International Journal of Pure and Applied Mathematics
Volume 119 No. 14 2018, 1549-1556
ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue

ijpam.eu

A HYBRID MODEL FOR RAINFALL PREDICTION USING BOTH PARAMETRIZED AND TIME SERIES MODELS

M. Swapna[a*], N. Sudhakar[b]
[a]Research Scholar, Department of CSE, Acharya Nagarjuna University, Guntur, Andhra Pradesh, India.
[b]Professor, Department of CSE, Bapatla Engineering College, Guntur, Andhra Pradesh, India.

Abstract

Prediction models of climate and rainfall are considered to be highly nonlinear and complicated, which requires classical, moderate and detailed models to obtain accurate prediction. Accurate and appropriate, area specific forecast of precipitation is very important for the countries thriving on agro based economy. Weather forecasting is an application to predict the atmospheric changes on a specific location, but sudden changes in the weather condition is a major threat for the predictions. Many tools and techniques have been introduced, still there is a need to predict accurate weather forecasting to avoid destruction of property and loss of lives. Multilayered Artificial Neural Network learning algorithm is commonly used in training. In this work, Long Short Term Memory(LSTM) Deep learning model is used to predict the forth coming rainfall in mm. In this paper, data collected during Hud-Hud from coastal Andhra Pradesh, severe cyclonic storm affected area. It also includes a hybrid model which takes time series model to predict weather Parameters like Temperature Max, Temperature Min, and Visibility etc... , and predicted parameters are tested with Parameterized prediction model which predicts next month rainfall. The results are enumerated in this paper.

Keywords: Natural Disasters, Rainfall Prediction, Deep Learning, LSTM, Time Series Prediction, Parameterized Prediction.

1. Introduction

Weather is the state of the atmosphere. Most weather occurs in the troposphere, or the lowest layer of the atmosphere. Weather is made up of multiple parameters, including air Temperature, atmospheric pressure, humidity, precipitation, solar radiation and wind. These factors can be measured to define typical weather patterns and to determine the quality of local atmospheric conditions. The aim of measuring weather is to find out the environmental conditions produced by different weather parameters has an impact on the quality of the surrounding ecosystem. Can temperature, pressure and humidity interact to form clouds and these clouds, in turn can reduce the solar radiation for plants, or increase precipitation, which can runoff into a body of water. High Temperature can increase the heat transfer to local bodies of water in addition to heating the air. Lack of precipitation affects not only the weather conditions, but also the soil moisture and water levels due to evaporation. Wind speed and direction can be indicative of a front moving into the area, or it can create waves and encourage a stratified water column to mix. Hence, Weather Forecasting is important. The implementation of Artificial Neural Network is initiated in 1964, an importance is given to Soft Computing methodology in weather forecasting [KumarAbhishek, 2012]. Both Feed-Forward Neural Networks and Recurrent Neural Networks can be used for Accurate Prediction. In the actual complex system, there are multiple variables evolving together and influencing each other, therefore multivariate prediction is more important [KumarAbhishek, 2012]. A Time Series model can be actually an integration of random and deterministic components [Pucheta, J, 2009]. If random components are eliminated then the deterministic components can then be easily modeled. Rainfall is an end product of number of complex atmospheric process which varies both in space and time. Hence time series Prediction is also important.

Most challenging part of flood forecasting is the lack of meteorological observations, particularly precipitation [Liu, J., 2012]. Precipitation Forecasts uses Numerical Weather Prediction models which still faces difficulties at

scales relevant for flood forecasting [Ebert, E. E. ,2001, Nam, D. H., 2014, Yucel, I, 2015]. Indeterminately Parameterized Prediction is important. So many models of Artificial Neural Networks have been developed for rainfall Prediction [Yucel, I, 2014]. There are two models Feed Forward Neural Network and Feedback Neural Network(Recurrent Neural Network). The main difference between Feed Forward Neural Network and Feedback Neural Networks are, in each neuron of feed-back network, the output of previous time step is forwarded as input of the next time step. This makes feed-back network be aware of time while the feed forward network has none. On Reading Multiple Papers in this domain, come to a conclusion that to predict the rainfall by using Feedback Artificial Neural Networks is a best option to obtain the results. In this paper Layer Recurrent Artificial Neural Networks are used to obtain a prediction model for the monthly rainfall of Costal Andhra Pradesh.

