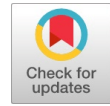


# Disaggregation of Rainfall Time Series using Artificial Neural Network In Case Of Limited Data



Saravanan Poomalai, Sivapragasam Chandrasekaran

**Abstract:** Temporal resolution of rainfall series needs to be necessarily less to use it in many engineering applications. But most of the simulated and observed rainfall series are coarser than 3 hours. Hence, it is imperative to disaggregate coarser rainfall to finer. The quantum of necessary fineness depends on application in which the rainfall data is going to be used. In this paper, the competency of Artificial Neural Network to disaggregate 3 hour rainfall into hourly, in case of limited data is verified. It is found that the disaggregation is viable with the constraint of limited data also. The rainfall is disaggregated using three models, of which, performance of the second model is much better than the others. Nonetheless the constraint of limited data, the rationale behind the better performance of the second model, is clearly discussed

**Keywords:** Artificial Neural Network, Rainfall disaggregation, Temporal resolution of rainfall

## I. INTRODUCTION

Since rainfall is directly influencing on flood and drought of a country, it is an essential meteorological parameter. Most of the simulated rainfall time series is readily available at coarser time interval (3 hourly data or larger). These coarser data are not sufficient for many hydrologic studies, for instance, preparation of Intensity-Duration-Frequency (IDF) curve. If the time interval of rainfall is finer, the IDF curve can be prepared for smaller duration of rainfall. Finer the time interval of rainfall, better the quality of IDF curve. Hence, the estimation of finer duration rainfall from coarser duration rainfall, called Rainfall disaggregation is necessary.

Many researchers disaggregated rainfall using various techniques. Sherien Fadhel et al., [1] used Nonparametric Prediction model and the method of fragments to disaggregate daily rainfall into 5-minutes rainfall. Westra et al.[2] used generalised additive model for rainfall disaggregation. Ormsbee [3] developed two empirical models for disaggregation, of which, the first one is for disaggregation for hourly rainfall into 20-min intervals and the other one is to disaggregate hourly rainfall to 30 min.

Researchers used artificial intelligence technique also for disaggregation and proved it to be superior to conventional

approach [4]. Among artificial intelligence technique, Artificial Neural Network (ANN) and Genetic Programming (GP) are widely used in disaggregation.

E. Hassanzadeh et al., [5] used GP to disaggregate projected coarser rainfall data to identify the possible changes in the IDF curve due to climate change. Burian et al., [6] attempted to reduce the temporal resolution of rainfall from hourly to finer than hourly by using ANN, where the preceding, current and successive hour rainfalls are used as input and 0,15,30,45 minute rainfall used as output. They compared their own two different ANN model with linear disaggregation method, also the method suggested by Ormsbee [3], and found ANN is superior to other methods. Golbahar Mirhosseini et al [7] tried to disaggregate rainfall using ANN. They employed 9 combinations of inputs derived from preceding, current and successive 3 hour rainfall, daily, monthly, and annual rainfall, daily maximum and minimum temperatures and elevation of the stations and 15,30,45,60 minute as output. They found best combination of inputs, using two different performance parameters

ANN is also successful tool in many water resources applications such as rainfall prediction, modelling rainfall-runoff, river stage forecasting [8, 9 and 10]

In this paper, an attempt is made to disaggregate 3 hours rainfall into hourly rainfall using ANN when the availability of data is limited. Though, other data like temperature, wind speed, humidity is available for the station, it is not used in this study. Hence, this model can be used for the station where the above mentioned data are not on hand. The predictability of ANN, with the use of cumulative rainfall alone as input, is verified and the favourable field circumstance to predict under limited data condition is discussed.

## II. STUDY AREA AND DATA SET

The study area is chosen as La Farge station in the Kickapoo River watershed, Wisconsin, USA (Figure 1). This basin is characterized by steep slopes (30 to 40 percent), rounded ridges and steep narrow valleys making it highly vulnerable to flooding. Floods on the Kickapoo River are usual. An average of one destructive flood per year is being observed. The basin consists of two rain gauges and three stream gauges. There are two rain gauges situated outside the basin. The scope of this study is restricted to develop a disaggregation model for Lafarge Rain gauge station.

