-Value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	1.062049378	0.1805024	5.883851843	5.2807E-09	0.707890497	1.41620826	0.707890497	1.41620826
AT	-0.004584916	0.000712659	-6.43353396	1.83981E-10	-0.005983205	-0.003186627	-0.005983205	-0.003186627
RH	0.000821297	0.000102151	8.040009349	2.25732E15	0.000620868	0.001021725	0.000620868	0.001021725
WS	0.011136836	0.004083972	2.726961992	0.006491335	0.003123785	0.019149887	0.003123785	0.019149887
WD	9.62413E-05	2.98131E.05	3.228150518	0.001281887	3.77458E-05	0.000154737	3.77458E-05	0.000154737
AP	-0.001056135	0.000190822	-5.53467018	3.87722E-08	-0.001430541	-0.000681729	-0.001430541	-0.000681729

Table 4 Details of new changing values of regression parameters

Iterations	**Atmospheric Temperature(AT) coefficient**	**Atmospheric Pressure(AP) coefficient**	**Wind Speed(SP) coefficient**	**Wind Direction(WD)coefficient**	**Relative Humidity(RH) coefficient**	**IntersectValues**
First Iteration	-0.004745663	-0.001252803	0.011136836	9.62413E-05	0.000614413	0.168502902
Second Iteration	-0.00511	0.000204769	0.005103	8.53421E-04	0.000591	0.189258971
Times Performed	3	3	3	3	3	3

Table 5 Results obtained after the third iteration

	Rainfall(RF)	**Atmospheric Temperature(AT)**	**Atmospheric Pressure(AP)**	**Wind Speed(SP)**	**Wind Direction(WD)**	**Relative Humidity(RH)**
First predicted value	0.001389	4.73E-13	1.34E-11	0.995667	0.0012818	2.19E-17
Second predicted value	0.027083	2.61E-12	3.31E-09	0.174452	0.0098131	2.61E-14
Error Percent	0.001389	-7.075736707	0.000204769	0.003754994	2.98131E-05	7.65255E-05

The average error percentage from table 1 to table 5 shows that the above test is around 7%, the negative sign indicates that the error is decreased overall.

Secondly when the value of rainfall (RF) is plotted against the Predicted value the scattered graph obtained has a squared regression of 0.213 which indicates also that the average error is minimal. The Multiple Linear Regression (MLR) graph of rainfall plotted against predicted values is shown in figure 3.

Figure 4 shows the comparison result of ANN and MLR regression. When the result is compared with the regression value of MLR, we will notice that the result of ANN which is 0.58 is better than that of MLR which is 0.213.