1.1 Failure Analysis

The atmosphere of our planet consists around 5,600 trillion tons of air. It can blast the ground below with lightning, torrential rain, heat waves, and tornadoes or carries it with a light breeze or dusting of snow flakes. When there is no modern weather forecasting tools, we need to depends on some of the observations like when bees stays in around their hives it means it is going to rain sometime that day, When we pour a cup of coffee which way the bubbles on the top go, Spider make their hub stronger when a storm is expected, If you add the chirps a cricket makes in a 14 seconds time period and add a number 40 you should come up with the temperature with one degree Fahrenheit, If all cows clustered together, a storm is brewing (if tighter cluster the worse the weather will be), Birds have a tendency to fly low because the air pressure starts falling due to an oncoming storm. etc.

By using above mentioned observation it is a complex task and we can predict but cannot be intimated prior. After that so many modern tools are introduced for weather forecasting. Some of them are Doppler Radar and High attitude balloons, Barometers measure air pressure, Anemometers etc. There are so many software tool also introduced to predict rainfall. But, there is always a need for the best predictive model for weather forecasting.

2. Study Area

Visakhapatnam (vizag) is, the financial capital of Indian State of Andhra Pradesh, India. Its population 2,035,922 making it 14th largest city in the Country. Visakhapatnam geographical location is amidst Eastern Ghats mountain range at the coast of Bay of Bengal. The city coordinates lies between 17.7041 N and 83.2977E. The City has tropical wet and dry climate . The annual mean temperature ranges between 24.7 – 30.6 ^{O} C (76-87^{O} F) with the maximum in the month of May and the minimum in the month of January. The minimum temperatures ranges between 20 – 27^{O} C (68-81^{O} F). The highest maximum temperature ever recorded was 42.0^{O} C (107.6^{O} F) in 1978, and the lowest was 20.0^{O} (68^{O} F) in 1904. It receives rainfall from the South – West and North – East monsoons and average rainfall recorded is 1,118.8 mm (44.05 in).

3. Materials and Methods

3.1 Hud-Hud Cyclone

Hud-hud cyclone the biggest disaster to have ever hit Andhra Pradesh. It was a strong tropical cyclone that caused immense loss to both the people and the government.It originated from a low pressure system with winds touching a speed of around 180-195 Kmph during landfall and waves surged as high as two to three meters .It all started with a cyclonic circulation in the Andaman sea on October 6 and slowly the furious winds got converted into a devasting cyclone on October 8.It turned out to be too severe on October 9.Hudhud had its major impact on the Jewel of East coast and The city of Destiny that is Vishakapatnam.Hudhud crossed the coast of AndhraPradesh at the noon of October12 over Visakhapatnam, with winds exceeding 185Km/ph Hudhud caused at least 120deaths within AndhraPradesh. Total damage costs estimated to be at least 10,000 crores. Despite causing extensive damage.

Fig. 1. Hudhud nearing landfall at peak strength on October 12, 2014

3.2 Artificial Neural Networks(ANNs)

Artificial Neural Networks are constructed based on human brain by making the connections using neurons and dendrites [Lee, B.W., 1998]. Each neuron is connected with Axons. Inputs from sensory organs are accepted by Dendrites. The series of electric impulses created by inputs travel along neural network. Then a neuron takes decision to transfer the signal to other neuron or block the signal to process forward. By this structure ANNs are constructed.

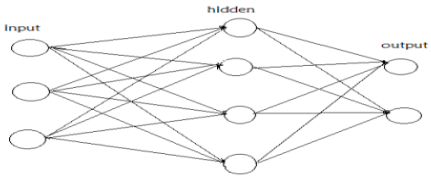

Fig. 2.Fully Connected Artificial Neural Network

There are 3 kinds of layers present in ANNs. They are Input Layer, Hidden Layer and Output Layer. Each layer consists fixed set of neurons depends on our application. There is no choice to decide how many neurons present in input and output layer because it depends on number of inputs and outputs of Data Set. But neuron selection by user is possible in hidden layer. ANN model allows us to select any number of hidden layers with any number of neurons in each hidden layer.