Manuscript published on 30 December 2019.

\* Correspondence Author (s)

**Saravanan Poomalai**, Center for Water Technology, Department of Civil Engineering, Kalasalingam Academy of Research and Education, Srivilliputtur, India. Email: p.saravanan@klu.ac.in

**Sivapragasam Chandrasekaran** \*, Center for Water Technology, Department of Civil Engineering, Kalasalingam Academy of Research and Education, Srivilliputtur, India. Email: sivapragasam@klu.ac.in

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

# Disaggregation of Rainfall Time Series using Artificial Neural Network In Case Of Limited Data

The rainfall observations for Lafarge are obtained from <https://www.ncdc.noaa.gov/cdo-web/>. From August 1948 to August 1949 there are 104 rainfall events having duration of

3 hours or more. The data of these 104 rainfall events are consolidated for analysis in this study.

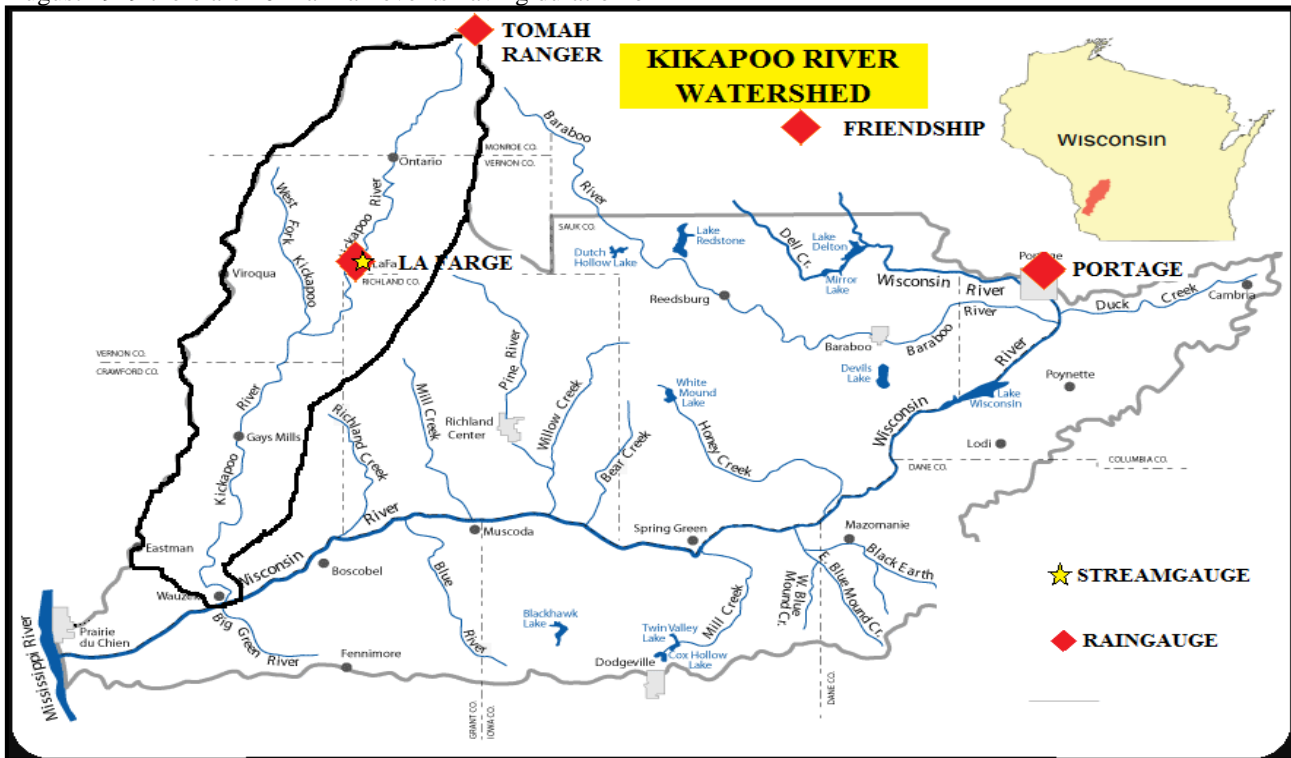


Fig. 1. Location of Study area

### III. ARTIFICIAL NEURAL NETWORK

Artificial Neural Network (ANN) is a regression tool inspired by human brain and nerve cells. Like the neurons of human brain, nodes or neurons of ANN are interconnected by links. The neurons of input layer receives external information, while the neurons of the hidden layer process the input and passes the results to the neurons of the output layer. Initially, random weights are assigned in each links. The learning of ANN happens by altering the weights of links. In the process of learning, design of architecture of ANN is first. After that, number of epochs need to be decided followed by assortment of data. Since, the assortment of number of data sets to use for training, testing and validation is highly affects the efficiency of the model, it needs to be decided wisely. The processing of input in hidden layer is depends on activation function. Hence, the choice of activation function is very important.