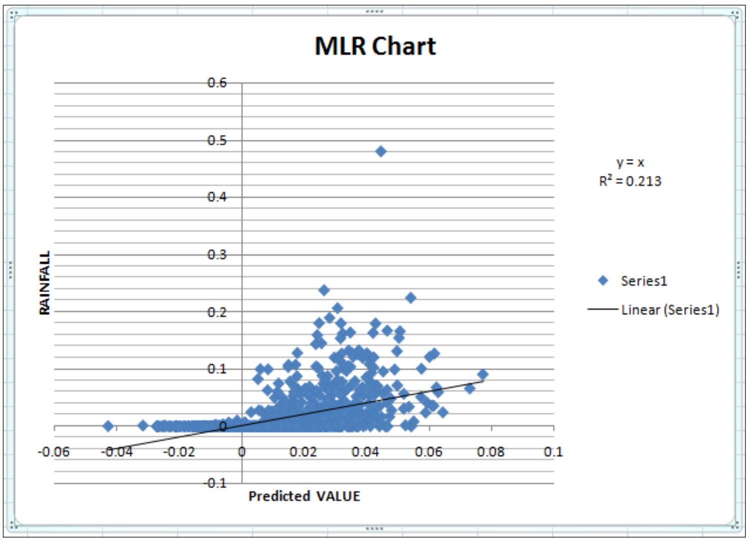

Figure 3 Scattered Multiple Linear Regression Graph

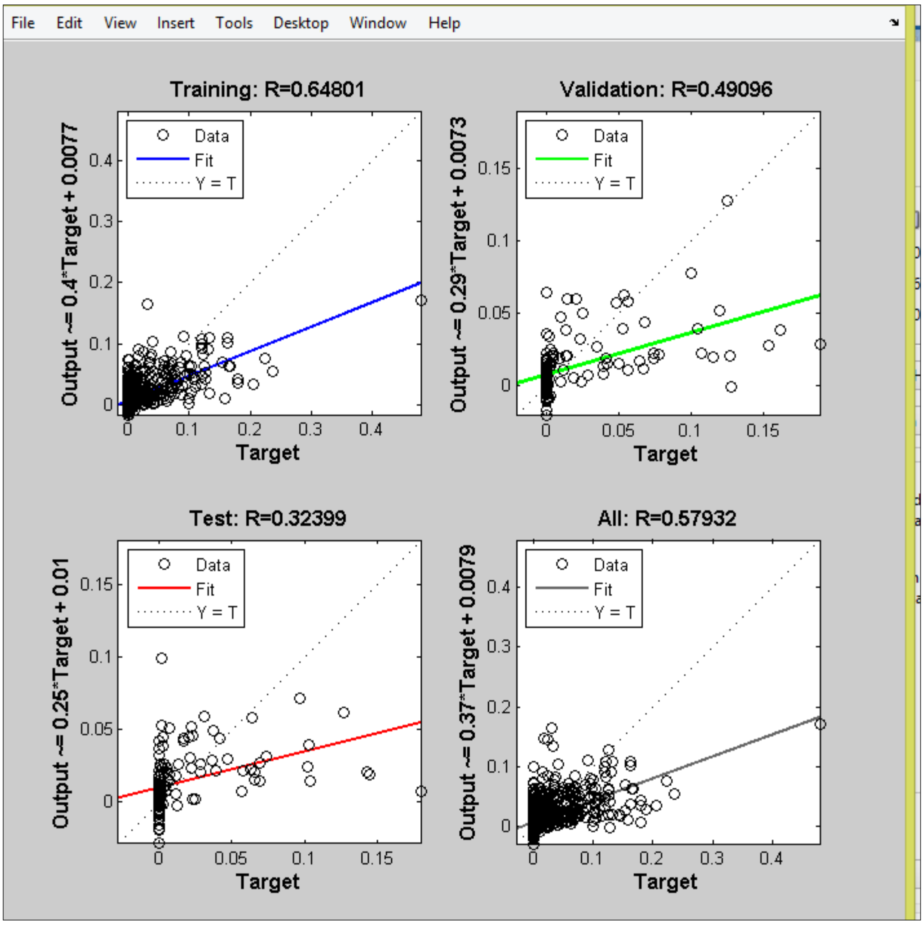

Figure 4 ANN regression plot

4. Conclusion

Rainfall is the major cause for many of the natural disasters like flash floods, droughts, tsunamis, loss of properties and lives. So in order to prevent these natural calamities, we should be able to predict the cause of the source. This can be overcome by automated rainfall forecasting system using optimized intelligent agents. The first approach is Artificial Neural network which has gained great popularity in weather prediction because of its simplicity and robustness. In this thesis, data is trained by LM algorithm. This is the fastest method among other weather forecasting methods. As there are many BP algorithm but among them Levenberg Marquardt BP is highly admired and used because of its speed and efficiency in learning. Secondly, ANN has a method which involves in training different network with different number of hidden layer neurons on the same dataset and at the end chooses the use the network that gives best performance. The second approach is multiple linear regressions which can take multiple months at a time as input and just forming a single equation which leads nearer to an accurate rainfall predicted. The problem with MLR is that too much Time is been wasted in analyzing the data and in generating the predicted values for each parameter. The two approaches can be used in other applications like, in schools to predict the average marks of their students, in sports to predict the scores or winning teams. This data is used to perform the necessary calculations to predict the rainfall from a particular district. Therefore, it is hoped that the methodology of runoff estimation using the ANN can be extended to catchments for which the gauge and discharge records are nonexistent.

Compliance with ethical standards

Acknowledgments

The success and outcome of this project were made possible by the guidance and support from my Supervisor and my team. I also want to appreciate and thank the management of Enugu State University of Science and Technology, Enugu, Nigeria for granting me the opportunity to do this program more especially their technical support.

Disclosure of conflict of interest

The authors declare that there is no conflict of interest.

References

[1] Beven, K., Asadullah, A., Bates, P., Blyth, E., Chappell, N., Child, S., Cloke, H., Dadson, S., Everard, N., Fowler, H. J.,Freer, J., Hannah, D. M., Heppell, K., Holden, J., Lamb, R., Lewis, H., Morgan, G., Parry, L. and Wagener, T. Developing observational methods to drive future hydrological science: can we make a start as a community? Hydrological Processes, 34 (3). pp. 868-873 ISSN 0885-6087, 2020.

[2] Fu Tak-chung. A review on time series data mining. Engineering Applications of Artificial Intelligence. vol. 24, pp 164–181, 2011.

[3] Poornima, S.; Pushpalatha, M. Prediction of rainfall using intensified LSTM based recurrent neural network with weighted linear units. Atmosphere 2019, 10, 668.

[4] He, R.; Zhang, L.; Chew, A.W.Z. Modeling and predicting rainfall time series using seasonal-trend decomposition and machine learning. Knowl. Based Syst. 2022, 251, 109125.

[5] Kala, A.; Vaidyanathan, S.G. Prediction of rainfall using artificial neural network. In Proceedings of the 2018 International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 11–12 July 2018; pp. 339–342.

[6] Wei, C.C.; Chou, T.H. Typhoon quantitative rainfall prediction from big data analytics by using the apache hadoop spark parallel computing framework. Atmosphere 2020, 11, 870.

[7] Samad, A.; Gautam, V.; Jain, P.; Sarkar, K. An approach for rainfall prediction using long short term memory neural network. In Proceedings of the 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA), Greater Noida, India, 30–31 October 2020; pp. 190–195.

[8] Kosko, B. Fuzzy cognitive maps. Int. J. Man Mach. Stud. 1986, 24, 65–75.

[9] Kociołek, M.; Strzelecki, M.; Obuchowicz, R. Does image normalization and intensity resolution impact texture classification? Comput. Med. Imaging Graph. 2020, 81, 101716.

[10] Behrooz, F.; Mariun, N.; Marhaban, M.H.; MohdRadzi, M.A.; Ramli, A.R. Review of control techniques for HVAC systems—Nonlinearity approaches based on Fuzzy cognitive maps. Energies 2018, 11, 495.

[11] Long, W.; Wu, T.; Xu, M.; Tang, M.; Cai, S. Parameters identification of photovoltaic models by using an enhanced adaptive butterfly optimization algorithm. Energy 2021, 229, 120750.

[12] Dong, X.; Chu, T.; Huang, T.; Ji, Z.; Wu, S. Noisy Adaptation Generates Lévy Flights in Attractor Neural Networks. Adv. Neural Inf. Process. Syst. 2021, 34, 16791–16804