3.*2.1. Feedback ANN*

Here, feedback loops are allowed. They are used in content addressable memories.

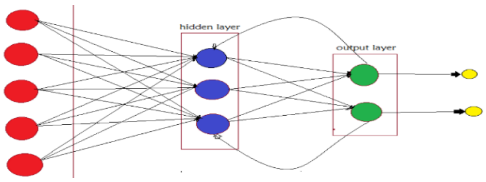

Fig. 3.FeedbackArtificial Neural Network

3.3. Deep Learning with LSTM

A Neural Network is highly structured and comes in collection of layers. First Layer is input layer, last layer is output layer, and all layers between first and last layers are hidden layers. If the network have more number of hidden layers is called a deep neural network(Deep Nets). when the data is huge a simple neural network gives some problems like vanishing gradient. To avoid these kind of problems we can go for Deep Nets. The main problem with Deep Nets are it takes much time to train the network.

3.3.1. Long Short Term Memory (LSTM)

For Data Science Industry one of the hardest problem to solve is sequence prediction problem. There is a most effective solution in deep learning for prediction problem is Long Short Term Memory (LSTM). LSTM has a special property of selectively remembering patterns for long durations of time. LSTM makes small modifications to the data by multiplications and additions. In LSTM the information flows through cell states.

There are three dependencies can be generalized to any problem

1. The previous cell state.
2. The previous hidden state.
3. The input at the current time step.
4.

Based on above three dependencies three gates are classified

1. Forget Gate (it is responsible for removing information from cell state)
2. Input Gate (it is responsible for the addition of information to the cell state)
3. Output Gate

The job of output gate is divided into 3 steps.

1. Creating a vector after applying tanh function
2. Making a filter to regulate the values using sigmoid function.
3. Multiplying the value and sending it out as a output and also to the hidden state of the next cell.

4. Prediction Models

To predict Rainfall there are two ways: (1) Time Series Prediction and (2) Parameterized Prediction

In Time Series Prediction both input and outputs are rainfall (in mm). In parameterized Prediction inputs will be parameters of weather like wind speed, humidity, pressure, Temperature, sea level etc. , output will be the

amount of rainfall in mm. Whatever the model, whether it is time series or parameterized, there are 2 phases training and testing. In training phase we need to give both input tuples and output tuples, in testing phase we need to give only input tuples and get output from the model.

Sample Data taken for both time series and parameterized Models

Table 1

Date	10/8/2014	10/9/2014	10/10/2014	10/11/2014	10/12/2014	10/13/2014
Temp Max	33	33	34	34	31.5	31
Temp Min	26	24.1	26.2	26.3	24	23
Wind	8.1	9.5	8.3	19.9	37.8	76
Pressure	1008.5	1008.7	1006.9	1002.9	996.4	971.5
Visibility	4.8	7.8	7	6.2	5.5	2.2
Rain fall	21	0	0	0	43	62

4.1. For Time series Prediction Model

if there are three inputs in a tuple one output both are amount of rainfall in mm like explained above. Suppose if , we want to find rainfall of June, 2005 the training tuple will be like following

Input(mm) **output(mm)**

May, 2005June, 2005
June, 2004
June, 2003
June, 2002
June, 2001
Testing tuple should be like following
Input (mm)
May, 2005
June, 2004
June, 2003
June, 2002
June, 2001

4.2. For Parameterized prediction Model

Training set will have tuples like following

Input **output**
Max Temperature rainfall (mm)
Min. Temperature
Wind
Pressure
Visibility
Ite. I am tryin

Testing tuple like following

Input
Max. Temperature
Min. Temperature
Wind
Pressure
Visibility

For both Time Series and Parameterized Prediction the network should be trained with data, that means we need to to give both inputs and outputs at training phase. After Training we need to give test data which contains only inputs, then trained network gives outputs. Then we need to compare outputs predicted from network compared with Actual Data. If we get very less error then we can use this model for further prediction, otherwise we need to change the network parameters like activation function, learning rate etc. After these modifications we again need to train the network, this process will be continued till we get the efficient network which generates less error.