There are varieties of methods used in the learning process of ANN. Of which, the feed-forward back propagation algorithm is used in this study. In the learning process the accumulated errors of all the nodes are calculated and by using this calculated error, the weights in the links are adjusted to minimize the error. After many cycles of this process, an optimal architecture is framed with minimal error, which is applied to validation set later. More details of ANN can be found in [11, 12, 13, and 14].

In this paper, the performance of ANN is evaluated by the graphs shown in figures 1, 2, and 3.

### IV. ANN MODEL DEVELOPMENT

#### A. MODEL 1A : (ANNM-1A)

In this category, the rainfall in first hour interval (1HR), rainfall in second hour interval (2HR) and rainfall in third hour interval (3HR) are predicted separately by means of 3 different models using cumulative 3 hour rainfall (C3HR) as input. The first model is developed to predict 1HR, Hence the input is C3HR and output is 1HR. While, the second model is developed to predict 2HR. Here, the input is C3HR and output is 2HR. Similarly, in the third model, the input and outputs are C3HR and 3HR respectively.

A three layered feed forwarded network with 7 hidden neurons gave the best output. Out of a total of 104 rainfall events of 3 hour duration, 33 is used for training, next 33 is for testing and remaining for validation. The activation function used in this case is Logistic.

#### B. MODEL 1B : (ANNM-1B)

The only difference between ANNM-1A and ANNM-1B is activation function. The activation function used in this model is "Linear". Number of data used for training, testing and validation, inputs and outputs, hidden nodes are all similar to the ANNM-1A.

#### C. MODEL 2 : (ANNM-2)

Since, the aim is to predict the individual hourly rainfall (1HR, 2HR, 3HR) from the cumulative 3 hour rainfall (C3HR), the input is C3HR and the output is 1HR, 2HR, 3HR. Hence, there is only one model in this category.