5. Results & Discussion

By using above process, 6 data sets are prepared. On those five time series data sets for maximum temperature, minimum temperature, wind, pressure and visibility, one data set for parameterized model to predict rain fall. The network for time series is constructed to find out the next month Maximum Temperature, Minimum Temperature, wind, pressure and visibility, the network is trained and tested with the data sets, the following are test results for both time series and parameterized prediction by using Recurrent Neural Network.

5.1. Regression Graph for Maximum Temperature using Time Series Model

This network is Feedback Neural Network with 12 neurons in hidden layer, input given to this model is time series data for Maximum Temperature. The network is trained with training algorithm, after testing average error rate is -0.0053, Root Mean Square Error is 0.7371.

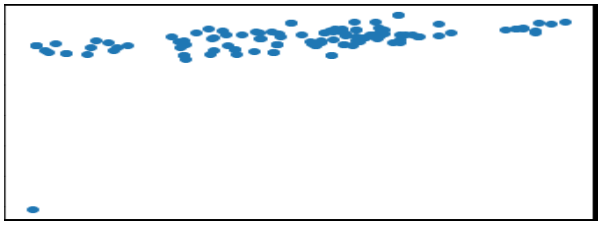

Fig. 5. Regression Graph for Time Series Maximum Temperature

5.2. Regression Graph for Minimum Temperature using Time Series Model

This network is Feedback Neural Network with 12 neurons in hidden layer, input given to this model is time series data for Minimum Temperature. The network is trained with training algorithm, after testing average error rate is -0.0887, Root Mean Square Error is 0.9553.

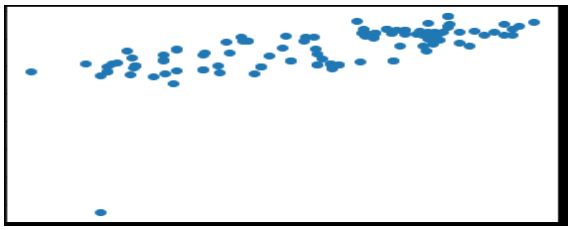

Fig. 6. Regression Graph for Time Series Minimum Temperature

5.3. Regression Graph for Wind using Time Series Model

This network is Feedback Neural Network with 12 neurons in hidden layer, input given to this model is time series data for Wind weather Parameter. The network is trained with training algorithm, after testing average error rate is -0.0359, Root Mean Square Error is 0.9292.

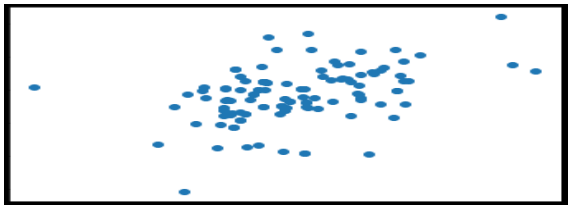

Fig. 7. Regression Graph for Time Series Wind

5.4. Regression Graph for Pressure using Time Series Model

This network is Feedback Neural Network with 12 neurons in hidden layer, input given to this model is time series data for Pressure weather Parameter. The network is trained with training algorithm, after testing average error rate is -0.0477, Root Mean Square Error is 0.6590

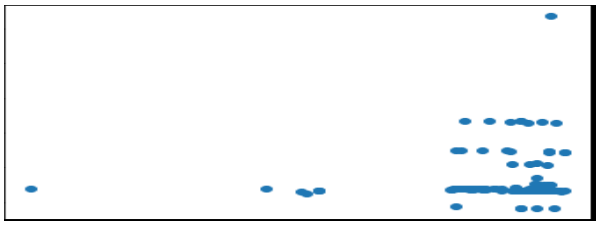

Fig. 8. Regression Graph for Time Series Pressure

5.5. Regression Graph for Visibility using Time Series Model

This network is Feedback Neural Network with 12 neurons in hidden layer, input given to this model is time series data for Visibility weather Parameter. The network is trained with training algorithm, after testing average error rate is -0.0979, Root Mean Square Error is 0.4959

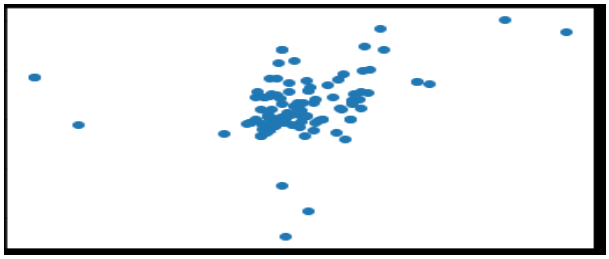

Fig. 9. Regression Graph for Time Series Visibility

5.6. Regression Graph for Rainfall using parameterized Model

This network is Feedback Neural Network with 12 neurons in hidden layer, input given to this model is parameterized data inputs as weather parameters output as rainfall. The network is trained with training algorithm, after testing average error rate is -0.1207, Root Mean Square Error is 1.2231.

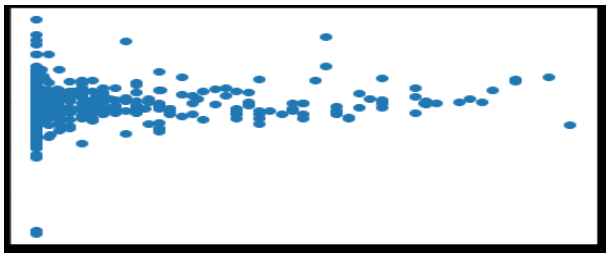

Fig. 10. Regression Graph for Parameterized model for rainfall prediction

5.7. Testing Results

Table 2

Testing Results

Weather Parameter	Actual Value	Predicted Value	Average Error	RMSE
Temperature Min	20.94516	20.6403	-0.0884	0.9553
Temperature Max	29.1871	28.4537	-0.0053	0.7371
Wind	5.125806	6.1271	-0.0359	0.9292
Pressure	1015.21	990.0418	-0.0477	0.659
Visibility	3.316129	3.663	-0.0979	0.4959
Rainfall	0.1935439	0.1149	-0.1207	1.2231

6. Conclusions

Weather forecasting using neural networks is a complex Application. A major limitation of every data analysis technique, including neural networks, is the need for a training data set that suitably represents the behavior of the system. In this paper an attempt has been made to forecast next month Average rainfall. The Limitation of Parameterized Prediction is that it can predict rainfall when actual values of weather parameters are available. By using time series we can predict next month rainfall, so we can use time series to predict weather parameter values, parameterized to predict rainfall. The model LSTM gives best results, and also noticed that when neurons in the hidden layer is increased the results are slightly better. Compared to all training algorithms LSTM is providing better results. . In results, there are some deviations from the actual rainfall. This may be due to various reasons like delay in commencement of monsoon, occurrences of cyclones etc..

The present work can be improved further by taking only cyclonic data set using same model. This study has demonstrated that multilayered Feedback neural networks trained using the LSTM algorithm can be effectively used to predict rainfall.

Acknowledgements

The authors are thankful to Dr. N.B. Venkateswarlu, professor, GVP college of Engineering for women, Visakhapatnam for his suggestions on this study. They are also thankful to Dr. V. Subramanyam, former DG, IMD, Visakhapatnam for his encouragement in this Study. They are also thankful to Dr. K L Sai Prasad, Dr. B Rajesh Babu, GVP College of Engineering for Women, Visakhapatnam for his comments to improve this paper. The contents and views are expressed in this paper are the views of authors.

References

[1] Chapter 6, Data Mining Concepts and Techniques, Han &kamber, 3rd Edition.

[2] Cohen, B., Saad, D. and Marom, E. [1997], "Efficient Training of Recurrent Neural Network with Time Delays,"Neural Networks,Vol. 10, No. 1, pp. 51-59.