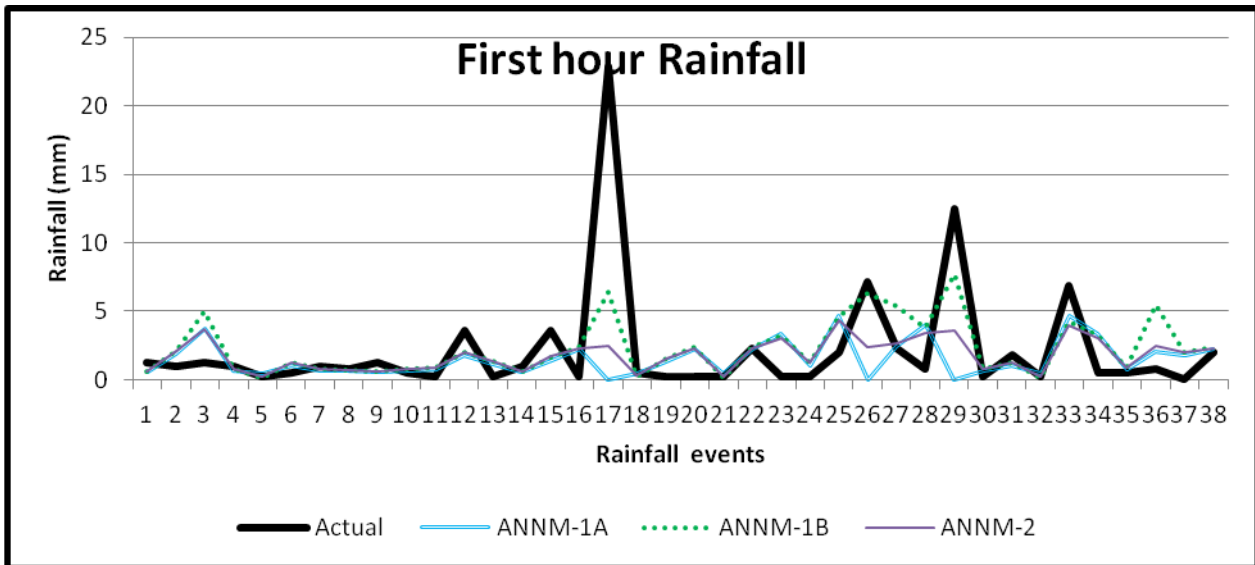


Fig. 2. Comparison of actual rainfall with the rainfall disaggregated by different ANN models for first hour rainfall

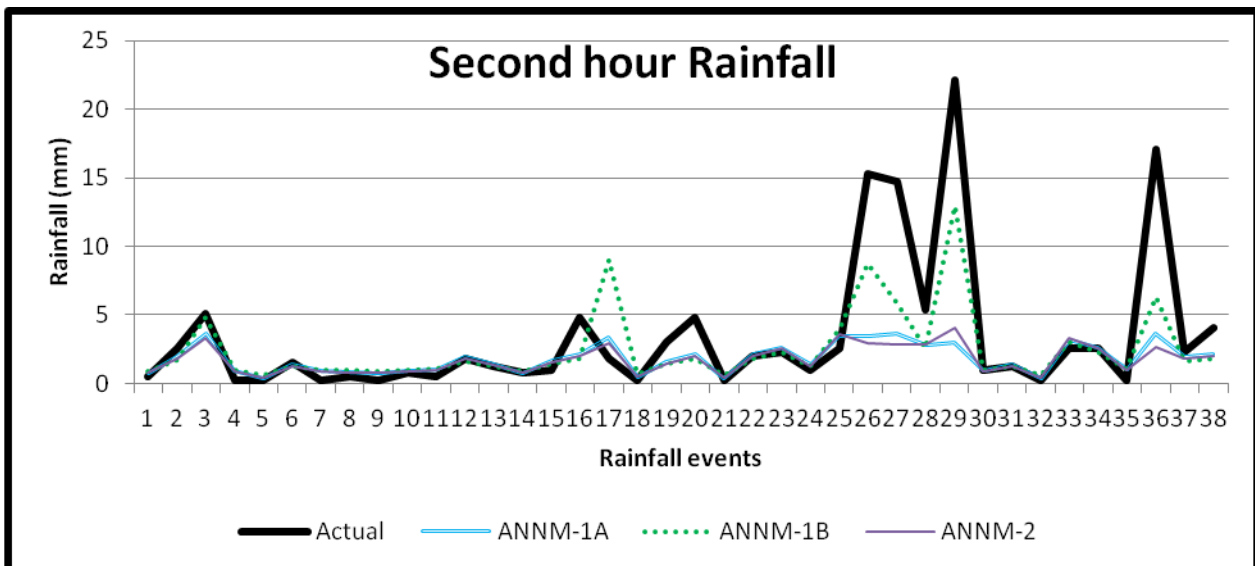


Fig. 3. Comparison of actual rainfall with the rainfall disaggregated by different ANN models for second hour rainfall