[3] KumarAbhishek, Abhay Kumar, Rajeev Ranjan, Sarthak Kumar,A Rainfall Prediction Model using Artificial Neural Network, 2012 IEEE Control and System Graduate Research Colloquium (ICSGRC 2012)

[4] J. Pucheta, C. Rodriguez Rivero, Member, IEEE, M. Herrera, C. Salas and V. Sauchelli, Rainfall Forecasting Using Sub sampling Nonparametric Methods, IEEE LATIN AMERICA TRANSACTIONS, VOL. 11, NO. 1, FEB. 2013.

[5] Ebert, E. E. (2001), Ability of a poor man's ensemble to predict the probability and distribution of precipitation, Mon. Weather Rev., 129, 2461–2480.

[6] Haykin, Neural networks:A comprehensive foundation, 2nd edn. 1998, Prentice Hall, New

[7] York.

[8] http://www.cwcvsk.gov.in/ India Meteorological Department Cyclone Warning Centre,Visakhapatnam

[9] https://en.wikipedia.org/wiki/Visakhapatnam#Geography

[10] https://en.wikipedia.org/wiki/Cyclone_Hudhud

[11] http://www.fondriest.com/environmental-measurements/parameters/weather/

[12] Lee, B.W., Liu, J. and Dai, H. [1998], "Forecasting From Low Quality Data with Applications in Weather Forecasting," Inter. J. Comput.Inform., Vol. 22, No. 3, pp. 351-358.

[13] Liu, J., M. Bray, and D. Han (2012), Sensitivity of the Weather Research and Forecasting (WRF) model to downscaling ratios and storm types in rainfall simulation, Hydrol. Processes, 26, 3012–3031.

[14] Lowrey, M. R. K., and Z. Yang (2008), Assessing the capability of a regional-scale weather model to simulate extreme precipitation patterns and flooding in central Texas, Weather Forecast., 23(6), 1102–1126.

[15] Luk, K.C.,(2001). An application of artificial neural networks for rainfall forecasting, Mathematical and Computer Modeling Journal, 33, 883-699.

[16] Meert, K. and Ludik, J. [1997], "A Multilayer Real-Time Recurrent Learning Algorithm for Improved Convergence," Proc.Artificial Neural Networks - ICANN'97, pp. 505-510.

[17] M. Riedmiller, and H. Braun. "A direct adaptive method for faster backpropagation learning: The RPROP algorithm," IEEE International Conference on Neural Networks, 1993, pp. 586-591.

[18] M. Swapna, N. Sudhakar (2013),Comparison of Neural Network Configurations in the Prediction of Monthly Rainfall over Coastal Andhra Pradesh, issn: 2278-7844, Published in IJAIR, Volume 2, Issue 1 Jan 2013.

[19] M. Swapna, N. Sudhakar (2016),An Artificial Recurrent Neural Network model for Weather Forecasting,International Journal of Science & Engineering Development Research,issn: 2455-2631 Published in Volume 1 Issue 9, September-2016.

[20] Nam, D. H., D. T. Mai, K. Udo, and A. Mano (2014), Short-term flood inundation prediction using hydrologic-hydraulic models forced with downscaled rainfall from global NWP, Hydrol. Processes, 28, 5844–5859.

[21] Pucheta, J., Patino, D. and Kuchen, B. "A Statistically Dependent Approach For The Monthly Rainfall Forecast from One Point Observations". In IFIP International Federation for Information Processing Volume 294, Computer and Computing Technologies in Agriculture II, Volume 2, eds. D. Li, Z. Chunjiang, (Boston:Springer), pp. 787–798.(2009).

[22] Yucel, I., A. Onen, K. K. Yilmaz, and D. J. Gochis (2015), Calibration and evaluation of a flood forecasting system: Utility of numerical weather prediction model, data assimilation and satellite-based rainfall, J. Hydrol.,523, 49–66.

[23] Yucel, I., and A. Onen (2014), Evaluation a mesoscale atmosphere model and a satellite-based algorithm in estimating extreme rainfall events in northwestern Turkey, Nat. Hazards Earth Syst. Sci., 14, 611–624.