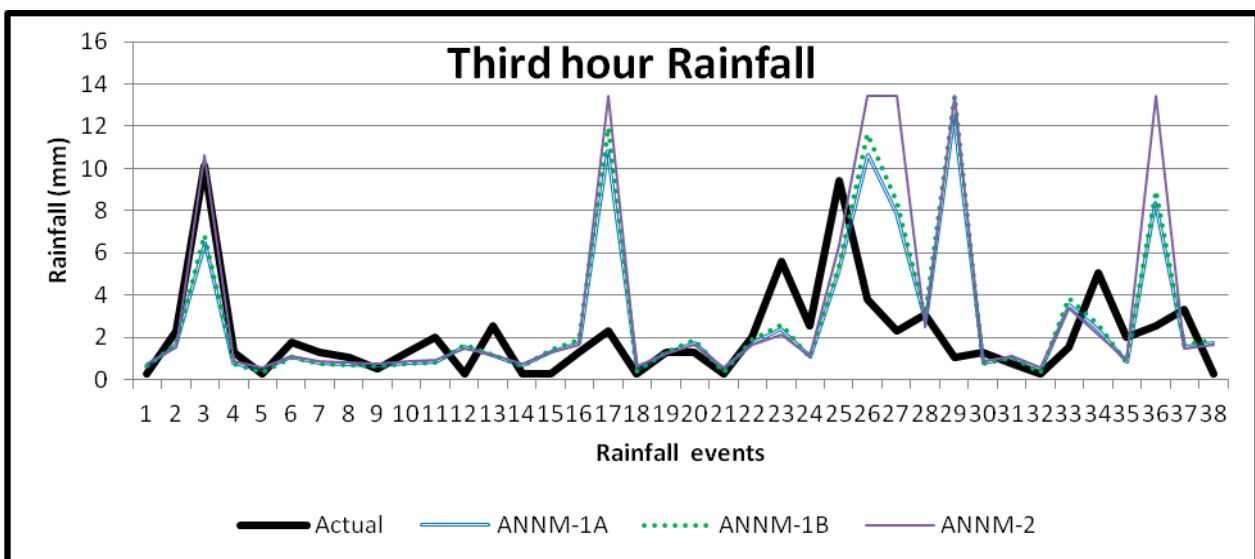


Fig. 4. Comparison of actual rainfall with the rainfall disaggregated by different ANN models for third hour rainfall

The model developed is three layer feed forward network with 7 hidden neurons. Here also first 33 events are used for training; next 33 events are used for testing and remaining 38 for validation out of 104 data. The activation function used in this category is Logistic.

## V. RESULTS AND DISCUSSIONS

In this section, the disaggregated of rainfalls by three ANN models (ANNM-1A, ANNM-1B and ANNM-2) are compared with the actual rainfall.

### A. MODEL 1A: (ANNM-1A)

By observing the figures 2, 3, 4, it is found that this model predicted the variations in the rainfall. Though the peak is underestimated in some points, some prediction is there in this model.

### B. MODEL 1B: (ANNM-1B)

In the estimation of peak this model is best among the other three models. This effect can be evidently seen in figures 1, 2 and 3. By comparison, we can say, this model mapped the rainfall with actual much better than others.

### C. MODEL 2: (ANNM-2)

This model failed to capture the rainfall variations. This shows peak when there is no peak in the actual rainfall and vice versa. The poor prediction of this model can be clearly seen from figures 1, 2 and 3.

Though, all other parameters are same, the prediction of ANNM-2 is very poor of compare to the ANNM-1A. Since the output in ANNM-2 is three, the ANN architecture is not capable to identify the pattern in the input and output. Consequently, it is decided to keep single output in the second model (ANNM-1A). Given that, it contains single output, it predicted the rainfall better than ANNM-2.

The amount of rainfall in the different interval of the rainfall may depend on temperature, wind speed, and other parameters also. If the temperature is more in a time interval more evaporation is possible during that interval, wind speed also may take away rainfall to some other location. Since, this study area is a valley, it may not affected by wind like normal terrain. Since, the altitude of the study area is much higher, the temperature is always less, consequently the variation also less. Hence the study area may not affected by temperature and wind speed. In view of this fact it is anticipated some linear relation between input and output. Hence the activation function of the ANNM-1B is kept as "linear". As expected the prediction is improved much better than the previous two models. The RMSE for ANNM-1A is 3.7 and ANNM-1B is 3.2

For successful disaggregation in limited data condition, the circumstances of the study need to be favorable. Further, if the situation of the field is not favorable, more inputs like temperature, elevation, daily rainfall etc., can be given as input to ANN to properly identify the pattern in the data.

## VI. CONCLUSION

The following conclusions can be arrived from the above mentioned discussions:

- Before prediction, the functional relation between input and output can be identified by proper observation of raw data. Then, the observed relation can be used to choose appropriate activation function to obtain better prediction.
- Though, the given input data is not enough for ANN to identify memory in the data, the prediction is possible, if the situation of the field is favorable.
- In highly random data like rainfall, if, more output is given for a single input, the ANN architecture may not be able to capture the pattern in the input and output. In such cases, less number of outputs can be given to improve prediction.

## REFERENCES

1. S. Fadhel, M.A. Rico-Ramirez, and D. Han, "Uncertainty of intensity-duration-frequency (IDF) curves due to varied climate baseline periods," *Journal of hydrology*, vol. 547, 2017, pp.600-612.
2. S. Westra, J.P. Evans, R. Mehrotra, and A. Sharma, "A conditional disaggregation algorithm for generating fine time-scale rainfall data in a warmer climate," *Journal of Hydrology*, vol. 479, 2013, pp.86-99.
3. L.E. Ormsbee, "Rainfall disaggregation model for continuous hydrologic modeling," *Journal of Hydraulic Engineering*, vol. 115(4), 1989, pp.507-525.
4. A. Weichert, and G. Bürger, "Linear versus nonlinear techniques in downscaling," *Climate Research*, vol.10(2),1998, pp.83-93.
5. E. Hassanzadeh, A. Nazemi, and A. Elshorbagy, "Quantile-based downscaling of precipitation using genetic programming: Application to IDF curves in Saskatoon". *Journal of Hydrologic Engineering*, vol.19(5),2013, pp.943-955.
6. S.J. Burian, S.R. Durrans, S. Tomić, R.L. Pimmel, and C.N. Wai, "Rainfall disaggregation using artificial neural networks," *Journal of Hydrologic Engineering*, vol. 5(3), 2000, pp.299-307.
7. G. Mirhosseini, P. Srivastava, and X. Fang, Developing rainfall intensity-duration-frequency curves for Alabama under future climate scenarios using artificial neural networks. *Journal of Hydrologic Engineering*, vol.19(11),2013, p.04014022.
8. C.Sivapragasam, P.Saravanan, S.Balamurali, S.Sumila, S.S.Keerthi, A.Farzana Farveen, "An investigation on the relationship between the Hurst exponent and the predictability of a rainfall time series". *Meteorol.Appl.*; 2019, pp.1-9. <https://doi.org/10.1002/met.1784>.
9. S. Riad, J. Mania, L. Bouchaou, and Y. Najjar, "Rainfall-runoff model using an artificial neural network approach," *Mathematical and Computer Modelling*, 40(7-8), 2004, pp.839-846.
10. K. Thirumalaiah, and M.C. Deo, "River stage forecasting using artificial neural networks". *Journal of Hydrologic Engineering*, 3(1), 1998, pp.26-32.
11. R.S. Govindaraju, "Artificial neural networks in hydrology. I: preliminary concepts," *J. Hydrol. Eng.* Vol. 5(2), 2000, pp.115-123.
12. R.S. Govindaraju, "Artificial neural networks in hydrology: II, hydrologic applications," 2000.
13. S.S. Haykin, "A *Comprehensive Foundation*". Tsinghua University Press, Beijing, 2001.
14. C. Sivapragasam, S. Vanitha, N. Muttill, K. Suganya, S.Suji, M.T. Selvi, S.J.Sudha, "Monthly flow forecast for Mississippi River basin using artificial neural networks". *Neural Comput. Appl.*, vol 24, 2014, pp. 1785-1793.

### AUTHORS PROFILE



**Dr.C.Sivapragasam**, has completed his graduation and post-graduation from IIT-Roorkee and IIT-Delhi respectively. He completed his PhD from NUS, Singapore. He has 2 years of industry experience and more than 17 years of teaching experience. He has completed 3 sponsored research projects from various funding agencies and published more than 70 papers in peer reviewed journals and conferences.



**Mr.P.Saravanan** did his Masters in Hydrology and Water Resources Engineering from College of Engineering, Guindy, Chennai. Currently he is doing PhD at Kalasalingam Academy of Research and Education. He has published 10 International Journal papers.